ECONOMICS: THEORY THROUGH APPLICATION



LibreTexts
LibreTexts

LibreTexts™

Economics: Theory Through Application

LibreTexts

This text is disseminated via the Open Education Resource (OER) LibreTexts Project (https://LibreTexts.org) and like the thousands of other texts available within this powerful platform, it is freely available for reading, printing, and "consuming."

The LibreTexts mission is to bring together students, faculty, and scholars in a collaborative effort to provide an accessible, and comprehensive platform that empowers our community to develop, curate, adapt, and adopt openly licensed resources and technologies; through these efforts we can reduce the financial burden born from traditional educational resource costs, ensuring education is more accessible for students and communities worldwide.

Most, but not all, pages in the library have licenses that may allow individuals to make changes, save, and print this book. Carefully consult the applicable license(s) before pursuing such effects. Instructors can adopt existing LibreTexts texts or Remix them to quickly build course-specific resources to meet the needs of their students. Unlike traditional textbooks, LibreTexts' web based origins allow powerful integration of advanced features and new technologies to support learning.



LibreTexts is the adaptable, user-friendly non-profit open education resource platform that educators trust for creating, customizing, and sharing accessible, interactive textbooks, adaptive homework, and ancillary materials. We collaborate with individuals and organizations to champion open education initiatives, support institutional publishing programs, drive curriculum development projects, and more.

The LibreTexts libraries are Powered by NICE CXone Expert and was supported by the Department of Education Open Textbook Pilot Project, the California Education Learning Lab, the UC Davis Office of the Provost, the UC Davis Library, the California State University Affordable Learning Solutions Program, and Merlot. This material is based upon work supported by the National Science Foundation under Grant No. 1246120, 1525057, and 1413739.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation nor the US Department of Education.

Have questions or comments? For information about adoptions or adaptions contact info@LibreTexts.org or visit our main website at https://LibreTexts.org.

This text was compiled on 06/14/2026

TABLE OF CONTENTS

Licensing

1: What Is Economics?

- 1.1: Microeconomics in a Fast-Food Restaurant
- 1.2: Macroeconomics in a Fast-Food Restaurant
- 1.3: What Is Economics, Really?
- 1.4: End-of-Chapter Material

2: Microeconomics in Action

- 2.1: Four Examples of Microeconomics
- 2.2: The Microeconomic Approach
- 2.3: End-of-Chapter Material

3: Macroeconomics in Action

- 3.1: Four Examples of Macroeconomics
- 3.2: Behind the Screens
- 3.3: Between News and Policy- The Framework of Macroeconomics
- 3.4: End-of-Chapter Material

4: Everyday Decisions

- 4.1: You and Your Choices
- 4.2: Individual Decision Making- How You Spend Your Income
- 4.3: Individual Demand
- 4.4: Individual Decision Making- How You Spend Your Time
- 4.5: End-of-Chapter Material

5: Life Decisions

- 5.1: Life Choices
- 5.2: Consumption and Saving
- 5.3: Using Discounted Present Values
- 5.4: Avoiding Risk
- 5.5: Embracing Risk
- 5.6: End-of-Chapter Material

6: eBay and craigslist

- 6.1: Buying on the Internet
- 6.2: craigslist and the Gains from Trade
- 6.3: eBay
- 6.4: Supply and Demand
- 6.5: Production Possibilities
- 6.6: End-of-Chapter Material

7: Where Do Prices Come From?

- 7.1: The Price of a Pill
- 7.2: The Goal of a Firm
- 7.3: The Revenues of a Firm
- 7.4: The Costs of a Firm
- 7.5: Markup Pricing- Combining Marginal Revenue and Marginal Cost
- 7.6: The Supply Curve of a Competitive Firm
- 7.7: End-of-Chapter Material

8: Why Do Prices Change?

- 8.1: Prices in the News
- 8.2: Market Supply and Market Demand
- 8.3: Using the Supply-and-Demand Framework
- 8.4: Another Perspective on Changing Prices
- 8.5: Three Important Markets
- 8.6: Beyond Perfect Competition
- 8.7: End-of-Chapter Material
- 8.8: Appendix- Algebraic Presentation of Supply and Demand

9: Growing Jobs

- 9.1: Changing Jobs
- 9.2: How Do Firms Decide How Many Hours of Labor to Hire?
- 9.3: Entry and Exit
- 9.4: Search
- 9.5: Government Policies
- 9.6: End-of-Chapter Material

10: Making and Losing Money on Wall Street

- 10.1: Financial Roller Coasters
- 10.2: A Walk Down Wall Street
- 10.3: The Value of an Asset
- 10.4: Asset Markets and Asset Prices
- 10.5: Efficient Markets
- 10.6: End-of-Chapter Material
- 10.7: Appendix- A General Formulation of Discounted Present Value

11: Raising the Wage Floor

- 11.1: Working at Minimum Wage
- 11.2: Nominal Wages and Real Wages
- 11.3: The Effects of a Minimum Wage
- 11.4: Minimum Wage Changes
- 11.5: The Minimum Wage and the Distribution of Income
- 11.6: Empirical Evidence on Minimum Wages
- 11.7: End-of-Chapter Material

12: Barriers to Trade and the Underground Economy

- 12.1: An Unusual Shopping List
- 12.2: How the Government Controls What You Buy and Sell
- 12.3: Limits on Trade across Borders

- 12.4: Government and the Labor Market
- 12.5: End-of-Chapter Material

13: Superstars

- 13.1: Rich and Richer
- 13.2: Facts about Inequality
- 13.3: The Sources of Inequality
- 13.4: Distributive Justice
- 13.5: Government Policy
- 13.6: End-of-Chapter Material

14: Cleaning Up the Air and Using Up the Oil

- 14.1: Dirty Travels
- 14.2: The Economics of Clean Air
- 14.3: Externalities
- 14.4: Renewable, Nonrenewable, and Accumulable Resources
- 14.5: End-of-Chapter Material
- 14.6: Appendix- An Example of the Hotelling Rule in Operation

15: Busting Up Monopolies

- 15.1: Working for Antitrust
- 15.2: Market Power and Monopoly
- 15.3: Patents and Copyright
- 15.4: Markets with a Small Number of Sellers
- 15.5: End-of-Chapter Material

16: A Healthy Economy

- 16.1: The Cost of Health Care
- 16.2: Supply and Demand in Health-Care Markets
- 16.3: Health Insurance
- 16.4: Government Policy
- 16.5: End-of-Chapter Material

17: Cars

- 17.1: Looking into Cars
- 17.2: The Demand for Automobiles
- 17.3: Supply of Cars
- 17.4: Market Outcomes in the Automobile Industry
- 17.5: Policy Issues
- 17.6: End-of-Chapter Material

18: The State of the Economy

- 18.1: The IMF Comes to Town
- 18.2: Measuring Economic Activity
- 18.3: Measuring Prices and Inflation
- 18.4: The Circular Flow of Income
- 18.5: The Meaning of Real GDP
- 18.6: End-of-Chapter Material

19: The Interconnected Economy

- 19.1: A Financial Crisis in the News
- 19.2: Housing Supply and Demand
- 19.3: Comparative Statics- Changes in the Price of Housing
- 19.4: Three Important Markets
- 19.5: Linkages across Markets
- 19.6: End-of-Chapter Material

20: Globalization and Competitiveness

- 20.1: Five Stories
- 20.2: The Production of Real GDP
- 20.3: Labor in the Aggregate Production Function
- 20.4: Physical Capital in the Aggregate Production Function
- 20.5: Other Inputs in the Aggregate Production Function
- 20.6: Accounting for Changes in GDP
- 20.7: Globalization and Competitiveness Revisited
- 20.8: End-of-Chapter Material

21: Global Prosperity and Global Poverty

- 21.1: Life Around the World
- 21.2: The Single-Person Economy
- 21.3: Four Reasons Why GDP Varies across Countries
- 21.4: The Accumulation of Physical Capital
- 21.5: Balanced Growth
- 21.6: The Role of International Institutions in Promoting Growth
- 21.7: End-of-Chapter Material

22: The Great Depression

- 22.1: Lessons from History
- 22.2: What Happened during the Great Depression?
- 22.3: The Great Depression- A Decrease in Potential Output?
- 22.4: The Components of GDP during the Great Depression
- 22.5: The Great Depression- A Decrease in Aggregate Spending?
- 22.6: Policy Interventions and the Great Depression
- 22.7: End-of-Chapter Material

23: Jobs in the Macroeconomy

- 23.1: Taking to the Streets
- 23.2: Unemployment
- 23.3: Job and Worker Flows
- 23.4: Hours Worked
- 23.5: The Government and the Labor Market
- 23.6: End-of-Chapter Material

24: Money- A User's Guide

- 24.1: The Color of Money
- 24.2: What Is Money?
- 24.3: Using Money to Buy Goods and Services
- 24.4: Using Money to Buy Other Monies- Exchange Rates

- 24.5: Using Money to Buy Assets- Interest Rates
- 24.6: End-of-Chapter Material

25: Understanding the Fed

- 25.1: Money and Power
- 25.2: Central Banks
- 25.3: The Monetary Transmission Mechanism
- 25.4: Monetary Policy, Prices, and Inflation
- 25.5: Monetary Policy in the Open Economy
- 25.6: The Tools of the Fed
- 25.7: The Fed in Action
- 25.8: End-of-Chapter Material

26: Inflations Big and Small

- 26.1: Rising Prices
- 26.2: The Quantity Theory of Money
- 26.3: Facts about Inflation and Money Growth
- 26.4: The Causes of Inflation
- 26.5: The Costs of Inflation
- 26.6: Policy Remedies
- 26.7: End-of-Chapter Material

27: Income Taxes

- 27.1: Tax Day
- 27.2: Basic Concepts of Taxation
- 27.3: The Kennedy Tax Cut of 1964
- 27.4: Income Taxes and Saving
- 27.5: The Reagan Tax Cut
- 27.6: End-of-Chapter Material

28: Social Security

- 28.1: Death and Taxes
- 28.2: Individual and Government Perspectives on Social Security
- 28.3: A Model of Consumption
- 28.4: Social Security in Crisis?
- 28.5: The Benefits and Costs of a Social Security System
- 28.6: Social Security in the Real World
- 28.7: End-of-Chapter Material

29: Balancing the Budget

- 29.1: $24,211,075,873,711.89
- 29.2: Deficits and Debt
- 29.3: The Causes of Budget Deficits
- 29.4: The Benefits of Deficits
- 29.5: The Costs of Deficits
- 29.6: The Ricardian Perspective
- 29.7: End-of-Chapter Material

30: The Global Financial Crisis

- 30.1: A World in Crisis
- 30.2: The Financial Crisis in the United States
- 30.3: From Financial Crisis to Recession
- 30.4: The Crisis in Europe and the Rest of the World
- 30.5: Currency Crises
- 30.6: End-of-Chapter Material

31: Toolkit

- 31.1: Introduction
- 31.2: Individual Demand
- 31.3: Elasticity
- 31.4: The Labor Market
- 31.5: Choices over Time
- 31.6: New Page
- 31.7: The Credit Market
- 31.8: Expected Value
- 31.9: Correcting for Inflation
- 31.10: Supply and Demand
- 31.11: Buyer Surplus and Seller Surplus
- 31.12: Efficiency and Deadweight Loss
- 31.13: Production Possibilities Frontier
- 31.14: Comparative Advantage
- 31.15: Pricing with Market Power
- 31.16: Pricing with Market Power
- 31.17: Comparative Statics
- 31.18: Production Function
- 31.19: Nash Equilibrium
- 31.20: Externalities and Public Goods
- 31.21: Foreign Exchange Market
- 31.22: Growth Rates
- 31.23: Mean and Variance
- 31.24: Correlation and Causality
- 31.25: The Credit (Loan) Market (Macro)
- 31.26: The Fisher Equation: Nominal and Real Interest Rates
- 31.27: The Aggregate Production Function
- 31.28: The Circular Flow of Income
- 31.29: Growth Accounting
- 31.30: The Solow Growth Model
- 31.31: The Aggregate Expenditure Model
- 31.32: Price Adjustment
- 31.33: Consumption and Saving
- 31.34: The Government Budget Constraint
- 31.35: The Life-Cycle Model of Consumption
- 31.36: Aggregate Supply and Aggregate Demand
- 31.37: The IS-LM Model

Index

Index

Glossary

Detailed Licensing

Licensing LicensingEdit Update titlecancel

*A detailed breakdown of this resource's licensing can be found in **Back Matter/Detailed Licensing**.*

CHAPTER OVERVIEW

1: What Is Economics?

1.1: Microeconomics in a Fast-Food Restaurant
1.2: Macroeconomics in a Fast-Food Restaurant
1.3: What Is Economics, Really?
1.4: End-of-Chapter Material

This page titled 1: What Is Economics? is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

1.1: Microeconomics in a Fast-Food Restaurant 1.1: Microeconomics in a Fast-Food RestaurantEdit Update titlecancel

Learning Objective

1. What kinds of problems do we study in microeconomics?

You watch another customer go to the counter and place an order. She purchases some fried chicken, an order of fries, and a Coca-Cola. The cost is €10. She hands over a bill and gets the food in exchange. It's a simple transaction; you have witnessed exchanges like it thousands of times before. Now, though, you think about the fact that this exchange has made both the customer and the store better off than they were previously. The customer has voluntarily given up money to get food. Presumably, she would do this only if having the food makes her happier than having the €10. KFC, meanwhile, voluntarily gave up the food to get the €10. Presumably, the managers of the store would sell the food only if they benefit from the deal as well. They are willing to give up something of value (their food) in exchange for something else of value (the customer's money).

Think for a moment about all the transactions that *could* have taken place but did not. For the same €10, the customer could have bought two orders of fried chicken. But she didn't. So even though you have never met the person, you know something about her. You know that—at this moment at least—she prefers having a Coca-Cola, fries, and one order of fried chicken to having two orders of fried chicken. You also know that she prefers having that food to any number of other things she could have bought with those euros, such as a movie theater ticket, some chocolate bars, or a book.

From your study of economics, you know that her decision reflects two different factors. The first is her tastes. Each customer likes different items on the menu. Some love the spicy fried chicken; others dislike it. There is no accounting for differences in tastes. The second is what she can afford. She has a budget in mind that limits how much she is willing to spend on fast food on a given day. Her decision about what to buy comes from the interaction between her tastes and her budget. Economists have built a rich and complicated theory of decision making from this basic idea.

You look back at the counter and to the kitchen area behind it. The kitchen, you now know, is an example of a *production process* that takes inputs and produces output. Some of the inputs are perhaps obvious, such as basic ingredients like raw chicken and cooking oil. Before you took the economics course, you might have thought only about those ingredients. Now you know that there are many more inputs to the production process, including the following:

- The building housing the restaurant
- The tables and chairs inside the room
- The people working behind the cash register and in the kitchen
- The people working at KFC headquarters managing the outlets in Paris
- The stoves, ovens, and other equipment in the kitchen used to cook the food
- The energy used to run the stoves, the ovens, the lighting, and the heat
- The recipes used to convert the ingredients into a finished product

The outputs of KFC are all the items listed on the menu. And, you realize, the restaurant provides not only the food but also an additional service, which is a place where you can eat the food. Transforming these inputs (for example, tables, chickens, people, recipes) into outputs is not easy. Let us examine one output—for example, an order of fried chicken. The production process starts with the purchase of some uncooked chicken. A cook then adds some spices to the chicken and places it in a vat of very hot oil in the huge pots in the kitchen. Once the chicken is cooked, it is placed in a box for you and served to you at the counter. That production process uses, to a greater or lesser degree, almost all the inputs of KFC. The person responsible for overseeing this transformation is the manager. Of course, she doesn't have to analyze how to do this herself; the head office provides a detailed organizational plan to help her.

KFC management decides not only what to produce and how to produce it but also how much to charge for each item. Before you took your economics course, you probably gave very little thought to where those prices on the menu came from. You look at the price again: €5 for an order of fried chicken. Just as you were able to learn some things about the customer from observing her decision, you realize that you can also learn something about KFC. You know that KFC wouldn't sell an order of fried chicken at that price unless it was able to make a profit by doing so. For example, if a piece of raw chicken cost €6, then KFC would obviously make a loss. So the price charged must be greater than the cost of producing the fried chicken.

KFC can't set the price too low, or it would lose money. It also can't set the price too high. What would happen if KFC tried to charge, say, €100 for an order of chicken? Common sense tells you that no one would buy it at that price. Now you understand that the challenge of pricing is to find a balance: KFC needs to set the price high enough to earn a good profit on each order sold but not so high that it drives away too many customers. In general, there is a trade-off: as the price increases, each piece sold brings in more revenue, but fewer pieces are sold. Managers need to understand this trade-off between price and quantity, which economists call *demand*. It depends on many things, most of which are beyond the manager's control. These include the income of potential customers, the prices charged in alternative restaurants nearby, the number of people who think that going to KFC is a cool thing to do, and so on.

The simple transaction between the customer and the restaurant was therefore the outcome of many economic choices. You can see other examples of economics as you look around you—for example, you might know that the workers earn relatively low wages; indeed, they may very well be earning minimum wage. Across the street, however, you see a very different kind of establishment: a fancy restaurant. The chef there is also preparing food for customers, but he undoubtedly earns a much higher wage than KFC cooks.

Before studying economics, you would have found it hard to explain why two cooks should earn such different amounts. Now you notice that most of the workers at KFC are young—possibly students trying to earn a few euros a month to help support them through college. They do not have years of experience, and they have not spent years studying the art of cooking. The chef across the street, however, has chosen to invest years of his life training and acquiring specialized skills and, as a result, earns a much higher wage.

The well-heeled customers leaving that restaurant are likewise much richer than those around you at KFC. You could probably eat for a week at KFC for the price of one meal at that restaurant. Again, you used to be puzzled about why there are such disparities of income and wealth in society—why some people can afford to pay €200 for one meal while others can barely afford the prices at KFC. Your study of economics has revealed that there are many causes: some people are rich because, like the skilled chef, they have abilities, education, and experience that allow them to command high wages. Others are rich because of luck, such as those born of wealthy parents.

Everything we have discussed in this section—the production process, pricing decisions, purchase decisions, and the employment and career choices of firms and workers—are examples of what we study in the part of economics called **microeconomics**. Microeconomics is about the behavior of individuals and firms. It is also about how these individuals and firms interact with each other through markets, as they do when KFC hires a worker or when a customer buys a piece of fried chicken. When you sit in a fast-food restaurant and look around you, you can see microeconomic decisions everywhere.

key takeaway

- In microeconomics, we study the decisions of individual entities, such as households and firms. We also study how households and firms interact with each other.

CHECKING YOUR UNDERSTANDING

1. List three microeconomic decisions you have made today.

This page titled 1.1: Microeconomics in a Fast-Food Restaurant is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

- **8.6: Beyond Perfect Competition** by Anonymous is licensed CC BY-NC-SA 3.0. Original source: https://2012books.lardbucket.org/books/theory-and-applications-of-economics.
- **Current page** by Anonymous is licensed CC BY-NC-SA 3.0. Original source: https://2012books.lardbucket.org/books/theory-and-applications-of-economics.

1.2: Macroeconomics in a Fast-Food Restaurant 1.2: Macroeconomics in a Fast-Food RestaurantEdit Update titlecancel

Learning Objective

1. What kinds of problems do we study in macroeconomics?

The economic decisions you witness inside Kentucky Fried Chicken (KFC) are only a few examples of the vast number of economic transactions that take place daily across the globe. People buy and sell goods and services. Firms hire and lay off workers. Governments collect taxes and spend the revenues that they receive. Banks accept deposits and make loans. When we think about the overall impact of all these choices, we move into the realm of macroeconomics. **Macroeconomics** is the study of the economy as a whole.

While sitting in KFC, you can also see macroeconomic forces at work. Inside the restaurant, some young men are sitting around talking and looking at the newspaper. It is early afternoon on a weekday, yet these individuals are not working. Like many other workers in France and around the world, they recently lost their jobs. Across the street, there are other signs that the economy is not healthy: some storefronts are boarded up because many businesses have recently been forced to close down.

You know from your economics class that the unemployed workers and closed-down businesses are the visible signs of the global downturn, or *recession*, that began around the middle of 2008. In a recession, several things typically happen. One is that the total production of goods and services in a country decreases. In many countries, the total value of all the goods and services produced was lower in 2008 than it was in 2007. A second typical feature of a recession is that some people lose their jobs, and those who don't have jobs find it more difficult to find new employment. And a third feature of most recessions is that those who do still have jobs are unlikely to see big increases in their wages or salaries. These recessionary features are interconnected. Because people have lower income and perhaps because they are nervous about the future, they tend to spend less. And because firms are finding it harder to sell their products, they are less likely to invest in building new factories. And when fewer factories are being built, there are fewer jobs available both for those who build factories and for those who work in them.

Down the street from KFC, a large construction project is visible. An old road and a nearby bridge are in the process of being replaced. The French government finances projects such as these as a way to provide more jobs and help the economy recover from the recession. The government has to finance this spending somehow. One way that governments obtain income is by taxing people. KFC customers who have jobs pay taxes on their income. KFC pays taxes on its profits. And customers pay taxes when they buy their food.

Unfortunately for the government, higher taxes mean that people and firms have less income to spend. But to help the economy out of a recession, the government would prefer people to spend more. Indeed, another response to a recession is to *reduce* taxes. In the face of the recession, the Obama administration in the United States passed a stimulus bill that both increased government spending *and* reduced taxes. Before you studied macroeconomics, this would have seemed quite mysterious. If the government is taking in less tax income, how is it able to increase spending at the same time? The answer, you now know, is that the government borrows the money. For example, to pay for the $787 billion stimulus bill, the US government issued new debt. People and institutions (such as banks), both inside and outside the United States, buy this debt—that is, they lend to the government.

There is another institution—called the monetary authority—that purchases government debt. It has specific names in different countries: in the United States, it is called the Federal Reserve Bank; in Europe, it is called the European Central Bank; in Australia, it is called the Reserve Bank of Australia; and so on. When the US government issues more debt, the Federal Reserve Bank purchases some of it. The Federal Reserve Bank has the legal authority to create new money (in effect, to print new currency) and then to use that to buy government debt. When it does so, the currency starts circulating in the economy. Similarly, decisions by the European Central Bank lead to the circulation of the euro notes and coins you saw being used to purchase fried chicken.

The decisions of the monetary authority have a big impact on the economy as well. When the European Central Bank decides to put more euros into circulation, this has the effect of reducing interest rates, which means it becomes cheaper for individuals to get a student loan or a mortgage, and it is cheaper for firms to buy new machinery and build new factories. Typically, another consequence is that the euro will become less valuable relative to other currencies, such as the US dollar. If you are planning a trip to the United States now that your class is finished, you had better hope that the European Central Bank doesn't increase the number of euros in circulation. If it does, it will be more expensive for you to buy US dollars.

Today, the world's economies are highly interconnected. People travel from country to country. Goods are shipped around the world. If you were to look at the labels on the clothing worn by the customers in KFC, you would probably find that some of the clothes were manufactured in China, perhaps some in Malaysia, some in France, some in the United States, some in Guatemala, and so on. Information also moves around the world. The customer sitting in the corner using a laptop might be in the process of transferring money from a Canadian bank account to a Hong Kong account; the person at a neighboring table using a mobile phone might be downloading an app from a web server in Illinois. This globalization brings many benefits, but it means that recessions can be global as well.

Your study of economics has taught you one more thing: the idea that you can take a trip to the United States would have seemed remarkable half a century ago. Despite the recent recession, the world is a much richer place than it was 25, or 50, or 100 years ago. Almost everyone in KFC has a mobile phone, and some people are using laptops. Had you visited a similar fast-food restaurant 25 years ago, you would not have seen people carrying computers and phones. A century ago, there was, of course, no such thing as KFC; automobiles were still a novelty; and if you cut your finger on the sharp metal edge of a table, you ran a real risk of dying from blood poisoning. Understanding why world economies have grown so spectacularly—and why not all countries have shared equally in this growth—is one of the big challenges of macroeconomics.

key takeaway

- In macroeconomics, we study the economy as a whole to understand why economies grow and why they sometimes experience recessions. We also study the effects of different kinds of government policy on the overall economy.

CHECKING YOUR UNDERSTANDING

1. If the government and the monetary authority think that the economy is growing too fast, what could they do to slow down the economy?

This page titled 1.2: Macroeconomics in a Fast-Food Restaurant is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

- **8.6: Beyond Perfect Competition** by Anonymous is licensed CC BY-NC-SA 3.0. Original source: https://2012books.lardbucket.org/books/theory-and-applications-of-economics.
- **Current page** by Anonymous is licensed CC BY-NC-SA 3.0. Original source: https://2012books.lardbucket.org/books/theory-and-applications-of-economics.

1.3: What Is Economics, Really? 1.3: What Is Economics, Really?Edit Update titlecancel

Learning Objective

1. What methods do economists use to study the world?

Economists take their inspiration from exactly the kinds of observations that we have discussed. Economists look at the world around them—from the transactions in fast-food restaurants to the policies of central banks—and try to understand how the economic world works. This means that economics is driven in large part by data. In microeconomics, we look at data on the choices made by firms and households. In macroeconomics, we have access to a lot of data gathered by governments and international agencies. Economists seek to describe and understand these data.

But economics is more than just description. Economists also build models to explain these data and make predictions about the future. The idea of a model is to capture the most important aspects of the behavior of firms (like KFC) and individuals (like you). Models are abstractions; they are not rich enough to capture all dimensions of what people do. Yet a good model, for all its simplicity, is still capable of explaining economic data.

And what do we do with this understanding? Much of economics is about policy evaluation. Suppose your national government has a proposal to undertake a certain policy—for example, to cut taxes, build a road, or increase the minimum wage. Economics gives us the tools to assess the likely effects of such actions and thus to help policymakers design good public policies.

This is not really what you thought economics was going to be about when you walked into your first class. Back then, you didn't know much about what economics was. You had a vague thought that maybe your economics class would teach you how to make money. Now you know that this is not really the point of economics. You don't have any more ideas about how to get rich than you did when you started the class. But your class has taught you something about how to make better decisions and has given you a better understanding of the world that you live in. You have started to think like an economist.

key takeaway

- Economists gather data about the world and then build models to explain those data and make predictions.

check your understanding

1. Suppose you were building a model of pricing at KFC. Which of the following factors would you want to make sure to include in your model? Which factors do you think would be irrelevant?
   - the age of the manager making the pricing decisions
   - the price of chicken
   - the number of customers who come to the store on a typical day
   - the price of apples
   - the kinds of restaurants nearby

This page titled 1.3: What Is Economics, Really? is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

- **8.6: Beyond Perfect Competition** by Anonymous is licensed CC BY-NC-SA 3.0. Original source: https://2012books.lardbucket.org/books/theory-and-applications-of-economics.
- **Current page** by Anonymous is licensed CC BY-NC-SA 3.0. Original source: https://2012books.lardbucket.org/books/theory-and-applications-of-economics.

1.4: End-of-Chapter Material 1.4: End-of-Chapter MaterialEdit Update titlecancel

In Conclusion

Economics is all around us. We all make dozens of economic decisions every day—some big, some small. Your decisions—and those of others—shape the world we live in. In this book, we will help you develop an understanding of economics by looking at examples of economics in the everyday world. Our belief is that the best way to study economics is to understand how economists think about such examples.

With this in mind, we have organized our book rather differently from most economics textbooks. It is built not around the theoretical concepts of economics but around different applications—economic illustrations as you encounter them in your own life or see them in the world around you. As you read this book, we will show you how economists analyze these illustrations, introducing you to the tools of economics as we proceed. After you have read the whole book, you will have been introduced to all the fundamental tools of economics, and you will also have seen them in action. Most of the tools are used in several different applications, thus allowing you to practice using them and gain a deeper understanding of how they work.

You can see this organization at work in our table of contents. In fact, there are two versions of the table of contents so that both students and instructors can easily see how the book is organized. The student table of contents focuses on the applications and the questions that we address in each chapter. The instructor table of contents lists the theoretical concepts introduced in each chapter so that instructors can easily see how economic theory is developed and used in the book.

We have also gathered all the tools of economics into a toolkit. You will see many links to this toolkit as you read the book. You can refer to the toolkit as needed when you want to be reminded of how a tool works, and you can also use it as a study aid when preparing for exams and quizzes.

exercises

1. A map is a model constructed by geographers and cartographers. Like an economic model, it is a simplified representation of reality. Suppose you have a map of your hometown in front of you. Think of one question about your town that you could answer using the map. Think of another question about your town for which the map would be useless.
2. Which of the following questions do you think would be studied by a macroeconomist and which by a microeconomist? (Note: we don’t expect you to be able to answer all these questions yet.)
   - What should the European Central Bank do about increasing prices in Europe?
   - What happens to the price of ice cream in the summer?
   - Should you take out a student loan to pay for college?
   - What happens when the US government cuts taxes and pays for these tax cuts by borrowing money?
   - What would happen to the prices of computers if Apple and Microsoft merged into a single firm?

Economics Detective

1. Look at a newspaper on the Internet. Find a news story about macroeconomics. How do you know that it is about macroeconomics? Find a news story about microeconomics. How do you know that it is about microeconomics?

This page titled 1.4: End-of-Chapter Material is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

2: Microeconomics in Action

2.1: Four Examples of Microeconomics

2.2: The Microeconomic Approach

2.3: End-of-Chapter Material

This page titled 2: Microeconomics in Action is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

2.1: Four Examples of Microeconomics 2.1: Four Examples of MicroeconomicsEdit Update titlecancel

Learning Objectives

1. What are two ways that you make economic choices all the time?
2. How do economists think about the way people react to a change in a rule?
3. What is the role of markets in an economy?

Here are four short and diverse illustrations of microeconomics you might encounter: deciding what to do with your time and money, buying or selling on eBay, visiting a large city, and reading about a soccer game. After you have finished your study of microeconomics, you will see these concepts very differently from the way you see them now. You may not know it, but your everyday life is filled with microeconomics in action.

Your Time and Money

Wouldn't you rather be doing something else with your time right now, instead of reading an economics textbook? You could be surfing on the Internet, reading blogs, or updating your Facebook profile. You could be reading a novel or watching television. You could be out with friends. But you aren't. You have made a choice—a decision—to spend time reading this chapter.

Your choice is an economic one. Economics studies how we cope with competing demands for our time, money, and other resources. You have only 24 hours each day, so your time is limited. Each day you have to divide up this time among the things you like or need to do: sleeping, eating, working, studying, reading, playing video games, hanging out in your local coffee shop, and so on. Every time you decide to do one thing instead of another, you have made an economic decision. As you study economics, you will learn about how you and other people make such choices, and you will also learn how to do a better job when making these decisions.

Money is also a limited resource. You undoubtedly have many things you would like to buy if money were no object. Instead you must choose among all the different things you like because your money—or, more precisely, your income—is a limited resource. Every time you buy something, be it a T-shirt, a breakfast bagel, or a new computer, you are choosing to forgo something else you could have bought instead. Again, these are economic decisions. Economics is about how you make choices. Whenever there is a limited resource—be it your time, the amount of oil reserves in the world, or tickets to the Super Bowl—and decisions to be made about how to use that resource, then economics is there to help. Indeed, the fundamental definition of economics is that it is the study of how we, as individuals and as a society, allocate our limited resources among possible alternative uses.

eBay and craigslist

Suppose you want to buy an MP3 player. There are many ways you can do this. You can go to a local store. You can look for stores on the Internet. You can also visit sites such as eBay (www.ebay.com) or craigslist (http://www.craigslist.org). eBay is an online auction site, meaning that you can look for an MP3 player and then bid against other potential buyers. The site craigslist is like an online version of the classified advertisements in a newspaper, so you can look to see if someone in your town or city is selling the player you want to buy. You can also use these sites if you want to sell something. Maybe you have some old baseball cards you want to sell. Perhaps you have a particular skill (for example, web design), and you want to sell your services. Then you can use sites such as eBay or craigslist as a seller instead of as a buyer.

We have said that economics is about deciding how to use your limited resources. It is also about how we interact with one another, and, more precisely, how we trade with one another. Adam Smith, the founder of modern economics, observed that humans are the only animal that makes bargains: "Nobody ever saw a dog make a fair and deliberate exchange of one bone for another with another dog."Adam Smith, *An Inquiry into the Nature and Causes of the Wealth of Nations* (New York: Modern Library, 1994 [1776]), 14. Barter or trade—the exchange of goods and services and money—is central to the world we live in today.

Economists often talk about trade taking place in markets. Some exchanges do literally take place in markets—such as a farmers' market where local growers bring produce to sell. Economists use the term more generally, though: a market is any institution that allows us to exchange one thing for another. Sites such as eBay and craigslist create markets in which we can transact. Normally, we exchange goods or services for money. Sometimes we exchange one good or service for another. Sometimes we exchange one type of money for another.

Most of the time, nobody forces you to buy anything, so when you give up some money in return for an MP3 player, you are presumably happier after the transaction than before. (There are some exceptions, of course. Can you think of any cases where you are forced to engage in an economic transaction?) Most of the time, nobody forces you to sell anything, so when you give up your time in return for some money, you are presumably happier after the transaction than before. Leaving aside the occasional mistake or the occasional regret, nearly every voluntary transaction makes both participants better off. Markets matter because they are a means for people to become happier.

Breathing the Air

Welcome to Mexico City! It is a wonderful place in many respects. But not in every way: from the picture you can see that Mexico City has some of the most polluted skies in the world.“Researchers to Scrutinize Megacity Pollution during Mexico City Field Campaign,” *University Corporation for Atmospheric Research*, last modified March 2, 2006, accessed January 22, 2011, www.ucar.edu/news/releases/2006/mirage.shtml.



Figure 2.1.1: The Skies of Mexico City. Source: This photo comes from the Center on Atmospheric Research, *http://www.ucar.edu*.

Mexico City was not always so polluted. Sadly, economic growth and population growth, together with the peculiarities of geography and climate, have combined to make its air quality among the worst you will encounter anywhere. Other cities around the world, from Beijing to Los Angeles, also experience significant air pollution, reducing the quality of life and bringing with it health risks and other costs.

It is hard to understand economists talking about the beauty and power of markets when you cannot breathe the air. So what is going wrong in Mexico City? Is it not full of people carrying out trades that make them better off? The problem is that transactions sometimes affect other people besides the buyer and the seller. Mexico City is full of gas stations. The owners of the gas stations are happy to sell gasoline because every transaction makes them better off. The owners of cars are happy to buy gasoline because every transaction makes them better off. But a side effect of all these transactions is that the air becomes more and more polluted.

Economics studies these kinds of problems as well. Economists seek to understand where and when markets work and where and when they don’t work. In those situations where markets let us down, economists search for ways in which economic policies can help.

Changing the Rules

We have explained that microeconomics studies choices and the benefits and problems that arise from trade. Perhaps most fundamentally, microeconomics studies how people respond to incentives. To illustrate the importance of incentives, here is an example of what can happen when they go wrong.

In February 1994, an extraordinary scene took place during a soccer match in the Caribbean. Grenada was playing Barbados, and with five minutes remaining in the match, Barbados was leading by two goals to one. As the seconds ticked away, it seemed clear that Barbados was going to win the match. Then, three minutes from the end of the game, the Barbados team did a remarkable thing. It intentionally scored an own goal, tying the game at two goals apiece.

After Grenada kicked off again, pandemonium ensued. The Grenada team tried not only to score against Barbados but also to score an own goal. Barbados desperately defended both its own goal and its opponents' goal. The spectacle on the field had very little to do with soccer as it is usually played.

To explain this remarkable sight, we must describe the tournament in which the two teams were playing. There were two groups of teams, with the winner of each group progressing to the final. The match between Barbados and Grenada was the last group game and would determine which two teams would be in the final. The results of the previous matches were such that Barbados needed to win by two goals to go to the final. If Barbados won by only one goal, then Grenada would qualify instead. But the tournament organizers had introduced an unusual rule. The organizers decided that if a game were tied, the game would go to "golden goal" overtime, meaning that the first team to score would win the game, and they had also decided that the winning team would then be awarded a two-goal victory.

As the game was drawing to a close, Barbados realized it was unlikely to get the two-goal win that it needed. The team reasoned that a tie was a better result than a one-goal victory because it gave them roughly a fifty-fifty chance of winning in extra time. So Barbados scored the deliberate own goal. Grenada, once it realized what had happened, would have been happy either winning or losing by one, so it tried to score in either goal. Barbados' strategy paid off. The game finished in a tie; Barbados scored in overtime and went on win the final.

The organizers should have consulted an economist before instituting the rules of the tournament. Economics has many lessons to teach, and among the most important is this: people respond to incentives. The change in the rules changed the incentives that the two teams faced. Because the tournament organizers had not realized that their rules could lead to a situation in which a team preferred a tie to a win, they failed to foresee the bizarre scene on the field."Football Follies," *Snopes.com*, last modified July 6, 2008, accessed January 22, 2011, http://www.snopes.com/sports/soccer/barbados.asp.

Key Takeaways

- You make economic decisions on the allocation of time by deciding how to spend each minute of the day. You make economic decisions on the allocation of your income by deciding how much to buy of various goods and services and how much to save.
- Economists study how changes in rules lead individual and firms to change their behavior. This is part of the theme in economics that incentives matter.
- Markets are one of the central ways in which individuals interact with each other. Market interactions provide a basis for the trade that occurs in an economy.

Exercises

1. When you are choosing how much time to allocate to studying, what incentives affect your decision? Does the decision depend on how much money you have? Does the decision depend on whether you have a quiz or an exam coming up in the course? If your instructor changed the rules of the course—for example, by canceling the final exam—would your choice change?
2. Instead of writing about air pollution in Mexico City, we could have written about water pollution from the 2010 oil spill in the Gulf of Mexico. Would that also be a good example of markets failing?

This page titled 2.1: Four Examples of Microeconomics is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

2.2: The Microeconomic Approach 2.2: The Microeconomic ApproachEdit Update titlecancel

Learning Objectives

1. What is the approach of microeconomics?
2. What are the big questions of economics?

There are several distinguishing features of the microeconomic approach to the world. We discuss them briefly and then conclude with a look at the big questions of economics.

Individual Choice

One element of the microeconomic approach is individual choice. Throughout this book, we explore how individuals make decisions. Economists typically suppose that individuals make choices to pursue their (broadly defined) self-interest given the incentives that they face.

We look at individuals in their roles both as members of households and as members of firms. Individuals in households buy goods and services from other households and—for the most part—firms. They also sell their labor time, mostly to firms. Managers of firms, meanwhile, make decisions in the effort to make their firms profitable. By the end of the book, we will have several frameworks for understanding the behavior of both households and firms.

Individuals look at the prices of different goods and services in the economy when deciding what to buy. They act in their own self-interest when they purchase goods and services: it would be foolish for them to buy things that they don't want. As prices change, individuals respond by changing their decisions about which products to buy. If your local sandwich store has a special on a breakfast bagel today, you are more likely to buy that sandwich. If you are contemplating buying an Android tablet computer but think it is about to be reduced in price, you will wait until the price comes down.

Just as consumers look at the prices they face, so do the managers of firms. Managers look at the wages they must pay, the costs of the raw materials they must purchase, and so on. They also look at the willingness of consumers to buy the products that they are selling. Based on all this information, they decide how much to produce and what to buy. Your breakfast bagel may be on special because the owner of your local sandwich shop got a good deal on bagels from the supplier. So the owner thinks that breakfast bagels can be particularly profitable, and to sell a lot of them, she sets a lower price than normal.

The buying and selling of a bagel may seem trivial, but similar factors apply to much bigger decisions. Potential students think about the costs and benefits of attending college relative to getting a full-time job. For some people, the best thing to do is to work full time. For others, it is better to go to school full time. Yet others choose to go to school part time and work part time as well. Presumably your own decision—whichever of these it may be—is one you made in your own best interests given your own specific situation.

From this discussion, you may think that economics is all about money, but economists recognize that much more than money matters. We care about how we spend our time. We care about the quality of the air we breathe. We care about our friends and family. We care about what others think of us. We care about our own self-image: what sort of a person am I? Such factors are harder to measure and quantify, but they all play a role in the decisions we make.

Markets

A second element of microeconomics has to do with how individual choices are interconnected. Economics is partly about how we make decisions as individuals and partly about how we interact with one another. Most importantly—but not exclusively—economics looks at how people interact by purchasing and selling goods and services.

In a typical transaction, one person (the buyer) hands over money to another (the seller). In return, the seller delivers something (a good or a service) to the buyer. For example, if you buy a chocolate bar for a dollar, then a dollar bill goes from your hands to those of the seller, and a chocolate bar goes from the seller to you. At the level of an individual transaction, this sounds simple enough. But the devil is in the details. In any given (potential) transaction, we can ask the following questions:

- How many? Will you buy 1, 2, or 10 chocolate bars? Or will you buy 0—that is, will the transaction take place at all?
- How much? How much money does the buyer give to the seller? In other words, what is the price?

You will see in different chapters of this book that the answers to these questions depend on exactly how buyers and sellers interact. We get a different answer depending on whether there are many sellers or only a few. We get a different answer if the good is sold at a retail store or at an auction. We get a different answer if buyers and sellers can or cannot negotiate. The exact way in which people exchange goods and services matters a great deal for the *how many?* and *how much?* questions and thus for the gains from trade in the economy.

The Role of Government

We have pointed out that individuals acting in their own self-interest benefit from voluntary trade. If you are not forced to buy or sell, then there is a presumption that every transaction makes the participants happier. What is more, markets are often a very effective institution for allowing people to meet and trade with one another. In fact, there is a remarkable result in economics that—under some circumstances—individuals acting in their own self-interest and trading in markets can manage to obtain *all* the possible benefits that can come from trading. Every transaction carried out is for the good, and every good transaction is carried out. From this comes a powerful recommendation: do whatever is possible to encourage trade. The phrase *under some circumstances* is not a minor footnote. In the real world, transactions often affect people other than the buyer and the seller, as we saw in our example of gas stations in Mexico City. In other cases, there can be problems with the way that markets operate. If there is only a small number of firms in a market, then managers may be able to set high prices, even if it means that people miss out on some of the benefits of trade. Later in this book, we study exactly how managers make these decisions. The microeconomic arguments for government intervention in the economy stem from these kinds of problems with markets. In many chapters, we discuss how governments intervene in an attempt to improve the outcome that markets give us. Yet it is often unclear whether and how governments should be involved. Pollution in Mexico City illustrates how complex these problems can be. First, who is responsible for the pollution? Some of it comes from people and firms outside the city and perhaps even outside the country. If pollution in Mexico City is in part caused by factories in Texas, who should deal with the problem: the Mexico City government, the Mexican government, the US government, or the Texas state legislature? Second, how much pollution should we tolerate? We could shut down all factories and ban all cars, but few people would think this is a sensible policy. Third, what measures can we use to combat air pollution? Should we simply place limits on production by firms and the amount of driving? Should we use some kind of tax? Is there a way in which we can take advantage of our belief that people, including the managers of firms, respond to incentives?

There are two traps that we must avoid. The first is to believe that markets are the solution to everything. There is no imaginable market in which the residents of Mexico City can trade with the buyers and sellers of gasoline to purchase the right amount of clean air. The second trap is to believe that the government can fix every market failure. Governments are collections of individuals who respond to their own incentives. They can sometimes make things better, but they can sometimes make things worse as well.

There is room for lots of disagreement in the middle. Some economists think that problems with markets are pervasive and that government can do a great deal to fix these problems. Others think that such problems are rare and that governmental intervention often does more harm than good. These disagreements result partly from different interpretations of the evidence and partly from differences in politics. Economists are as prone as everyone else to view the world through their own ideological lens. As we proceed, we do our best to present the arguments on controversial issues and help you understand why even economists sometimes come to differing conclusions about economic policy.

Incentives

Perhaps our story of the Barbados-Grenada soccer game did not seem related to economics. Economists believe, though, that the decisions we make reflect the incentives we face. Behavior that seems strange—such as deliberately scoring an own goal in a soccer game—can make perfect sense once you understand the underlying incentives. In the economic world, it is often governments that make the rules of the game; like the organizers of soccer tournaments, governments need to be careful about how the rules they set can change people's behavior.

Here is an example. In some European countries, laws are in place that give a lot of protection to workers and keep them from being unfairly fired by their employers. The intentions of these laws are good; some of their consequences are not so beneficial. The laws also make firms more reluctant to *hire* workers because they are worried about being stuck with an unsuitable employee. Thus these laws probably contribute to higher unemployment.

Incentives affect all transactions. When you buy a breakfast bagel on sale, both you and the owner of the sandwich shop are responding to the incentives that you face. The owner responds to the lower price of bagels. You respond to the lower price of the

sandwich. Economists think that we can understand a great deal about people's behavior if we have a good understanding of the incentives that they face.

Notice that not everyone makes the same choices. There are two main reasons for this:

- People have different desires or tastes. Some people like bagels; others hate them. Some people like being students; others would prefer to work rather than study.
- People have different incentives. Some people face very different job prospects and thus make different decisions about schooling. If you have this great idea for a new web product (for example, the next Google or Facebook), then you might be wise to spend your time on this project instead of studying.

The Big Questions of Economics

To conclude our introduction to microeconomics, let us look at the big picture of what happens in an economy. An economy possesses some *resources*. These include the time and abilities of the people who live in the economy, as well as natural resources such as land, mineral deposits, and so on. An economy also possesses some *technologies*. A technology is a means of changing, or transforming, one set of things into other things. For example, we have a technology for making tea. This technology takes cold water, energy, and dried leaves and transforms them into a hot beverage. Finally, an economy, of course, contains its people, and these people like to consume things. Economics studies all aspects of this process. It considers the following:

- **What** goods and services are produced in an economy?
- **How** are these goods and services produced?
- **Who** gets to consume these goods and services?

These questions concern the *allocation of resources*.

The *what* in the first question reflects the choice among the multitude of goods and services an economy could produce. Think for a moment about the clothes you are wearing right now, the food you have eaten today, and the activities you undertake during a typical day. Someone made those clothes; someone prepared that food. Somehow, society must decide how much of each type of good and service to produce.

The *how* in the second question reflects competing ways to produce goods and services. Take a basic commodity such as rice. A large amount of rice is produced in the United States on large-scale, mechanized farms. A large amount of rice is also produced in Vietnam, but the production methods are very different. In Vietnam, people do much more work manually rather than by machine. A big part of the *how* question is deciding what mix of resources and what technologies should be used to produce goods and services. The answer in a rich country such as the United States is frequently different from the answer in a poor country such as Vietnam. Indeed, the answer may be different in different states in the United States or in the same place at different times.

The *who* in the third question concerns the distribution of goods and services in the economy. Suppose you were responsible for the distribution of all goods and services to your family. If there are 4 people in your family and each consumed 50 products in a typical day, you would have to make about 200 allocation decisions each day. It would be a very hard task. Yet somehow the economies of the world allocate billions of products to billions of people.

These three questions are answered in the world partly through individual decisions. The way in which you allocate your time each day is part of the allocation of resources in the economy. If each of us lived alone, engaging in subsistence farming and not interacting with others, then we would each determine our own allocation of resources. Because we interact with others, however, these questions are also answered in part by the way in which society is organized. Most of us produce only a few goods but consume many. We *specialize in production* and *generalize in consumption*. To do so, we must exchange what we produce with others. Most of these exchanges take place as a result of individual decisions in different kinds of markets. It is the operation of these countless markets that determines the allocation of goods and services in the economy. Remarkably, these markets somehow coordinate the decisions of the billions of people in the world economy.

Some of these exchanges are controlled by the government. In some economies, the government plays a very active role; in others, it intervenes less. When a government makes decisions about the allocation of resources, this is another mechanism in the production of goods and the distribution to individuals.

Key Takeaways

- The approach of microeconomics starts with the decisions of an individual about the allocation of time and income. The impact of incentives on individual choices is a key part of economics. The approach of microeconomics then looks at the interactions of individuals directly and in markets.
- Economics answers the questions of what goods and services are produced, how they are produced, and who consumes them.

check your understanding

1. We said that most people *specialize in production* and *generalize in consumption*. What goods or services (if any) do you produce? What are the most important goods and services that you consume?
2. Police protection is a service provided by most governments. What are the *what*, *how*, and *who* aspects of the provision of this service?

This page titled 2.2: The Microeconomic Approach is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

2.3: End-of-Chapter Material 2.3: End-of-Chapter MaterialEdit Update titlecancel

In Conclusion

Our book is built around economic topics. Examples of these topics include the decisions you make in your everyday life, auctions such as those you see on eBay, whether you can make money on Wall Street, where jobs come from, and health care. As we introduce and discuss these applications, we remain keenly aware of the key themes in microeconomics: individuals responding to incentives, markets as the basis for interactions among firms and households, and the role of government intervention.

Throughout this book, we emphasize the measurement and interpretation of economic data. Understanding how to read charts and tables of economic data is a critical skill for anyone who wants to be a sophisticated consumer of economic and political news.

Mastering microeconomics involves both understanding the tools that microeconomists use and knowing how and when those tools should be applied. In this book, you will learn about these tools by example; you will see them in use as we study different questions in economics. At the same time, you will learn about many topics that should interest you as engaged and aware citizens of the world. We hope that, after reading this book, you will both better understand what it is that economists do and be better informed about the world in which we all live.

There is a considerable amount of core material in microeconomics that we use repeatedly as we tackle different problems. We highlight these core elements in the chapters and also gather them together in the toolkit. You can read any and every chapter in the book without necessarily having to refer to the toolkit, but you may often find it to be a helpful reference.

exercises

1. Think about the last item of clothing you bought for yourself. How much did you spend on it? List three other things that you like and could have bought with (approximately) the same amount of money. Why did you decide to buy the clothing rather than one of the things you just listed?
2. How have you spent the previous 24 hours? How much time did you spend sleeping? How much time did you spend working? What else could you have done with your time? Why are you reading this chapter instead of doing something else with your time?
3. Think about a game or sport that you enjoy. What rule of that game could be changed? How would this change in the rules affect the way in which the players behave?
4. When we discussed individual choice, we talked mainly about the choices of an individual person. However, in economics we often talk about the choice of a household consisting of two or more people. In what ways are the choices of a household different from the choices of an individual? In what ways are they similar?
5. Can you think of examples of economic choices that are made by the government?

Economics Detective

1. We explained the social problem of air pollution in Mexico as a situation where markets have failed to bring about good outcomes. Instead of writing about pollution, we could have written about other social problems, such as crime, illiteracy, or obesity. Browse the Internet to find another example of a social problem—either from this list or something else that interests you. Write one paragraph that explains the problem and another that discusses if and how the government might solve the problem.

This page titled 2.3: End-of-Chapter Material is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

3: Macroeconomics in Action

3.1: Four Examples of Macroeconomics
3.2: Behind the Screens
3.3: Between News and Policy- The Framework of Macroeconomics
3.4: End-of-Chapter Material

This page titled 3: Macroeconomics in Action is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

3.1: Four Examples of Macroeconomics 3.1: Four Examples of MacroeconomicsEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. How might you encounter macroeconomics?
2. What are the main indicators of the macroeconomy?
3. What are the primary macroeconomic policy tools of the government?

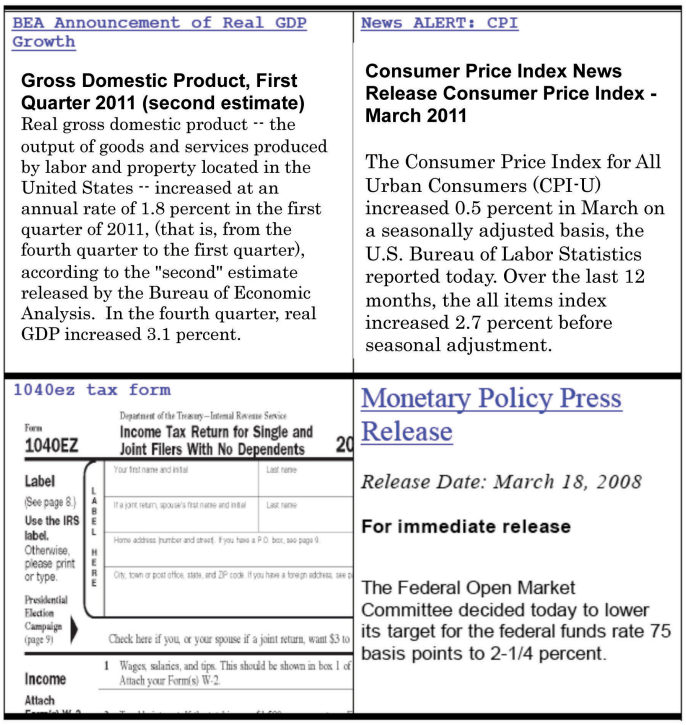


Figure 3.1.1

The four screens in Figure 3.1.1 are diverse illustrations of macroeconomics as you might encounter it:

- An evening news show presents a story about whether the economy is in a recession.
- You wonder why prices seem to be higher now than they were a few years ago.
- You sit down to fill out your tax return.
- You make payments on a car loan or a student loan.

By the time you have finished this book, you will see these examples very differently from the way you do right now. You may not know it, but your everyday life is filled with macroeconomics in action.

Economic Activity in the United States

The top left screen in Figure 3.1.1 is tuned to the Bureau of Economic Analysis (BEA; http://www.bea.gov), which is a part of the US government. A newspaper article or blog that reports such news from the BEA is telling us about the state of the macroeconomy. The report from the BEA tells you how the economy has been doing over the previous three months. More specifically, it describes what has happened to something called **real gross domestic product (real GDP)**.

As you will soon learn, real GDP is a measure of the overall level of economic activity within an economy. We won't worry for the moment about exactly what GDP means or how it is measured. Looking at the BEA announcement (www.bea.gov/newsreleases/national/gdp/2011/gdp1q11_2nd.htm), you can see that in the first quarter of 2011, real GDP increased by 1.8 percent, whereas in the fourth quarter of 2010, it increased by 3.1 percent. Because real GDP increased in both quarters, we

know that the economy is growing. However, it grew much more slowly in the first quarter of 2011 than in the final quarter of 2010.

You might wonder why you would bother to listen to this report. Perhaps it looks rather dry and boring. Yet the performance of the economy has a direct impact on how easy it is to find a job if you are looking for one, how likely you are to lose your job if you are already employed, how much you will earn, and what you can buy with the income you receive from working. Overall economic activity is directly linked to the well-being of everyone in the economy, including yourself. Should you be worried when you see that real GDP is growing much more slowly than before? After you have read this book, we hope you will know the answer.

Because real GDP is such a general measure of economic activity, it can also be used to compare how economies throughout the world are performing. If you have traveled to other countries, you may have observed big differences in people's standards of living. If you go to Canada, France, or Japan, you will generally see relatively prosperous people who can afford decent food, clothing, and shelter. If you go to Laos, Guatemala, or Malawi, you will see people living in severe poverty. To understand these differences, we need to understand what determines real GDP in an economy.

Inflation in the United States

The top right screen in Figure 3.1.1 reports on another economic variable that comes up all the time in the news: the **rate of inflation**. You have probably never visited the Bureau of Labor Statistics (BLS; http://www.bls.gov) website from which we took this quotation. But you have certainly heard a news story, perhaps on television or your car radio, telling you about the inflation rate.

After the BLS releases a report such as this one (http://www.bls.gov/news.release/cpi.nr0.htm), news programs will note that the inflation rate reported in March 2011 was 2.7 percent. This means that, on average, prices in the economy are 2.7 percent greater than they were a year ago. If you bought a jacket for $100 last year, you should expect the same jacket to cost about $102.70 right now. Not every single good and service increases by exactly this amount, of course. But, on average, prices are now 2.7 percent higher.

A news report like this tells us that the things we buy have become more expensive. This matters to all of us. If your income has not increased over the last year, this inflation report tells you that you are worse off now than you were last year because you can no longer buy as much with your income.

Most of the time, you will hear news reports about inflation only for the country in which you are living. Occasionally, you might also hear a news report about inflation somewhere else. In early 2008, you might well have heard a news report that the inflation rate in Zimbabwe was over 100,000 percent. You would probably find it difficult to imagine living in a country where prices increase so quickly, and you might reasonably wonder how two different countries in the world could have such different rates of inflation. When you have finished this book, you will know the answer to this question.

Fiscal Policy in Action

The bottom left screen in Figure 3.1.1 is something you may have seen before. It is a US tax form. Residents of the United States must file this form or one like it every year by April 15. If you live in another country, you almost certainly have to file a similar form. As individuals, we typically see this form as a personal inconvenience, and we don't think much about what it means for the economy as a whole. But this is much more than a form. It is a manifestation of decisions made by the government about how much tax you and everyone else should pay.

Decisions about how much to tax and how much to spend are known as **fiscal policy**. The fiscal policy adopted by a government affects your life in more ways than you can easily imagine. It not only tells you how much gets taken out of your paycheck, but it also affects real GDP and much more. It affects how likely you are to be unemployed in the future and how much money you will receive from the government if you do lose your job. It affects the interest rate you must pay on your car loan or student loan. It affects the tax rates you will pay 20 years from now and your likelihood of receiving social security payments when you retire.

Monetary Policy in Action

The bottom right screen in Figure 3.1.1 draws the attention of individuals and businesses all around the world. Every six weeks a group called the Federal Open Market Committee (FOMC) meets in Washington, DC, to make decisions on the course of US **monetary policy**. Their decisions affect the interest rates we pay on loans, including car loans, student loans, and mortgages. Their decisions also influence the level of economic activity and the inflation rate. The FOMC could, if it chose, create very high inflation by allowing rapid growth in the amount of money in the economy. It could, if it chose, create high rates of unemployment.

It is a powerful organization. There are other similar organizations elsewhere in the world: every country conducts monetary policy in some form, and most have some equivalent of the FOMC.

International Channels

Figure 3.1.1 shows the kind of economic news you might see in the United States. If you are living or traveling in a different country, you would see similar announcements about real GDP, inflation, and economic policy. Using the Internet, it is also easy to check news sources in other countries. If you start reading about economics on the Internet, you will come to appreciate the global nature of economics. You can read stories in the United States about monetary policy in China or fiscal policy in Portugal. And you can read news stories in other countries about economic policy in the United States. In the modern globalized world, economic connections across countries are impossible to ignore.

Figure 3.1.2 presents two stories that show globalization at work. Both share a common theme: the effects of a March 20, 2008, decision by the FOMC to cut the target federal funds rate. The graph at the top of Figure 3.1.2 shows the market price of the euro—the currency used in most of Europe—in terms of the British pound. When you travel, you typically exchange one currency for another. For example, an American tourist traveling to France would buy euros with dollars to have money to spend in France. If that same tourist then wanted to travel from France to London, she might take some of her euros and buy British pounds. The graph tells the price she would have paid in February and March of 2008.

You can see that, over a little more than a week, the euro became much more valuable relative to the pound. Most notably, there was a big increase in the price of the euro between March 9 and March 19, and then prices settled down a bit. This was a wild week for the international economy. In the United States, the Federal Reserve announced major financial support for Wall Street firms on March 16 and then reduced interest rates on March 19. Around the same time, the European Central Bank (ECB) and the Bank of England in London were also taking actions to try to calm the financial markets. At least for a period of time, they seemed to succeed in stopping the rapid rise of the euro against the British pound. It is striking that much of the financial action was taking place in the United States, yet the markets in which Europeans trade currencies were also affected.

The story at the bottom of Figure 3.1.2 discusses the response of Asian stock markets to the action of the US Federal Reserve. Markets all over the world increased in value after the action of the FOMC. The actions of the Fed matter well beyond the borders of the United States. Bankers and businesspeople all over the globe are “Fed watchers.”

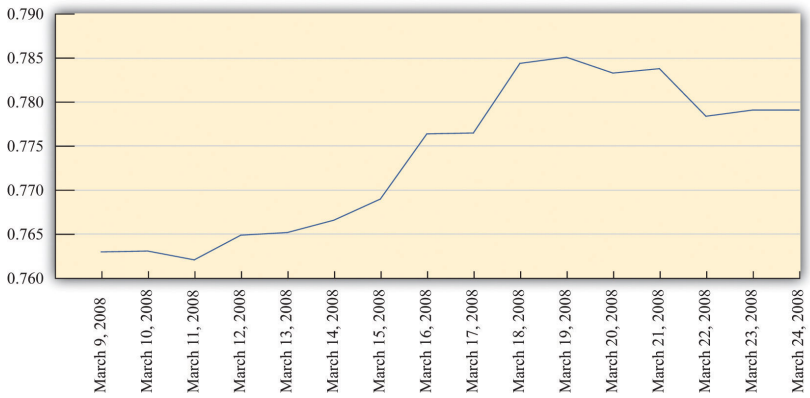


Figure 3.1.2: Price of Euro in British Pounds, March 2008. Source: *http://www.oanda.com*.

Asian Stocks Rise after Fed Cut

TOKYO (AP)—Asian stock markets rose Wednesday as investors welcomed a hefty U.S. interest rate cut…

Japan’s benchmark Nikkei 225 index climbed 2.5 percent to close at 12,260.44 after rising more than 3 percent earlier. Hong Kong’s Hang Seng index, which rose as much as 3 percent earlier, closed up 2.3 percent at 21,866.94.

Australia’s main index jumped 4 percent, and markets in South Korea, China and India also rose.“Asian Stocks Rise after Fed Cut,” MSNBC.com, March 19, 2008, accessed June 27, 2011, http://www.msnbc.msn.com/id/23703748/ns/business-eye_on_the_economy.

Key Takeaways

- You encounter macroeconomics everyday through the news about the state of the macroeconomy, the price you pay for goods and services, the tax you pay on income, and the effects of macroeconomic policy on interest rates. Macroeconomic events and policies in other countries affect you as well.
- Real GDP, the rate of inflation, and the rate of unemployment are three primary indicators of the state of the macroeconomy.
- The government influences the macroeconomy through its level of spending, taxes, and control of the money supply.

check your understanding

1. What do we mean by “real” when we talk about GDP?
2. How might the state of the macroeconomy in another country, such as China, or in a group of countries, such as the European Union, affect the macroeconomy of the United States?

This page titled 3.1: Four Examples of Macroeconomics is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

- **Current page** by Anonymous is licensed CC BY-NC-SA 3.0. Original source: https://2012books.lardbucket.org/books/theory-and-applications-of-economics.
- **23.3: Job and Worker Flows** by Anonymous is licensed CC BY-NC-SA 3.0. Original source: https://2012books.lardbucket.org/books/theory-and-applications-of-economics.

3.2: Behind the Screens 3.2: Behind the ScreensEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. How has real GDP changed over the past 40 years?
2. What is inflation and how does it affect the macroeconomy?
3. How can we see fiscal and monetary policy in action?

Let's look at Figure 3.2.1 again in a bit more detail.

The State of the Economy

The top two panels in Figure 3.2.1 provide information on some key indicators of the state of the economy. The announcement from the Bureau of Economic Analysis (BEA) concerns one of the most closely watched indicators of the macroeconomy: real gross domestic product (real GDP). This is a measure of the goods and services produced by an economy in a year. We discuss real GDP in every macroeconomic application in this book.

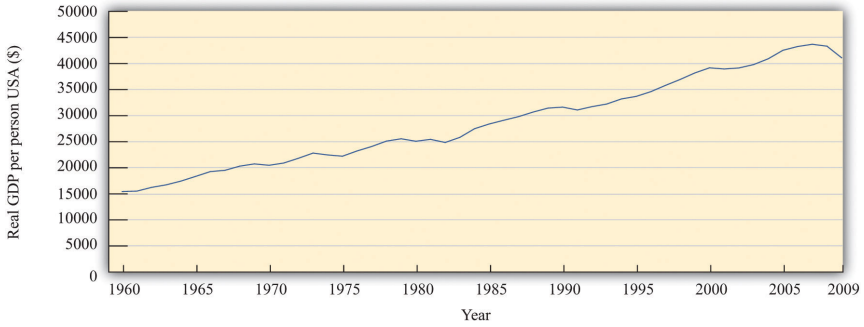


Figure 3.2.1: Real GDP per Person in the United States, 1960–2009. Source: Alan Heston, Robert Summers, and Bettina Aten, Penn World Table Version 7.0, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, May 2011.

Figure 3.2.1 shows real GDP per person (often called real GDP per capita) from 1960 to 2009. Pictures like this one show up all the time in newspapers, in magazines, on television, or on the Internet. One of the things you will learn in your study of macroeconomics is how to interpret such economic data. We devote an entire chapter to understanding exactly how real GDP is measured. For now, we draw your attention to some details to help you appreciate what the graph means.

The horizontal axis indicates the year. Real GDP per person is shown on the vertical axis. To read this graph, you would look at a particular year on the horizontal axis, such as 2000, and then use the curve to see that the real GDP per person in 1965 was about $39,000.

If you look at this picture, the single most notable thing is that real GDP per person has been increasing. It was about 2.6 times larger in 2009 than in 1960. This tells us that, on average, the typical individual in the United States was 2.6 times richer in 2000 compared to 1960. The increase in GDP is not caused by the fact that there are more people in the economy because the figure shows GDP *per person*. The increase in GDP is not because prices are going up: the word *real* in this discussion means that it has been *corrected for inflation*.In the bottom right of the picture, you can see the phrase *Data in 1996 dollars*. This means that the numbers in the table are based on how much a dollar would have bought in 1996. Do not worry if you do not understand exactly what this phrase means right now. Chapter 18 will provide much more detail.

Another thing you can see from the picture is that the growth of the economy has not been smooth. Sometimes the economy grows fast; sometimes it grows more slowly. Sometimes there are even periods in which the economy shrinks rather than grows. From this figure, you can see that real GDP per person decreased in the mid-1970s, the mid-1980s, and most notably in 2008 and 2009. During these times, people were becoming poorer on average, not richer.

We keep using the phrase *on average*. This reminds us that, even though the economy as a whole has been getting richer, the picture doesn't tell us anything about how those gains have been shared across the economy. In fact, some people became a lot richer over this period, while many others saw only small gains, and some became poorer.

We see this uneven distribution very clearly when the economy shrinks. When that happens, one of the things we also observe is that more people in the economy are unemployed—that is, they are looking for a job but unable to find one. The burden of an economic downturn is borne disproportionately by those who lose their jobs.

Although this figure displays the history of the US economy over these 50 years, similar figures can be constructed for other countries around the world. They do not all look identical, but the pattern of uneven growth that we observe for the United States is one that we also see for most other countries. However, it is not true everywhere. We will also see examples of countries that have become poorer rather than richer in recent decades.

Real GDP is the most frequently watched indicator of economic performance. A second key indicator is the one in the top right screen of Figure 3.2.1: the inflation rate. The Bureau of Labor Statistics (BLS) collects information on prices on an ongoing basis; each month it releases information on how fast prices are changing. The rate at which prices are changing is the inflation rate. Other countries similarly have government agencies entrusted with gathering information about the inflation rate and other economic indicators.

It may seem that the job of the BLS is pretty easy: get information on prices and report it. Their task is, in fact, rather complex. In part, it is difficult because there are so many goods and services in the economy. So when we say that prices are increasing, we must decide which goods and services we are talking about. In addition, new goods appear, and obsolete goods disappear; the BLS must take this into account. And the quality of goods changes as well. If the price of a computer increases, is this an example of inflation or does it reflect an increase in the quality of the computer?

What are the implications of an inflation announcement? All else being the same, higher prices mean that we are unable to afford goods and services we were able to buy when prices were lower. But “all else” is not the same. Generally when prices increase, wages also increase. This means that the overall effects of inflation on our ability to buy goods and services are not self-evident.

Another implication of inflation is the policy response it elicits. The monetary authorities in the United States and many other countries are focused on ensuring that inflation does not get out of control. A report of inflation might therefore lead to a response by a monetary authority. Inflation affects us directly through the prices we pay and the wages we receive and indirectly through the policy response it induces.

Though not included in our screens, another significant variable also indicates the state of the macroeconomy: the rate of unemployment. The BLS (http://www.bls.gov/news.release/empsit.toc.htm) reports the **unemployment rate** on a monthly basis. It measures the fraction of people in the labor force who do not have a job. When real GDP is relatively high, then the unemployment rate tends to be lower than average, but when real GDP decreases, more people find themselves out of a job.

The Making of Fiscal and Monetary Policy

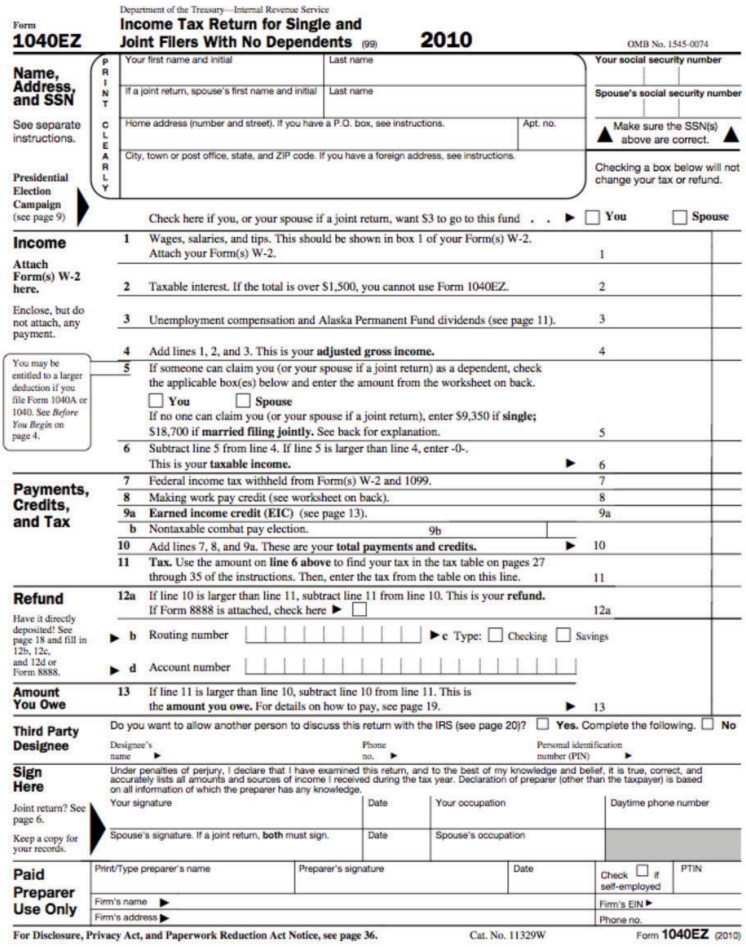
The top screens in Figure 3.2.1 provide information that flows to the policymakers in an economy. These policymakers carefully watch the state of the economy and then, if appropriate, take actions. The bottom screens in Figure 3.2.1 show policy in action.

Fiscal Policy

For individuals and firms paying taxes in the United States, April 15 is an important day because tax forms are due for the previous calendar year. Each year US citizens fill out their tax forms and either make tax payments or receive reimbursements from the government.

The tax day differs across countries, but the experience is much the same everywhere: individuals and firms must pay taxes to the government. This is one of the key ways in which citizens interact with their governments.

A more complete version of the 1040EZ form for 2010 is shown in Figure 3.2.2.



Department of the Treasury—Internal Revenue Service

Form **1040EZ** **Income Tax Return for Single and Joint Filers With No Dependents** (99) **2010** OMB No. 1545-0074

Name, Address, and SSN See separate instructions. PRINT CLEARLY

Your first name and initial | Last name | Your social security number

If a joint return, spouse's first name and initial | Last name | Spouse's social security number

Home address (number and street). If you have a P.O. box, see instructions. | Apt. no. | Make sure the SSN(s) above are correct.

City, town or post office, state, and ZIP code. If you have a foreign address, see instructions. | Checking a box below will not change your tax or refund.

Presidential Election Campaign (see page 9) Check here if you, or your spouse if a joint return, want $3 to go to this fund . . ▶ ☐ You ☐ Spouse

Income

Attach Form(s) W-2 here. Enclose, but do not attach, any payment.

1	Wages, salaries, and tips. This should be shown in box 1 of your Form(s) W-2. Attach your Form(s) W-2.	1
2	Taxable interest. If the total is over $1,500, you cannot use Form 1040EZ.	2
3	Unemployment compensation and Alaska Permanent Fund dividends (see page 11).	3
4	Add lines 1, 2, and 3. This is your **adjusted gross income.**	4
5	If someone can claim you (or your spouse if a joint return) as a dependent, check the applicable box(es) below and enter the amount from the worksheet on back. ☐ You ☐ Spouse If no one can claim you (or your spouse if a joint return), enter $9,350 if **single;** $18,700 if **married filing jointly.** See back for explanation.	5
6	Subtract line 5 from line 4. If line 5 is larger than line 4, enter -0-. This is your **taxable income.** ▶	6

You may be entitled to a larger deduction if you file Form 1040A or 1040. See *Before You Begin* on page 4.

Payments, Credits, and Tax

7	Federal income tax withheld from Form(s) W-2 and 1099.	7
8	Making work pay credit (see worksheet on back).	8
9a	**Earned income credit (EIC)** (see page 13).	9a
b	Nontaxable combat pay election. 9b	
10	Add lines 7, 8, and 9a. These are your **total payments and credits.** ▶	10
11	**Tax.** Use the amount on **line 6 above** to find your tax in the tax table on pages 27 through 35 of the instructions. Then, enter the tax from the table on this line.	11

Refund

Have it directly deposited! See page 18 and fill in 12b, 12c, and 12d or Form 8888.

12a	If line 10 is larger than line 11, subtract line 11 from line 10. This is your **refund.** If Form 8888 is attached, check here ▶ ☐	12a
▶ b	Routing number ▶ c Type: ☐ Checking ☐ Savings	
▶ d	Account number	

Amount You Owe

13	If line 11 is larger than line 10, subtract line 10 from line 11. This is the **amount you owe.** For details on how to pay, see page 19. ▶	13

Third Party Designee Do you want to allow another person to discuss this return with the IRS (see page 20)? ☐ **Yes.** Complete the following. ☐ **No**

Designee's name ▶ | Phone no. ▶ | Personal identification number (PIN) ▶

Sign Here Under penalties of perjury, I declare that I have examined this return, and to the best of my knowledge and belief, it is true, correct, and accurately lists all amounts and sources of income I received during the tax year. Declaration of preparer (other than the taxpayer) is based on all information of which the preparer has any knowledge.

Joint return? See page 6. Keep a copy for your records.

Your signature | Date | Your occupation | Daytime phone number

Spouse's signature. If a joint return, **both** must sign. | Date | Spouse's occupation

Paid Preparer Use Only Print/Type preparer's name | Preparer's signature | Date | Check ☐ if self-employed | PTIN

Firm's name ▶ | Firm's EIN ▶

Firm's address ▶ | Phone no.

For Disclosure, Privacy Act, and Paperwork Reduction Act Notice, see page 36. Cat. No. 11329W Form **1040EZ** (2010)

Figure 3.2.2: Form 1040EZ

From the perspective of an individual filling out this form, the task is to get the data correct and determine exactly what figures go where on the form. This is no small challenge. From the perspective of economists working for the government, the tax form is an instrument of fiscal policy. Embedded in the tax form are various tax rates that must be paid on the different types of income you earn.

Where do these tax revenues go? The government collects taxes to finance its purchases of goods and services in the economy—such as roads, schools, and national defense—and also to make transfers to households, such as unemployment insurance.

The tax forms we fill out change each year, sometimes quite significantly. The tax rates households and firms confront are changed by governmental decisions. The government alters tax rates to affect the level of economic activity in the economy. It uses these tools when, in its judgment, the level of economic activity (as measured by real GDP, the unemployment rate, and other variables we will learn about) is insufficient. This is a delicate assessment that requires an understanding of the meaning and measurement of satisfactory economic performance and a deep understanding of how the economy works.

For example, consider the winter of 2008. Policymakers working in the White House and on Capitol Hill kept careful track of the state of the economy, looking as we just did at announcements from the BEA and the BLS on output and inflation. Eventually, they concluded that economic activity was not at a high enough level. They took actions to increase output by reducing taxes through the American Recovery and Reinvestment Act of 2009 (www.irs.gov/newsroom/article/0,,id=204335,00.html). The idea is as follows: when people pay less in taxes, they have more income available to spend, so they will purchase more goods and services. The link between the legislation and you as an individual is through tax forms like the one shown in Figure 3.2.2.

Monetary Policy

The bottom right screen in Figure 3.2.1 shows a decision of the Federal Open Market Committee (FOMC) to reduce a key interest rate by three-fourths of a percentage point to 2.25 percent. As we shall see in our study of monetary policy, a reduction in interest

rates is a tool to increase economic activity. Lower interest rates make it cheaper for households and firms to borrow, so they spend more on goods and services. The FOMC action was taken on account of weak economic conditions in the United States, but its consequences were felt worldwide.

Other monetary authorities likewise look at the state of their economies and adjust their monetary policy. The following is part of a statement from the European Central Bank (ECB), the monetary policy authority for the European Union. It was part of a press conference held in April 2005 in which Jean-Claude Trichet, president of the ECB, and Lucas Papademos, vice president of the ECB, provided a statement about economic outlook for Europe and the stance of monetary policy.

All in all, we have not changed our assessment of risks to price stability over the medium term. So far, we have seen no significant evidence of underlying domestic inflationary pressures building up in the euro area. Accordingly, we have left the key ECB interest rates unchanged. Both nominal and real rates are at exceptionally low levels, lending ongoing support to economic activity. However, upside risks to price stability over the medium term remain and continued vigilance is therefore of the essence.

I shall now explain our assessment in more detail, turning first to the **economic analysis**. Recent data and survey indicators on economic activity have been mixed. In general they point to ongoing economic growth at a moderate pace over the short term, with no clear signs as yet of a strengthening in underlying dynamics.

Looking further ahead, the conditions remain in place for moderate economic growth to continue. Global growth remains solid, providing a favourable environment for euro area exports. On the domestic side, investment is expected to continue to be supported by very favourable financing conditions, improved profits and greater business efficiency. Consumption growth should develop in line with real disposable income growth. However, at the same time, persistently high oil prices in particular pose downside risks to growth.

[…]"Introductory Statement with Q&A," European Central Bank, April 7, 2005, accessed June 27, 2011, http://www.ecb.int/press/pressconf/2005/html/is050407.en.html.

Statements such as this are reported in the business press and widely read. Businesspeople all over the world closely follow the actions of central banks. That is, the people interested in this statement by the ECB were not only European citizens but also individuals in the United States and other countries. Likewise, when the Fed takes action, the news shows up on televisions and computer screens across the world.

The ECB quotation mentions several key economic variables: inflation, real interest rates, nominal interest rates, economic activity, investment, exports, consumption growth, and real disposable income growth. These variables are also important indicators of the state of the economy, as we can tell from the fact that they play such a prominent role in the ECB assessment.

The economists at the ECB need to know the current state of the economy when deciding on what policies to pursue. But there are compelling reasons for others to care about these variables as well. Suppose, for example, that you are an investor contemplating an investment in Spain. Your interest is in making profit from producing a good in Spain and selling it in that country and others. The profitability of the investment in Spain depends on the overall state of the Spanish economy and its neighbors in the European Union who are the target group for your sales.

For you as an investor, the ECB statement contains vital information about the state of the European economy. It also contains information on the likely conduct of monetary and fiscal policy in Europe. These factors matter for you simply because they impact the profitability of your investment. Thus you want to understand the statements from the ECB, starting with the definitions of key macroeconomic variables.

By now, you may well have a number of questions. What exactly are these monetary authorities in Europe and the United States? Where do they come from and what are their powers? How exactly do their actions have so much influence on our lives? Answering these questions is one of our tasks in this book. We devote two full chapters to the determination and the influence of monetary policy in the economy.

Key Takeaways

- Real GDP has grown on average over the past 50 years, but the growth is not always constant: sometimes the economy grows quickly and sometimes real GDP grows slowly (or not at all).
- The inflation rate measures the percent change in prices. If prices are increasing, then a unit of currency, such as a dollar, buys fewer goods and services. During a period of inflation, the monetary authority may take action to reduce the inflation rate.

- Each year, the income taxes we pay to the government reflect its choice of fiscal policy. The policy meetings of the FOMC in the United States, the ECB of the European Monetary Union, and other central banks around the world are examples of monetary policy.

check your understanding

1. Which of the macroeconomic variables discussed would a fiscal authority pay attention to?
2. Do the ECB and the FOMC always make the same policy decision?
3. Is a change in the tax code an example of fiscal or monetary policy?

This page titled 3.2: Behind the Screens is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

3.3: Between News and Policy- The Framework of Macroeconomics 3.3: Between News and Policy- The Framework of MacroeconomicsEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What is the methodology of macroeconomics?
2. What is the role of models in the making of macroeconomic policy?

We have seen the news and policy in action. But there is a vital piece missing: given the economic news, how do policymakers know what to do? The answer to this question is at the heart of this book. The basic methodology of macroeconomics is displayed in Figure 3.3.1. Macroeconomics involves the interplay of theory, data, and policy. We have already seen two of these components in Figure 3.3.1. Two screens highlighted data we have on the macroeconomy, and two screens highlighted policy actions.

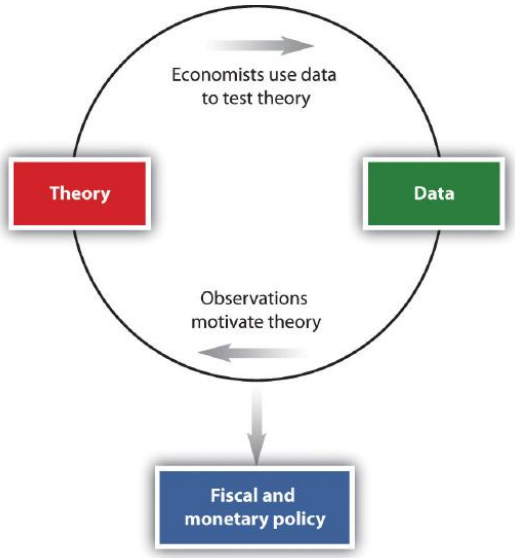


Figure 3.3.1: Macroeconomics Methodology

The answer to the question "how do policymakers know what to do?" is on the top left of Figure 3.3.1: theory. Macroeconomists typically begin by observing the world and then try to develop a theoretical framework to explain what they have seen. (An old joke says that the definition of an economist is "someone who sees something happen in practice and wonders whether on earth it is possible in theory.") Usually, a theory developed by economists has a mathematical foundation—expressed by either equations or diagrams. There is even a bit of art here: the theoretical framework must be simple enough to work with yet realistic enough to be useful.

We hinted at these theories in our earlier discussion when we explained that both monetary policy and fiscal policy affect the economy by changing the willingness of households and firms to purchase goods and services. In our applications chapters, we develop these ideas and explain the frameworks that policymakers use when deciding on their policies.

Our frameworks—or models, as they are often called—are tested by their ability to match existing data and provide accurate predictions about new data. Models are constantly refined so that they can do a better job of matching facts. After many rounds of interaction between theory and data, a useful framework emerges. This then becomes the basis for policymaking.

How do policymakers know about the theories devised by economists? Politicians are typically not expert economists. In most countries, a large number of trained economists are employed as advisors to the government. These individuals have studied economic theory and are also familiar with economic statistics, allowing them to provide the link between the economic frameworks and the actual implementation of policy.

The big challenge for economists is to understand the links from policy to the aggregate economy. When you first learned to drive, you were presumably introduced to all the instruments in the car: the steering wheel, the accelerator, the brake, the mirrors, and so forth. At the same time, you were learning the rules of the road. For many, the instruments of the car are easy enough to grasp, and

the rules of the road are reasonably intuitive. The difficulty (and this is why driving schools make money) is in making the connection between the controls in the car and the outcome you wish to achieve while driving. The same is true of economic modeling: policy tools are not very difficult to understand, yet it can take decades of experience to truly understand how to use these tools effectively.

Economists and businesspeople hope, for example, that the current chairman of the Federal Reserve, Ben Bernanke, has this understanding, as discussed in the following news article excerpt.

Economic View: Bernanke's Models, and Their Limits

In terms of intellect, Ben S. Bernanke may be to the Federal Reserve what John G. Roberts Jr. is to the Supreme Court. And like Chief Justice Roberts, Mr. Bernanke, the nominee to replace Alan Greenspan at the Fed, has left a paper trail worth studying. What can it tell us about the sort of Fed chairman he would be?

In general, Mr. Bernanke's work has been solidly in the mainstream—a mainstream he has helped define since he began publishing papers in major economic journals since 1981. He has written repeatedly about ways of using mathematical models of a dauntingly complex economy to set monetary policy. When he has strayed from that subject, his conclusions have sometimes raised eyebrows.

[…]

These topics, however, are not at the core of what Mr. Bernanke would be concerned with at the Fed. There, his opinions about domestic monetary policy would be more important. One tenet of Mr. Bernanke's philosophy could not be clearer: that the central bank should use a model, not just hunches, to decide about interest rates and the money supply.

This is how he put it in 1997 in a paper with Michael Woodford, now a professor of political economy at Columbia: "We conclude that, although private-sector forecasts may contain information useful to the central bank, ultimately the monetary authorities must rely on an explicit structural model of the economy to guide their policy decisions."

[…]Daniel Altman, "Economic View: Bernanke's Models, and Their Limits," *New York Times*, October 30, 2005, accessed June 27, 2011, http://www.nytimes.com/2005/10/30/business/yourmoney/30econview.html.

Key Takeaways

- The methodology of macroeconomics involves the interplay between data and models.
- Abstract models provide policymakers with a framework to understand what is happening in the macroeconomy and also a way to predict the effects of policy actions.

check you understanding

1. Why are economic models always being refined?
2. If a theory is inconsistent with some but not all observations, could it still be useful for policymaking purposes?

This page titled 3.3: Between News and Policy- The Framework of Macroeconomics is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

3.4: End-of-Chapter Material 3.4: End-of-Chapter MaterialEdit Update titlecancel

In Conclusion

Our book is built around economic topics, such as the income tax code, the social security system, the determination of monetary policy in Europe, and the contrasting economic health of different countries.

Throughout this book, we will emphasize the measurement and interpretation of economic data. Understanding how to read charts and tables of economic data is a critical skill for anyone who wants to be a sophisticated consumer of economic and political news. We also explain both policy tools and their links to economic outcomes. Understanding these links requires a model of the economy. We introduce models as needed, in the context of their applications. Mastering macroeconomics involves both understanding the tools that macroeconomists use and knowing how and when those tools should be applied. In this book, you will learn about these tools by example: you will see them in use as we study different questions in economics. At the same time, you will learn about many topics that should interest you as engaged and aware citizens of the world. We hope that, after reading this book, you will both better understand what it is that economists do and be better informed about the world in which we all live.

As you proceed through the chapters, you will often see reference to our *toolkit*. This is a collection of some of the most important tools that we use over and over in different chapters. Each tool is fully introduced somewhere in the book, but you can also use the toolkit as a reference when working through different chapters. In addition, it can serve as a study aid when you are preparing for quizzes and examinations.

We try to avoid getting too hung up on the mathematical expression of our theories (although the math will usually be lurking in the background where you can't quite see it). In particular, our applications chapters contain very little mathematics. This means that you can read and understand the applications without needing to work through a lot of mathematics. Compared to our applications chapters, our toolkit contains slightly more formal versions of the frameworks that we develop. You will refer to the tools over and over again as we progress through the book, for the same tool is often used to shed light on all sorts of different questions.

key links

- Bureau of Economic Analysis: http://www.bea.gov
- Bureau of Labor Statistics: http://www.bls.gov
- Board of Governors of the Federal Reserve System: http://www.federalreserve.gov
- European Central Bank: http://www.ecb.int/home/html/index.en.html

exercises

1. Provide updated information for at least one of the four screens in Figure 3.4.1.
2. Use the Internet to find an article (for example, magazine, newspaper, publication of an economics research group) that contains a graph of real GDP for a country other than the United States. What purpose does the picture serve in the article? Why do you think it was included?
3. Find a statement about monetary policy from the monetary authority in the United States, Canada, or Australia. What are some of the indicators of the state of the economy that are used in the policy statement?
4. The article on Bernanke's model contained the following quote: "We conclude that, although private-sector forecasts may contain information useful to the central bank, ultimately the monetary authorities must rely on an explicit structural model of the economy to guide their policy decisions." What do you think is meant by this statement?

This page titled 3.4: End-of-Chapter Material is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

4: Everyday Decisions

4.1: You and Your Choices
4.2: Individual Decision Making- How You Spend Your Income
4.3: Individual Demand
4.4: Individual Decision Making- How You Spend Your Time
4.5: End-of-Chapter Material

This page titled 4: Everyday Decisions is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

4.1: You and Your Choices 4.1: You and Your ChoicesEdit Update titlecancel

Economics is about you. It is about how you make choices. It is about how you interact with other people. It is about the work you do and how you spend your leisure time. It is about the money you have in your pocket and how you choose to spend it. Because economics is about your choices plus everyone else's, this is where we begin. As far as your own life is concerned, you are the most important economic decision maker of all. So we begin with questions you answer every day:

- What will I do with my money?
- What will I do with my time?

Economists don't presume to tell you what you *should* do with your time and money. Rather, studying economics can help you better understand your own choices and make better decisions as a consequence. Economics provides guidelines about how to make smart choices. Our goal is that after you understand the material in this chapter, you will think differently about your everyday decisions.

Decisions about spending money and time have a key feature in common: *scarcity*. You have more or less unlimited desires for things you might buy and ways that you might spend your time. But the time and the money available to you are limited. You don't have enough money to buy everything you would like to own, and you don't have enough time to do everything you would like to do.

Because both time and money are scarce, whenever you want more of one thing, you must accept that you will have less of something else. If you buy another game for your Xbox, then you can't spend that money on chocolate bars or movies. If you spend an hour playing that game, then that hour cannot be spent studying or sleeping. Scarcity tells us that everything has a cost. The study of decision making in this chapter is built around this tension. Resources such as time and money are limited even though desires are essentially limitless.

Road Map

We tackle the two questions of this chapter in turn, but you will see that there are close parallels between them. We begin by looking at spending decisions. Although we have said that money is scarce, a more precise statement is that you have limited *income*. (Economists usually use the term "money" more specifically to mean the assets, such as currency in your wallet or funds in your checking account, that you use to buy things.) Because your income is limited, your spending opportunities are also limited. We show how to use the prices of goods and services, together with your income, to analyze what spending decisions are *possible* for you. Then we think about what people's wants and desires look like. Finally, we put these ideas together and uncover some principles about how to make choices that will best satisfy these desires.

Your decisions about what to buy therefore depend on how much income you have and the prices of goods and services. Economics summarizes these decisions in a simple way by using the concept of *demand*. We show how demand arises from the choices you make. Demand is one of the most useful ideas in economics and lies at the heart of almost everything we study in this book.

Finally, we turn to the decision about how to spend your time. Again, we begin with the idea that your resources are limited: there are only so many hours in a day. As with the spending decision, you have preferences about how to spend your time. We explain the principles of good decision making in this setting. Based on this analysis, we introduce another central economic idea—that of *supply*.

Economics is both *prescriptive* and *descriptive*. Economics is prescriptive because it tells you some rules for making good decisions. Economics is descriptive because it helps us explain the world in which we live. As well as uncovering some principles of good decision making, we discuss whether these are also useful descriptions of how people actually behave in the real world.

This page titled 4.1: You and Your Choices is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

4.2: Individual Decision Making- How You Spend Your Income 4.2: Individual Decision Making- How You Spend Your IncomeEdit Update titlecancel

Learning Objectives

1. What are an individual's budget set and budget line?
2. What is an opportunity cost?
3. How do people make choices about how much to consume?
4. What features do we expect most people's preferences to have?
5. What does it mean to make rational choices?

We start with the decision about how to spend your income. We want to know what possibilities are available to you, given that your income is limited but your desires are not.

The Budget Set

We describe your personal decision making on a month-by-month basis (although we could equally well look at daily, weekly, or even annual decisions because the same basic ideas would apply). Suppose you receive a certain amount of income each month—perhaps from a job or a student grant. The government takes away some of this income in the form of taxes, and the remainder is available for you to spend. We call the income that remains after taxes your **disposable income**.

You may want to put aside some of this income for the future; this is your savings. The remainder is your consumption, which is your spending on all the goods and services you buy this month: rent, food, meals out, movies, cups of coffee, CDs, music downloads, DVD rentals, chocolate bars, books, bus rides, haircuts, and so on. Figure 4.2.1 shows this process.

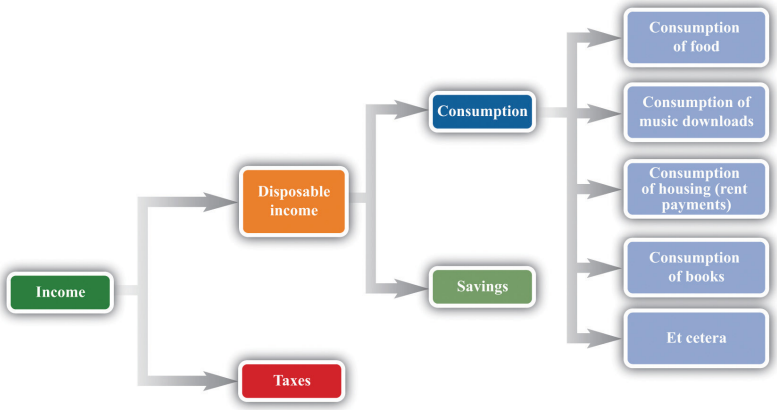


Figure 4.2.1: What You Do with Your Income

Here is a schematic view of what happens to your income.

This view of your paycheck involves several economic decisions. Some of these are decisions made by the government. Through its tax policies, the government decides how much of your income it takes from you and how much is left as disposable income. You make other decisions when you allocate your disposable income among goods and services today and in the future. You choose how to divide your disposable income between consumption this month and saving for the future. You also decide exactly how much of each good and service you purchase this month. We summarize your ability to purchase goods and services by your **budget set**.

Toolkit: Section 31.1

The budget set is a list of all the possible combinations of goods and services that are affordable, given both income and the prices of all goods and services. It is defined by

$$total\ spending \leq disposable\ income.$$

Begin by supposing you neither save nor borrow. We can construct your budget set in three steps.

1. Look at spending on each good and service in turn. For example, your monthly spending on cups of coffee is as follows:

$$spending\ on\ coffee = number\ of\ cups\ purchased \times price\ per\ cup.$$

   A similar equation applies to every other good and service that you buy. Your spending on music downloads equals the number of downloads times the price per download, your spending on potato chips equals the number of bags you buy times the price per bag, and so on.

2. Now add together all your spending to obtain your total spending:

$$total\ spending = spending\ on\ coffee + spending\ on\ downloads + \ldots,$$

   where … means including the spending on every different good and service that you buy.

3. Observe that your total spending cannot exceed your income after taxes:

$$total\ spending \leq disposable\ income.$$

   You are consuming within your budget set when this condition is satisfied.

In principle, your list of expenditures includes every good and service you could ever imagine purchasing, even though there are many goods and services you never actually buy. After all, your spending on Ferraris every month equals the number of Ferraris that you purchase times the price per Ferrari. If you buy 0 Ferraris, then your spending on Ferraris is also $0, so your total spending does include all the money you spend on Ferraris.

Imagine now that we take some bundle of products. *Bundle* here refers to any collection of goods and services—think of it as being like a grocery cart full of goods. The bundle might contain 20 cups of coffee, 5 music downloads, 3 bags of potato chips, 6 hours of parking, and so on. If you can afford to buy this bundle, given your income, then it is in the budget set. Otherwise, it is not.

The budget set, in other words, is a list of all the possible collections of goods and services that you can afford, taking as given both your income and the prices of the goods and services you might want to purchase. It would be very tedious to write out the complete list of such bundles, but fortunately this is unnecessary. We merely need to check whether any given bundle is *affordable* or not. We are using affordable not in the casual everyday sense of “cheap” but in a precise sense: a bundle is affordable if you have enough income to buy it.

It is easiest to understand the budget set by working though an example. To keep things really simple, suppose there are only two products: chocolate bars and music downloads. An example with two goods is easy to understand and draw, but everything we learn from this example can be extended to any number of goods and services.

Suppose your disposable income is $100. Imagine that the price of a music download is $1, while the price of a chocolate bar is $5. Table 4.2.1 "Spending on Music Downloads and Chocolate Bars" shows some different bundles that you might purchase. Bundle number 1, in the first row, consists of one download and one chocolate bar. This costs you $6—certainly affordable with your $100 income. Bundle number 2 contains 30 downloads and 10 chocolate bars. For this bundle, your total spending on downloads is $30 (= 30 × $1), and your total spending on chocolate bars is $50 (= 10 × $5), so your overall spending is $80. Again, this bundle is affordable. You can imagine many other combinations that would cost less than $100 in total.

Table 4.2.1: Spending on Music Downloads and Chocolate Bars

Bundle	Number of Downloads	Price per Download ($)	Spending on Downloads ($)	Number of Chocolate Bars	Price per Chocolate Bar ($)	Spending on Chocolate Bar ($)	Total Spending ($)
1	1	1	1	1	5	5	6
2	30	1	30	10	5	50	80
3	50	1	50	10	5	50	100
4	20	1	20	16	5	80	100
5	65	1	65	7	5	35	100
6	100	1	100	0	5	0	100
7	0	1	0	20	5	100	100

Bundle	Number of Downloads	Price per Download ($)	Spending on Downloads ($)	Number of Chocolate Bars	Price per Chocolate Bar ($)	Spending on Chocolate Bar ($)	Total Spending ($)
8	50	1	50	11	5	55	105
9	70	1	70	16	5	80	150
10	5,000	1	5,000	2,000	5	10,000	15,000

Bundles 3, 4, 5, 6, and 7 are special because they are affordable if you spend *all* your income. For example, you could buy 50 downloads and 10 chocolate bars (bundle 3). You would spend $50 on music downloads and $50 on chocolate bars, so your total spending would be exactly $100. Bundle 4 consists of 20 downloads and 16 chocolate bars; bundle 5 is 65 downloads and 7 chocolate bars. Again, each bundle costs exactly $100. Bundle 6 shows that, if you chose to buy nothing but downloads, you could purchase 100 of them without exceeding your income, while bundle 7 shows that you could buy 20 chocolate bars if you chose to purchase no downloads. We could find many other combinations that—like those in bundles 3–7—cost exactly $100.

Bundles 8, 9, and 10 are *not* in the budget set. Bundle 8 is like bundle 3, except with an additional chocolate bar. Because bundle 3 cost $100, bundle 8 costs $105, but it is not affordable with your $100 income. Bundle 9 costs $150. Bundle 10 shows that you cannot afford to buy 5,000 downloads and 2,000 chocolate bars because this would cost $15,000. There is quite literally an infinite number of bundles that you cannot afford to buy.

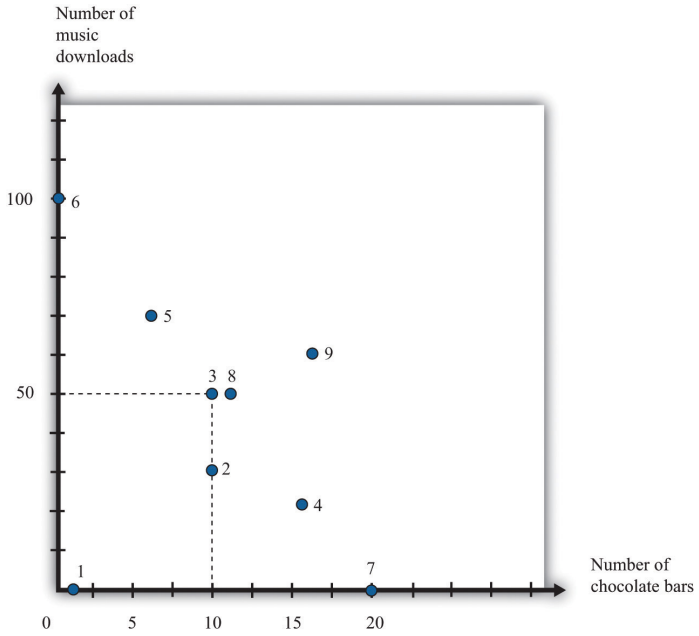


Figure 4.2.2: Various Bundles of Chocolate Bars and Downloads

This figure shows the combinations of chocolate bars and music downloads from Table 4.2.1 "Spending on Music Downloads and Chocolate Bars".

Figure 4.2.2 illustrates the bundles from Table 4.2.1 "Spending on Music Downloads and Chocolate Bars". The vertical axis measures the number of music downloads, and the horizontal axis measures the number of chocolate bars. *Any point on the graph therefore represents a consumption bundle—a combination of music downloads and chocolate bars*. We show the first nine bundles from Table 4.2.1 "Spending on Music Downloads and Chocolate Bars" in this diagram. (Bundle 10 is several feet off the page.) If you inspect this figure carefully, you may be able to guess for yourself what the budget set looks like. Look in particular at bundles 3, 4, 5, 6, and 7. These are the bundles that are just affordable—that cost exactly $100. It appears as if these bundles all lie on a straight line, which is in fact the case. All the combinations of downloads and chocolate bars that are just affordable represent a straight line.

Meanwhile, the bundles that are affordable with income to spare—like bundles 1 and 2—are below the line, and the bundles you cannot afford—like bundles 8, 9, and 10—are above the line. Building on these discoveries, we find that the budget set is a triangle

(Figure 4.2.3).

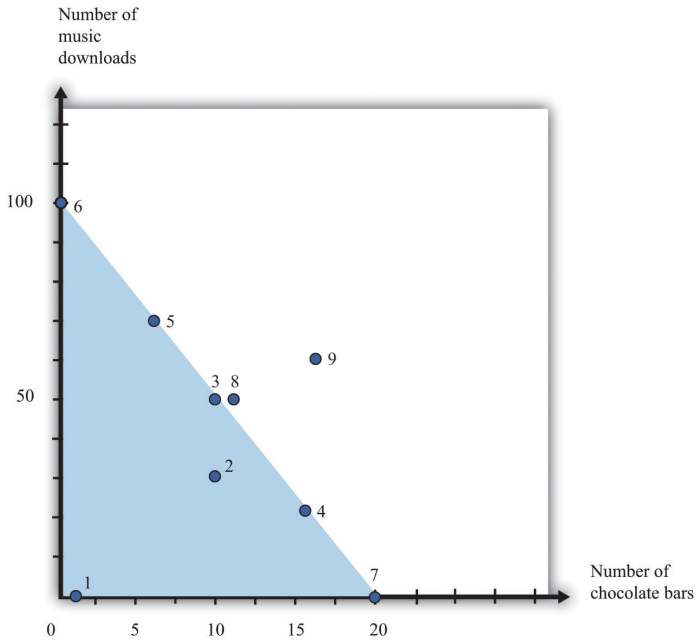


Figure 4.2.3: The Budget Set. The bundles that are affordable are in the ***budget set***, *shown here as a triangle.*

Every point—that is, every combination of downloads and chocolate bars—that lies on or inside this triangle is affordable. Points outside the triangle are not affordable, so they are not in the budget set.

What Have We Assumed?

We now have a picture of the budget set. However, you might be curious about whether we have sneaked in any assumptions to do this. This is a *Principles of Economics* book, so we must start by focusing on the basics. We do our best throughout the book to be clear about the different assumptions we make, including their importance.

- We have assumed that there are only two products. Once we have more than two products, we cannot draw simple diagrams. Beyond this, though, there is nothing special about our downloads-and-chocolate-bar example. We are using an example with two products simply because it makes our key points more transparent. We can easily imagine a version of Table 4.2.1 "Spending on Music Downloads and Chocolate Bars" with many more goods and services, even if we cannot draw the corresponding diagram.
- We assume that you cannot consume negative quantities of downloads or chocolate bars. In our diagram, this means that the horizontal and vertical axes give us two sides of the triangle. This seems reasonable: it is not easy to imagine consuming a negative quantity of chocolate bars. (If you started out with some chocolate bars and then sold them, this is similar to negative consumption.)
- An easier way to look at this is to add any money you get from selling goods or services to your income. Then we can focus on buying decisions only.
- By shading in the entire triangle, we suppose that you can buy fractional quantities of these products. For example, the bundle consisting of 17.5 downloads and 12.7 chocolate bars is inside the triangle, even though iTunes, for example, would not allow you to purchase half a song, and you are unlikely to find a store that will sell you 0.7 chocolate bars. For the most part, this is a technical detail that makes very little difference, except that it makes our lives much easier.
- We have supposed that the price per unit of downloads and chocolate bars is the same no matter how few or how many you choose to buy. In the real world, you may sometimes be able to get quantity discounts. For example, a store might have a “buy two get one free” offer. In more advanced courses in microeconomics, you will learn that we can draw versions of Figure 4.2.3 that take into account such pricing schemes.
- We assume no saving or borrowing. It is easy to include saving or borrowing in this story, though. We think of borrowing as being an addition to your income, and we think of saving as one more kind of spending. Thus if you borrow, the budget set is described by

$$total\ spending \leq disposable\ income + borrowing.$$

If you save, the budget set is described by

$$total\ spending\ +\ spending\ \leq\ disposable\ income.$$

The Budget Line

Continuing with our two-goods example, we know that

$$spending\ on\ chocolate\ =\ number\ of\ chocolate\ bars\ \times\ price\ of\ a\ chocolate\ bar$$

and

$$spending\ on\ downloads\ =\ number\ of\ downloads\ \times\ price\ of\ download.$$

When total spending is exactly equal to total disposable income, then

$$(number\ of\ chocolate\ bars\ \times\ price\ of\ a\ chocolate\ bar)\ +\ (number\ of\ downloads\ \times\ price\ of\ download)\ =\ disposable\ income.$$

Toolkit: Section 31.1

The budget line lists all the goods and services that are affordable, given prices and income, *assuming you spend all your income*.

The difference between the definitions of the budget set and the **budget line** is that there is an inequality in the budget set and an equality in the budget line:

$$total\ spending\ =\ disposable\ income.$$

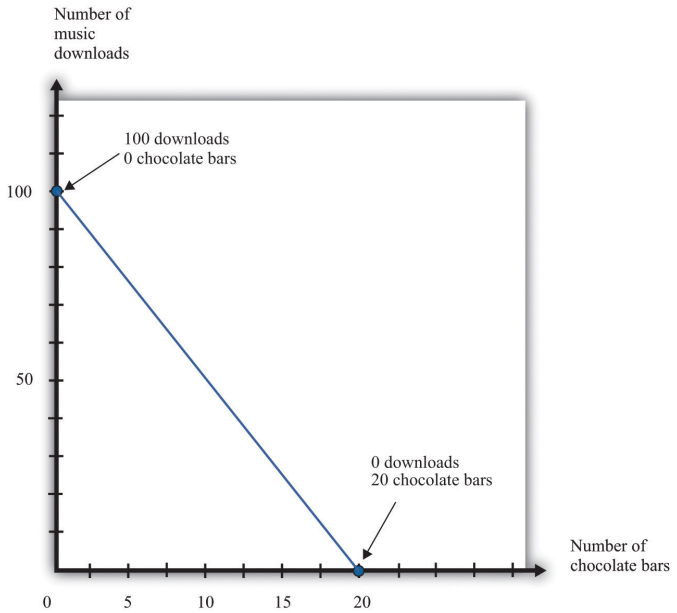


Figure 4.2.4: The Budget Line. The bundles that are exactly affordable are on the ***budget line***.

In the two-goods example, the budget line is the outside edge of the budget set triangle, as shown in Figure 4.2.4. What information do we need to draw the budget line? If we know both prices and the total amount of income, then this is certainly enough. In fact, we need only two pieces of information (not three) because basic mathematics tells us that it is enough to know two points on a line: once we have two points, we can draw a line. In practice, the easiest way to draw a budget line is to find the *intercepts*—the points on each axis. These correspond to how much you can obtain of each product if you consume 0 of the other. If you don't buy any chocolate bars, you have enough income to buy 100 downloads. If the number of chocolate bars is 0, then the budget line becomes

$$number\ of\ downloads\ \times\ price\ of\ download\ =\ disposable\ income,$$

so

Similarly, if the number of downloads is 0,

So we have two points on the budget line: (1) 0 chocolate bars and 100 downloads and (2) 0 downloads and 20 chocolate bars.

Another way to describe the budget line is to write the equation of the line in terms of its intercept (on the vertical axis) and its slope:To derive this equation, go back to the budget line and divide both sides by the price of a download:Rearranging, we get the equation in the text.

The intercept is which answers the following question: “How many downloads can you obtain if you buy no chocolate?” As we have already seen, this is 100 in our example.

The slope is

which answers the following question: “What is the rate at which you can trade off downloads for chocolate bars?” In our example, this is −5. If you give up 1 chocolate bar, you will have an extra $5 (the price of a chocolate bar), which allows you to buy 5 more downloads.

The negative slope of the budget line says that to get more downloads, you must give up some chocolate bars. The *cost* of getting more downloads is that you no longer have the *opportunity* to buy as many chocolate bars. More generally, economists say that the **opportunity cost** of an action is what you must give up to carry out that action. Likewise, to get more chocolate bars, you must give up some downloads. The opportunity cost of buying a chocolate bar is that you do not have that the money available to purchase downloads. The idea of opportunity cost pervades economics.

You may well have heard the following quotation that originated in economics: “There is no such thing as a free lunch.” This statement captures the insight that everything has an opportunity cost, even if it is not always obvious who pays. Economists’ habit of pointing out this unpleasant truth is one reason that economics is labeled “the dismal science.”Although economists may dislike this characterization of their profession, they can take pride in its origin. The term was coined by Thomas Carlyle about 150 years ago, in the context of a debate about race and slavery. Carlyle criticized famous economists of the time, such as John Stuart Mill and Adam Smith, who argued that some nations were richer than others not because of innate differences across races but because of economic and historical factors. These economists argued for the equality of people and supported the freedom of slaves.

We said that a goal of this chapter is to help you make good decisions. One ingredient of good decision making is to understand the trade-offs that you face. Are you thinking of buying a new $200 mobile phone? The cost of that phone is best thought of, not as a sum of money, but as the other goods or services that you could have bought with that $200. Would you rather have 200 new songs for your existing phone instead? Or would you prefer 20 trips to the movies, 40 ice cream cones, or $200 worth of gas for your car? Framing decisions in this way can help you make better choices.

Your Preferences

Your choices reflect two factors. One is what you can afford. The budget set and the budget line are a way of describing the combinations of goods and services you can afford. The second factor is what you like, or—to use the usual economic term—your *preferences*.

Economists don’t pretend to know what makes everyone happy. In our role as economists, we pass no judgment on individual tastes. Your music downloads might be Gustav Mahler, Arctic Monkeys, Eminem, or Barry Manilow. But we think it is reasonable to assume three things about the preferences that underlie your choices: (1) more is better, (2) you can choose, and (3) your choices are consistent.

More Is Better

Economists think that you are never satisfied. No matter how much you consume, you would always like to have more of something. Another way of saying this is that every good is indeed “good”; having more of something will never make you less happy. This assumption says nothing more than people don’t usually throw their income away. Even Bill Gates is not in the habit of burning money.

“More is better” permits us to focus on the budget line rather than the budget set. In Figure 4.2.3, you will not choose to consume at a point inside the triangle of the budget set. Instead, you want to be on the edge of the triangle—that is, on the budget line itself. Otherwise, you would be throwing money away. It also allows us to rank some of the different bundles in Table 4.2.1 "Spending on Music Downloads and Chocolate Bars". For example, we predict you would prefer to have bundle 3 rather than bundle 2 because it

has the same number of chocolate bars and more downloads. Likewise, we predict you would prefer bundle 8 to bundle 3: bundle 8 has the same number of downloads as bundle 3 but more chocolate bars.

By the way, we are not insisting that you must *eat* all these chocolate bars. You are always allowed to give away or throw away anything you don't want. Equally, the idea that more is better does not mean that you might not be sated with one particular good. It is possible that one more chocolate bar would make you no happier than before. Economists merely believe that there is always *something* that you would like to have more of.

"More is better" does *not* mean that you necessarily prefer a bundle that costs more. Look at bundles 7 and 9. Bundle 7 contains 0 downloads and 20 chocolate bars; it costs $100. Bundle 9, which contains 70 downloads and 16 chocolate bars, costs $150. Yet someone who loves chocolate bars and has no interest in music would prefer bundle 7, even though its market value is less.

You Can Choose

Economists suppose that you can always make the comparison between any two bundles of goods and services. If you are presented with two bundles—call them A and B—then the assumption that "you can choose" says that one of the following is true:

- You prefer A to B.
- You prefer B to A.
- You are equally happy with either A or B.

Look back at Table 4.2.1 "Spending on Music Downloads and Chocolate Bars". The assumption that "you can choose" says that if you were presented with any pair of bundles, you would be able to indicate which one you liked better (or that you liked them both equally much). This assumption says that you are never paralyzed by indecision.

"More is better" allows us to draw some conclusions about the choices you would make. If we gave you a choice between bundle 3 and bundle 8, for example, we know you will choose bundle 8. But what if, say, we presented you with bundle 4 and bundle 5? Bundle 5 has more downloads, but bundle 4 has more chocolate bars. "You can choose" says that, even though we may not know *which* bundle you would choose, you are capable of making up your mind.

Your Choices Are Consistent

Finally, economists suppose that your preferences lead you to behave consistently. Based on Table 4.2.1 "Spending on Music Downloads and Chocolate Bars", suppose you reported the following preferences across combinations of downloads and chocolate bars:

- You prefer bundle 3 to 4.
- You prefer bundle 4 to 5.
- You prefer bundle 5 to 3.

Each choice, taken individually, might make sense, but all three taken together are not consistent. They are contradictory. If you prefer bundle 3 to bundle 4 and you prefer bundle 4 to bundle 5, then a common-sense interpretation of the word "prefer" means that you should prefer bundle 3 to bundle 5.

Consistency means that your preferences must not be contradictory in this way. Put another way, if your preferences are consistent and yet you made these three choices, then at least one of these choices must have been a mistake—a bad decision. You would have been happier had you made a different choice.

Your Choice

We have now looked at your opportunities, as summarized by the budget set, and also your preferences. By combining opportunities and preferences, we obtain the economic approach to individual decision making. Economists make a straightforward assumption: they suppose you look at the bundles of goods and services you can afford and choose the one that makes you happiest. If the claims we made about your preferences are true, then you will be able to find a "best" bundle of goods and services, and this bundle will lie on the budget line. We know this because (1) you can compare any two points and (2) your preferences will not lead you to go around in circles.

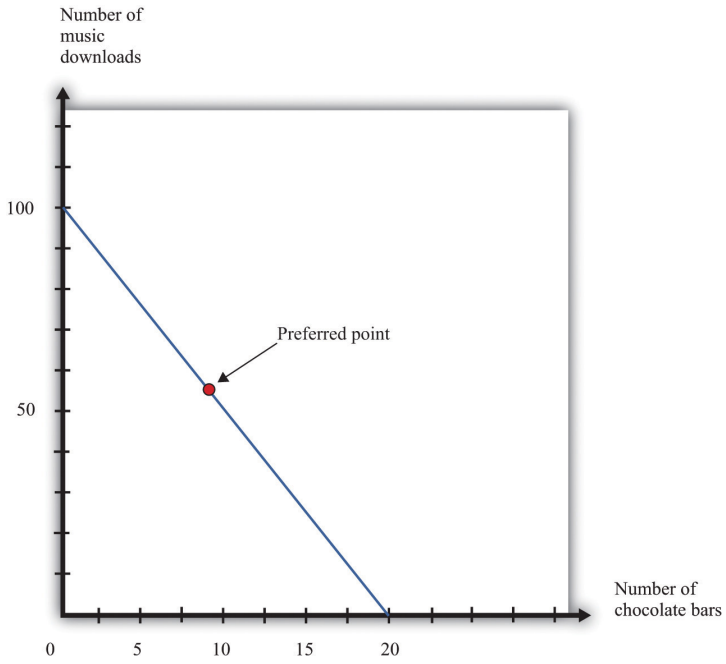


Figure 4.2.5: Choosing a Preferred Point on the Budget Line

An individual's preferred point reflects opportunities, as given by the budget line, and preferences. The preferred point will lie on the budget line, not inside, because of the assumption that more is better.

In Figure 4.2.5, we indicate an example of an individual's preferred point. The preferred point is on the budget line and—by definition—is the best combination for the individual that can be found in the budget set. At the preferred point, the individual cannot be better off by consuming any other *affordable* bundle of goods and services.

There is one technical detail that we should add. It is possible that an individual might have more than one preferred point. There could be two or more combinations on the budget line that make an individual equally happy. To keep life simple, economists usually suppose that there is only a single preferred point, but nothing important hinges on this.

Rationality

Economists typically assume **rationality** of decision makers, which means that people can do the following:

- evaluate the opportunities that they face
- choose among those opportunities in a way that serves their own best interests

Is this a good assumption? Are people really as rational as economists like to think they are? We would like to know if people's preferences do satisfy the assumptions that we have made and if people behave in a consistent way. If we could hook someone up to a machine and measure his or her preferences, then we could evaluate our assumptions directly. Despite advances in neurobiology, our scientific understanding has not reached that point. We see what people do, not the preferences that lie behind these choices. Therefore, one way to evaluate the economic approach is to look at the choices people make and see if they are consistent with our assumptions.

Imagine you have an individual's data on download and chocolate bar consumption over many months. Also, suppose you know the prices of downloads and chocolate bars each month and the individual's monthly income. This would give you enough information to construct the individual's budget sets each month and look for behavior that is inconsistent with our assumptions. Such inconsistency could take different forms.

- She might buy a bundle of goods inside the budget line and throw away the remaining income.
- In one month, she might have chosen a bundle of goods—call it bundle A—in preference to another affordable bundle—call it bundle B. Yet, in another month, that same individual might have chosen bundle B when she could also have afforded bundle A.

The first option is inconsistent with our idea that "more is better." As for the second option, it is generally inconsistent to prefer bundle A over bundle B at one time yet prefer bundle B over bundle A at a different time. (It is not *necessarily* inconsistent,

however. The individual might be indifferent between bundles A and B, so she doesn't care which bundle she consumes. Or her preferences might change from one month to the next.)

Inconsistent Choices

Economists are not the only social scientists who study how we make choices. Psychologists also study decision making, although their focus is different because they pay more attention to the processes that lie behind our choices. The decision-making process that we have described, in which you evaluate each possible option available to you, can be cognitively taxing. Psychologists and economists have argued that we therefore often use simpler rules of thumb when we make decisions. These rules of thumb work well most of the time, but sometimes they lead to biases and inconsistent choices. This book is about economics, not psychology, so we will not discuss these ideas in too much detail. Nevertheless, it is worth knowing something about how our decision making might go awry.

On occasion, we make choices that are apparently inconsistent. Here are some examples.

The endowment effect. Imagine you win a prize in a contest and have two scenarios to consider:

1. The prize is a ticket to a major sporting event taking place in your town. After looking on eBay, you discover that equivalent tickets are being bought and sold for $500.
2. The prize is $500 cash.

Rational decision makers would treat these two situations as essentially identical: if you get the ticket, you can sell it on eBay for $500; if you get $500 cash, you can buy a ticket on eBay. Yet many people behave differently in the two situations. If they get the ticket, they do not sell it, but if they get the cash, they do not buy the ticket. Apparently, we often feel differently about goods that we actually have in our possession compared to goods that we could choose to purchase.

Mental states. We may be in a different mental state when we buy a good from when we consume it. If you are hungry when you go grocery shopping, then you may buy too much food. When we buy something, we have to predict how we will be feeling when we consume it, and we are not always very good at making these predictions. Thus our purchases may be different, depending on our state of mind, even if prices and incomes are the same.

Anchoring. Very often, when you go to a store, you will see that goods are advertised as "on sale" or "reduced from" some price. Our theory suggests that people simply look at current prices and their current income when deciding what to buy, in which case they shouldn't care if the good used to sell at a higher price. In reality, the "regular price" serves as an anchor for our judgments. A higher price tends to increase our assessment of how much the good is worth to us. Thus we may make inconsistent choices because we sometimes use different anchors.

How should we interpret the evidence that people are—sometimes at least—not quite as rational as economics usually supposes? Should we give up and go home? Not at all. Such findings deepen our understanding of economic behavior, but there are many reasons why it is vital to understand the behavior of rational individuals.

1. Economics helps us make better decisions. The movie *Heist* has dialogue that sums up this idea:

D. A. Freccia:	You're a pretty smart fella.
Joe Moore:	Ah, not that smart.
D. A. Freccia:	[If] you're not that smart, how'd you figure it out?
Joe Moore:	I tried to imagine a fella smarter than myself. Then I tried to think, "what would he do?"

   Most of us are "not that smart"; that is, we are not smart enough to determine what the rational thing to do is in all circumstances. Knowing what someone smarter would do can be very useful indeed.The quote comes from the Internet Movie Database (http://www.imdb.com). We first learned of the scene from B. Nalebuff and I. Ayres, *Why Not?* (Boston: Harvard Business School Press, 2003), 46. Further, if we understand the biases and mistakes to which we are all prone, then we can do a better job of recognizing them in ourselves and adjusting our behavior accordingly.

2. Rationality imposes a great deal of discipline on our thinking as economists. If we suppose that people are irrational, then anything is possible. A better approach is to start with rational behavior and then see if the biases that psychologists and economists have identified are likely to alter our conclusions in a major way.

3. Economics has a good track record of prediction in many settings. A lot of the time, even if not all the time, the idea that people behave rationally seems more right than wrong.

More Complicated Preferences

People may be rational yet have more complicated preferences than we have considered.

Fairness. People sometimes care about fairness and so may refuse to buy something because the price seems unfair to them. In one famous example, people were asked to imagine that they are on the beach and that a friend offers to buy a cold drink on their behalf.See Richard Thaler, "Mental Accounting and Consumer Choice, *Marketing Science* 4 (1985): 199–214. They are asked how much they are willing to pay for this drink. The answer to this question should not depend on where the drink is purchased. After all, they are handing over some money and getting a cold drink in return. Yet people are prepared to pay more if they know that the friend is going to buy the drink from a hotel bar rather than a local corner store. They think it is reasonable for hotels to have high prices, but if the corner store charged the same price as the hotel, people think that this is unfair and are unwilling to pay.

Altruism. People sometimes care not only about what they themselves consume but also about the well-being of others. Such altruism leads people to give gifts, to give to charity, to buy products such as "fair-trade" coffee, and so on.

Relative incomes. Caring about the consumption of others can take more negative forms as well. People sometimes care about whether they are richer or poorer than other people. They may want to own a car or a barbecue grill that is bigger and better than that of their neighbors.

More complicated preferences such as these are not irrational, but they require a more complex framework for decision making than we can tackle in a *Principles of Economics* book.We say more about some of these ideas in Chapter 13.

Key Takeaways

- The budget set consists of all combinations of goods and services that are affordable, and the budget line consists of all combinations of goods and services that are affordable *if you spend all your income*.
- The opportunity cost of an action (such as consuming more of one good) is what must be given up to carry out that action (consuming less of some other good).
- Your choices reflect the interaction between what you can afford (your budget set) and what you like (your preferences).
- Economists think that most people prefer having more to having less, are able to choose among the combinations in their budget set, and make consistent choices.
- Rational agents are able to evaluate their options and make choices that maximize their happiness.

Exercises

1. Suppose that all prices and income were converted into a different currency. For example, imagine that prices were originally in dollars but were then converted to Mexican pesos. Would the budget set change? If so, explain how. If not, explain why not.
2. Assume your disposable income is $100, the price of a music download is $2, and the price of a chocolate bar is $5. Redo Table 4.2.1 "Spending on Music Downloads and Chocolate Bars". Find (or create) three combinations of chocolate bars and downloads that are on the budget line. Find a combination that is not affordable and another combination that is in the budget set but not on the budget line.
3. What is the difference between your budget set and your budget line?

This page titled 4.2: Individual Decision Making- How You Spend Your Income is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

4.3: Individual Demand 4.3: Individual DemandEdit Update titlecancel

Learning Objectives

1. What is a demand curve, and what is the law of demand?
2. What is the decision rule for choosing how much to buy of two different goods?
3. What is the decision rule for choosing to buy a single unit of a good?

Now that we have a framework for thinking about your choices, we can now explain one of the most fundamental economic ideas: demand. Here we focus on the demand of a single individual.In Chapter 6, we develop the idea of the market demand curve, which combines the demands of many individuals. We use two different ways of thinking about your demand for a good or a service. One approach builds on the idea of the budget set. The other focuses on how much you would be willing to pay for a good or service. In combination, they give us a detailed understanding of how economic decisions are made.

Individual Demand for a Good

As you visit stores at different times, you undoubtedly notice that the prices of goods and services change. At the same time, your income may also change from one month to the next. So if we were to look at your budget set monthly, we would typically find it changing from one month to the next. We would then expect that you would choose different combinations of goods and services from one month to the next.

To keep things simple, suppose we are still in a world of two goods—downloads and chocolate bars—and that you do no saving. We will describe your demand for chocolate bars. (If you like, you can think of downloads as representing all the other goods and services you consume.) Given prices and your income, you pick the best point on the budget line. Look again at the "preferred" point in Figure 4.3.5. One way to interpret this point is that it tells us how many chocolate bars you will buy, given your income and given the price of a chocolate bar and other goods. Using this as a reference point, we now ask how your choice will change as income changes and then as the price of a chocolate bar changes.

Changes in Income

Imagine that your income increases. Figure 4.3.1 shows what happens to the budget line. Higher income means that you can afford to buy more chocolate bars and more downloads, so the budget line shifts outward. The slope of the budget line is unchanged because there is no change in the price of a chocolate bar relative to downloads.

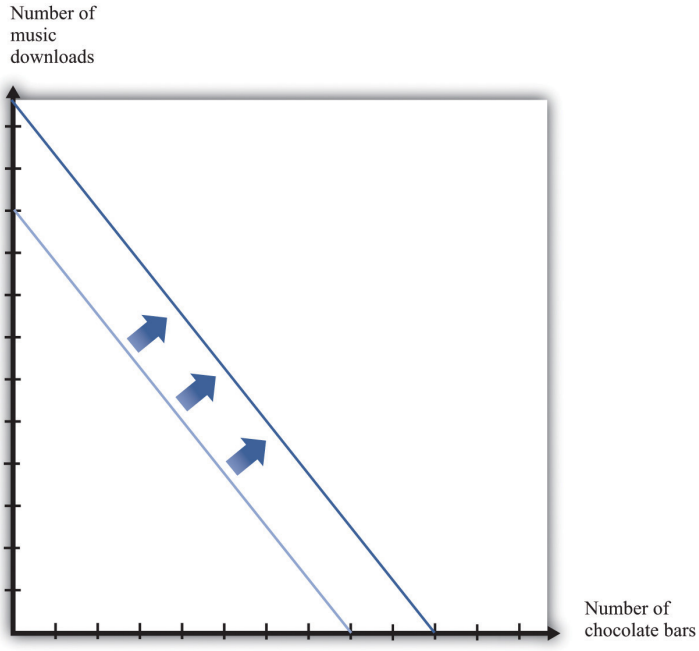


Figure 4.3.1: An Increase in Income. An increase in income shifts the budget line outward.

What happens to your consumption of chocolate bars? There are two possibilities (Figure 4.3.2): the increase in income leads you to consume either more chocolate bars or fewer chocolate bars. Both are plausible, and either is possible. (Of course, you might also choose exactly the same amount as before.) We might think that the more normal case is that higher income would lead to higher chocolate bar consumption. If a good has a property that you will consume more of it when you have higher income, we call it a **normal good**.

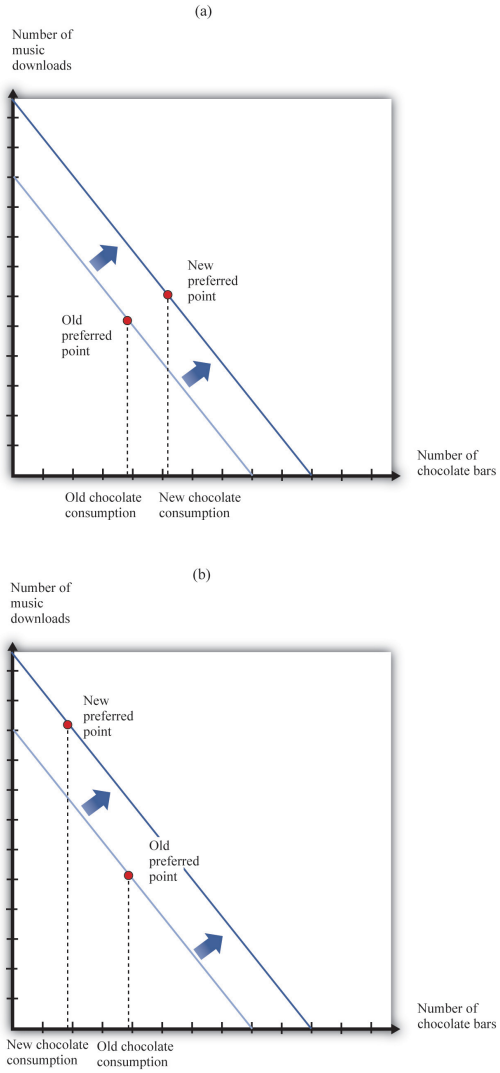


Figure 4.3.2: The Consequences of an Increase in Income. In response to an increase in income, two things are possible: the consumption of chocolate bars may increase (a) or decrease (b).

Under some circumstances, higher income leads to *lower* consumption. For example, suppose you are surviving in college on a diet consisting of mostly instant noodles. When you graduate college and have higher income, you can afford better things to eat, so you will probably consume a smaller quantity of instant noodles. Economists call such products **inferior goods**. If a particular product exists in several qualities (cheap versus expensive cuts of meat, for example), we often find that the low-quality version is an inferior good. A good might be normal for one consumer and inferior for another. More precisely, therefore, we say that a *good* is inferior if, on average, higher income leads to lower consumption.

Economists make a further distinction among different kinds of normal goods. If you spend a *larger fraction of your income* on a particular good as your income increases, then we say that the good is a **luxury good**. Another way of saying this is that for a luxury good, the percentage increase in consumption is bigger than the percentage increase in income.

We can also define these ideas in terms of the **income elasticity of demand**, which is a measure of how sensitive demand is to changes in income. For an inferior good, the income elasticity of demand is negative: higher income leads to lower consumption.

For a normal good, the income elasticity of demand is positive. And for a luxury good, the income elasticity of demand is greater than one.

> **Toolkit: Section 31.2**
>
> For more discussion of normal, inferior, and luxury goods, see the toolkit.

The distinctions among these different kinds of goods are crucial for managers of firms. To predict sales, managers need to know whether the products they are selling are normal, inferior, or luxury goods. Firms that sell inferior goods tend to do well when the economy as a whole is doing poorly and vice versa. By contrast, firms selling luxury goods will do particularly poorly when the economy as a whole is performing poorly.

Changes in Price

Now we look at what happens if there is a change in the price of a chocolate bar. Suppose the price decreases. First, let us analyze what this means in terms of our picture. Remember that the intercepts of the budget line tell you how much you can have of one good if you consume none of the other. If you consume no chocolate bars, then a decrease in the price of a chocolate bar has no effect on your consumption. You can consume exactly the same number of downloads as before. The intercept on the vertical axis therefore does not move. However, a decrease in the price of a chocolate bar means that, if you consume only chocolate bars, then you can have more than before. The intercept on the horizontal axis moves outward. One way to see this is to remember that this intercept is given by

Figure 4.3.3 illustrates a decrease in the price of a chocolate bar. The new budget line lies outside the old budget line. Any bundle you could have bought at the old prices is still affordable now, and you can also get more. A decrease in the price of a chocolate bar, in other words, makes you better off. In addition, the slope of the budget line changes: it is flatter than before. The slope of the budget line reflects the way in which the market allows you to trade chocolate bars for other products. If you choose to consume fewer downloads, the reduction in the price of a chocolate bar means that you get relatively more chocolate bars in exchange.

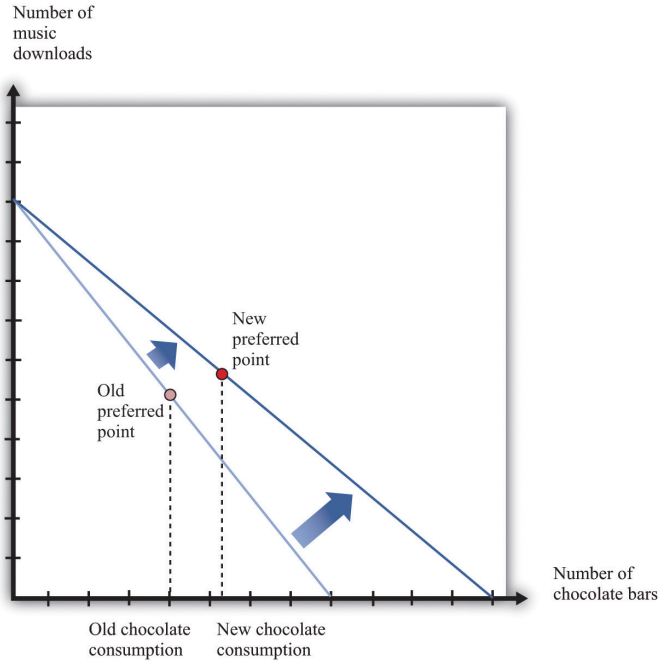


Figure 4.3.3: A Decrease in the Price of a Chocolate Bar. A decrease in the price of a chocolate bar causes the budget line to *rotate.* *The result is an increase in the quantity of chocolate bars consumed.*

Figure 4.3.3 also shows a new consumption point. In response to the decrease in the price of a chocolate bar, we see an increase in the consumption of chocolate bars. The idea that people almost always consume more of a good when its price decreases is one of the fundamental ideas of economics. Indeed, it is sometimes called the **law of demand**. It is certainly intuitive that lower prices lead people to consume more. There are two reasons why we expect to see this result.

First, if a good (for example, chocolate bars) decreases in price, it becomes cheaper *relative to other goods*. Its opportunity cost—that is, the amount of other goods you must give up to get a chocolate bar—has decreased. The lower opportunity cost means that there is a **substitution effect** away from other goods and toward chocolate bars. Second, a decrease in the price of a chocolate bar also means that you can afford more of *everything*, including chocolate bars. Provided chocolate bars are a normal good, this **income effect** will also lead you to want to consume more chocolate bars. If chocolate bars are inferior goods, the income effect leads you to want to consume fewer chocolate bars.

In Figure 4.3.3, the substitution effect is reflected by the fact that the budget line changes slope. The flatter budget line tells us that the opportunity cost of chocolate bars in terms of downloads has decreased. The income effect shows up in the fact that the new budget set includes the old budget set. You can consume your previous bundle of goods and still have some income left over to buy more.

Based on the idea of the law of demand, we can construct your **individual demand curve** for chocolate bars. We do so by drawing the budget set for each different price of a chocolate bar, seeing how much you buy, and then plotting this data. For example, we might find that your purchases of chocolate bars look like those in Table 4.3.1 "Demand for Chocolate Bars". If we plot these points on a graph and then “fill in the gaps,” we get a diagram like Figure 4.3.4. This is your demand curve for chocolate bars. It tells how many chocolate bars you would purchase at any given price. The law of demand means that we expect this curve to slope downward. If the price increases, you consume less. If the price decreases, you consume more.

Table 4.3.1: Demand for Chocolate Bars

Price per Bar ($)	Quantity of Chocolate Bars Bought
1	12
2	6
3	4
4	3
5	2.4
6	2

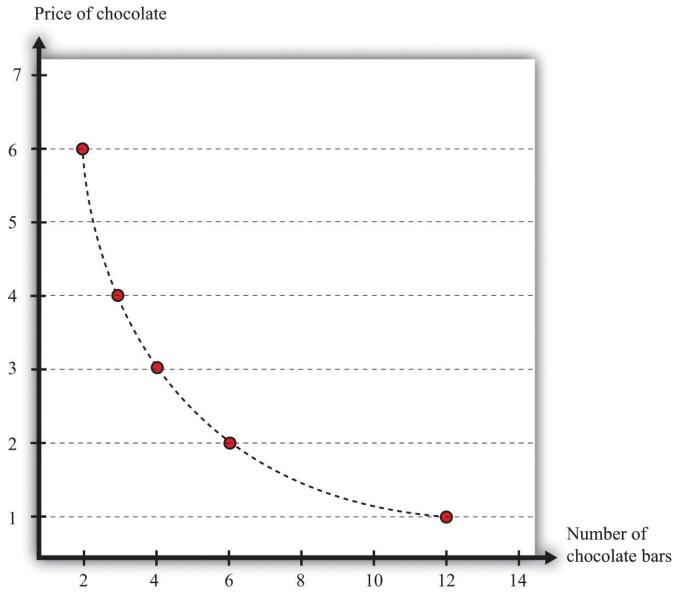


Figure 4.3.4: The Demand Curve. contains an example of some observations on demand. At different prices from $1 to $6, we see the number of chocolate bars purchased. If we fill in the gaps, we obtain a demand curve.

Toolkit: Section 31.1

The *individual demand curve* is drawn on a diagram with the price of a good on the vertical axis and the quantity demanded on the horizontal axis. It is drawn for a given level of income.

We must be careful to distinguish between movements along the demand curve and shifts in the demand curve. Suppose there is a change in the price of a chocolate bar. Then, as we explained earlier, the budget line shifts, and the quantity demanded of both chocolate bars and downloads will change. This appears in Figure 4.3.4 as a *movement along the demand curve*. For example, if the price of a chocolate bar decreases from $4 to $3, then, as in Table 4.3.1 "Demand for Chocolate Bars", the quantity demanded increases from three bars to four bars. The demand curve does not change; we simply move from one point on the line to another.

If a change in anything other than the price of a chocolate bar causes you to change your consumption of chocolate bars, then there is a *shift in the demand curve*. For example, suppose you get a pay raise at your job, so you have more income. As long as chocolate bars are a normal good, this increase in income will cause your demand curve for chocolate bars to shift outward. This means that, at any given price, you will buy more of the good when income increases. In the case of an inferior good, an increase in income will cause the demand curve to shift inward. You will buy less of the good when income increases. We illustrate these two cases in Figure 4.3.5.

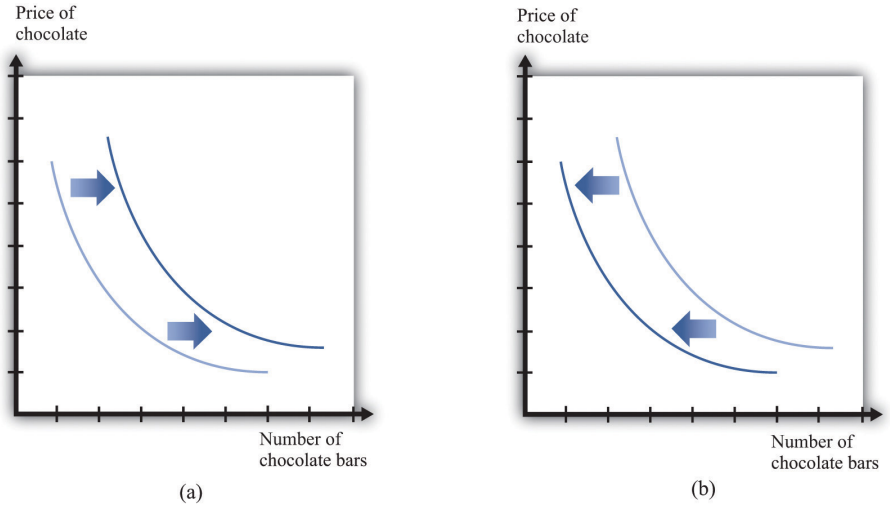


Figure 4.3.5: Shifts in the Demand Curve: Normal and Inferior Goods. (a) If income increases and chocolate bars are a normal good, then the individual demand curve will shift to the right. At every price, a greater quantity of chocolate bars is demanded. (b) If income increases and chocolate bars are inferior goods, then the individual demand curve will shift to the left. In the event of a decrease in income, the two cases are reversed.

Exceptions to the Law of Demand

The law of demand is highly intuitive and is supported by lots of research for all sorts of different goods and services. We can take it as a reliable fact that in almost all circumstances, the demand curve will indeed slope downward. Yet we might still wonder if there are any exceptions, any cases in which the demand curve slopes upward. There are indeed a few such exceptions.

- **Giffen goods.** We explained that the law of demand comes from both a substitution effect and—for normal goods—an income effect. But for inferior goods, the income effect acts in the opposite direction to the substitution effect: a decrease in price makes people better off, which is an incentive to consume less of the good. Theoretically, it is possible for this income effect to be stronger than the substitution effect. Although conceivable, Giffen goods are extremely rare; indeed, economists are unsure if they actually exist outside of textbooks. For the income effect to overwhelm the substitution effect, the good in question must form a very large part of the overall consumption bundle. This might arise in extremely poor economies where people spend a large part of their income on a staple food, as was found in a recent experiment conducted in rural China.See Robert Jensen and Nolan Miller, “Giffen Behavior: Theory and Evidence” (Harvard University, John F. Kennedy School of Government Working Paper RWP 07-030, July 2007). The researchers gave families subsidies to buy rice, making it cheaper, and found that consumption of rice did indeed decrease.
- **Status goods.** Some luxury products are purchased mainly for their status appeal—for example, Rolls-Royce automobiles, Louis Vuitton handbags, and Gucci shoes. The high prices of these goods contribute to their exclusivity. Price is an *attribute* of

a good, so a higher price can make a good seem more, not less, attractive. Again, it is theoretically possible that this could lead the demand curve to slope upward, at least for some range of prices. Although high prices increase the appeal of status goods, it is rare for this effect to be strong enough to outweigh the more basic income and substitution effects.

- **Judging quality by price.** Implicitly, we have supposed that people are well informed about the products they buy. In many cases, though, we must purchase goods and services with only imperfect knowledge of their quality. In this situation, we may use price as an indicator of the quality of a good. Of course, marketers are well aware that we do this and will often try to use price as a signal of the quality of their brand or products. The upshot is that people may be more willing to buy at a higher price.

In these three situations, it is conceivable that we might observe a higher price being associated with a higher quantity of the good being purchased. You should not overestimate the significance of these cases, however, because (1) most products do not fall into any of these categories, and (2) the substitution effect is still in operation for all goods, as is the income effect for all normal goods.

Changes in Other Prices

So far, we have said that the number of chocolate bars you want to buy is affected by income and the price of a chocolate bar. Changes in the prices of other goods also have an impact. In general, an increase in a price of another good could cause the demand for chocolate bars to increase or decrease.

Goods are **substitutes** if an increase in the price of one good leads to increased consumption of the other good. CDs and music downloads are one example: if the price of CDs increases, you will obtain more music through downloads.

Goods are **complements** if an increase in the price of one good leads to decreased consumption of the other good. For example, DVDs and DVD players are complements. If the price of DVD players decreases, more people will buy DVD players. As a result, more people will want to buy DVDs.

Toolkit: Section 31.2

To learn more about substitutes and complements, see the toolkit for formal definitions. (These definitions are presented in terms of the *cross-elasticity of demand*, which is a measure of how responsive the quantity demanded is to changes in the price of another good.)

The Valuation Approach to Demand

There is another way of thinking about demand. Instead of focusing attention on the budget set and the budget line, we can think more directly about your preferences. Imagine you are asked the following question in an interview:

What is the *maximum* amount you would be willing to pay for one chocolate bar?

In answering this question, you should not worry about what would be a reasonable or fair price for a chocolate bar or even about the price at which this chocolate bar might actually be available. You should simply decide how much you want chocolate bars. For example, you might think that you would really like to have at least one chocolate bar, so you would be willing to pay up to $12 for one bar—that is, you would be happier with one more chocolate bar and up to $12 less in income. This is your **valuation** of one chocolate bar.

Now we ask you some more questions.

What is the *maximum* amount that you would be willing to pay for two chocolate bars? Three chocolate bars?…

Perhaps you decide you would be willing to pay $18 for two bars, $22 for three bars, and so on. If we kept asking such questions, we might get the first two columns of Table 4.3.2 "Valuation and Marginal Valuation". We can also plot this valuation (part [a] of Figure 4.3.6). Your valuation is increasing; you always like having more chocolate bars because “more is better.”

Table 4.3.2: Valuation and Marginal Valuation

Quantity of Chocolate Bars Bought	Valuation ($)	Marginal Valuation ($)
0	0.00	
1	12.00	12.00
2	18.00	6.00

Quantity of Chocolate Bars Bought	Valuation ($)	Marginal Valuation ($)
3	22.00	4.00
4	25.00	3.00
5	27.40	2.40
6	29.40	2.00

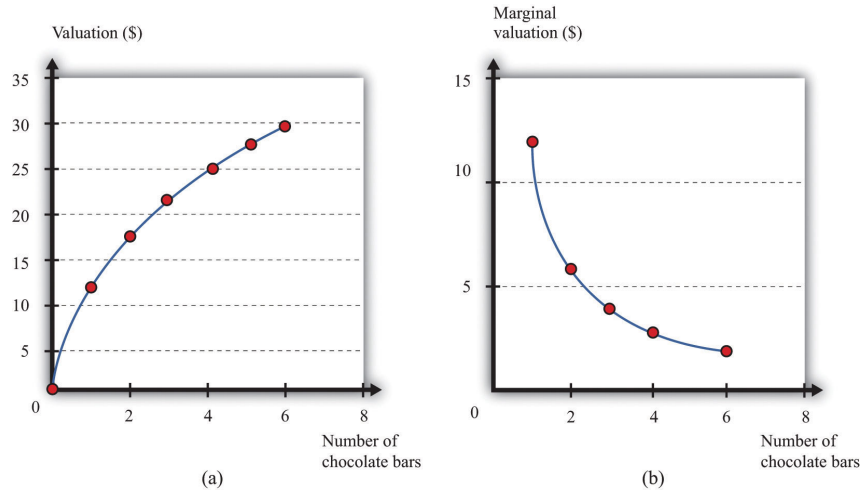


Figure 4.3.6: Valuation of a Good. Your ***valuation*** *of a good is the maximum amount that you would be willing to pay, purely on the basis of your desire for the good.*

Because you are willing to pay $12 for one bar and $18 for two bars, we know you would be willing to pay an *additional* $6 for the second chocolate bar. Similarly, if you have two chocolate bars, you would be willing to pay an additional $4 for a third bar. We call the *change* in your valuation your **marginal valuation** (see the third column of Table 4.3.2 "Valuation and Marginal Valuation", which is graphed in part [b] of Figure 4.3.6). Notice that marginal valuation decreases as the quantity of chocolate bars increases. The change in your valuation gets smaller as you obtain more chocolate bars. We can see the same thing in part (a) of Figure 4.3.6 from the fact that the valuation curve gets flatter as the quantity of chocolate bars increases.

Toolkit: Section 31.1

An individual's valuation of some quantity of a product is the maximum amount the individual would be willing to pay to obtain that quantity. An individual's marginal valuation of some good is the maximum amount the individual would be willing to pay to obtain one *extra* unit of that product.

It usually seems reasonable to think that marginal valuations will indeed decrease in this way. If you don't have any chocolate bars, then the first bar is worth a lot to you—$12 in our example. But if you already have five chocolate bars, then the sixth bar is worth only $2 to you. As you obtain more and more of any given product, each additional unit is less and less valuable. For most people, we expect that most products will exhibit such *diminishing marginal valuation*.

Part (b) of Figure 4.3.6 may seem familiar. It is the demand curve for chocolate bars you saw previously in Figure 4.3.4. This is not an accident or a coincidence. There is a simple decision rule to tell you how much you should buy: when your marginal valuation is greater than the price, you should buy more of the good, stopping only when the marginal valuation of the good has dropped to the level of the price. For example, suppose that chocolate bars are selling for $3.99. You should definitely buy the first chocolate bar, because it is worth $10 to you and will cost you only $3.99. You should buy the second bar as well because it is worth an additional $8 to you; likewise you should buy the third and fourth bars. You don't buy the fifth bar because it is worth only $3 to you, which is less than what it costs. Thus your decision rule is as follows:

- Buy until the marginal valuation of a good equals the price of the good.

Because the demand curve, by definition, tells you how much you buy at a given price, it is the same as the marginal valuation curve.

Combining the Two Approaches to Demand

We have presented two different ways of thinking about consumer decisions, but the underlying choice is the same. To see how the two approaches are linked, rewrite the decision rule for chocolate bars as buy until

You have a similar decision rule for downloads: buy until

Combining these two equations, we see that

which we can rearrange as

The ratio on the right-hand side of this expression should look familiar. Earlier, we found that the slope of the budget line is

So the marginal valuation of a chocolate bar divided by the marginal valuation of downloads equals minus the slope of the budget line.

What does this mean? The budget line tells us the rate at which the market allows you to trade off chocolate bars for downloads. If you consume one fewer chocolate bar, the number of dollars you will get is equal to the price of a chocolate bar. These dollars will buy you downloads. The ratio of marginal valuations describes how you view the trade-off between chocolate bars and downloads. If you are making good decisions about how to spend your money, then the rate at which you are happy to trade off chocolate bars for downloads equals the rate at which the market allows you to make such trades. If this were not true, then you could make yourself happier by choosing a different bundle on the budget line.

Making Decisions at the Margin

To make good decisions, you need to understand the trade-offs you are making. To put it another way, you need to recognize that every purchase has an opportunity cost, which is summarized by the budget line. If you want more chocolate bars, you must consume less of something else. You also need to find the right point on the budget line—the point that makes you happiest. Most of the time, economists simply assume that you are able to make this decision correctly on the basis of the three assumptions about your preferences that we introduced earlier.

You can also use this theory to help you think about the decisions you make. Suppose you are facing the budget line we discussed earlier and plan to buy 8 chocolate bars and 60 downloads (as in Figure 4.3.5). In principle, you need to compare that bundle with every other combination on the budget line. In practice, it is enough—most of the time at least—to compare it with nearby bundles. For example, if you prefer this bundle to 7 chocolate bars and 65 downloads, and you also prefer this bundle to 9 chocolate bars and 55 downloads, then you can be reasonably confident that you have found the best bundle. If a small change won't make you happier, then neither will a large change.

Unit Demand

So far we have considered situations where you might buy multiple units of a good—for example, 20 music downloads or 5 chocolate bars. To keep things simple, we also supposed that you could buy fractional amounts, such as 20.7 downloads, or 3.25 chocolate bars. This assumption gives a decision rule for purchase: buy until marginal valuation equals price.

Some purchase decisions are better thought of as "buy or don't buy." Large, infrequent purchases fall into this category. Think, for example, about the decision to buy a new car, a new microwave, or an expensive vacation. You won't buy five microwaves because they are cheap. The decision rule for purchase is even easier in this case: buy if your valuation of the good exceeds the price of the good. In fact, this is really no different from our earlier decision rule. Because you are only ever thinking about buying one unit, your valuation and your marginal valuation are the same thing (look back at the first two rows of Table 4.3.2 "Valuation and Marginal Valuation"). And because in this case it does *not* make sense to suppose you can buy fractional amounts of the good, you cannot keep buying until your marginal valuation decreases all the way to the price.

If a buyer is interested in purchasing one and only one unit of a good, the **unit demand curve** tells us the price at which he is willing to buy. Below his valuation, he buys the good. Above his valuation, he does not buy the good. A unit demand curve is shown in Figure 4.3.7. In this example, the buyer's valuation is $3,000.

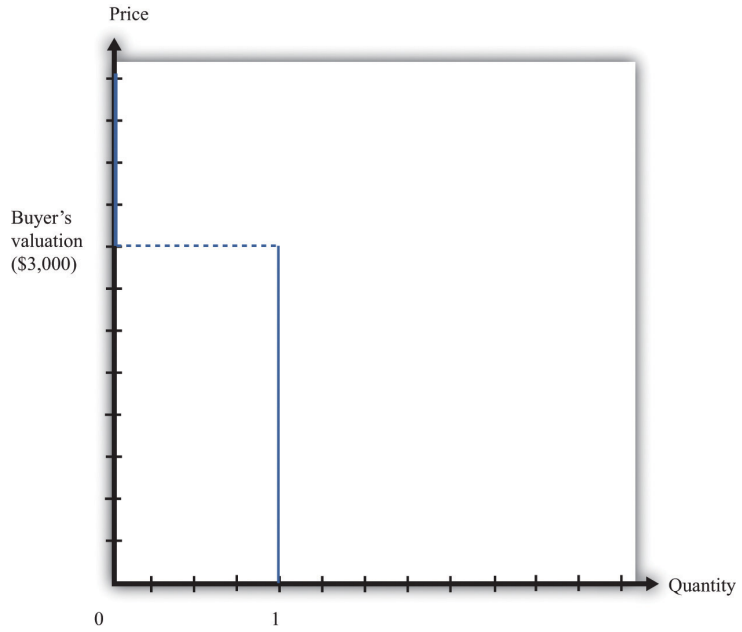


Figure 4.3.7: Unit Demand. The buyer follows this decision rule: "Buy if the price is less than valuation." If the price is greater than $3,000, the buyer will not purchase. The quantity demanded is zero. If the price is less than $3,000, the buyer will purchase one unit. No matter how low the price decreases, the buyer will not want more than a single unit. This is an example of *unit demand.*

To see where such a valuation could come from, look at Figure 4.3.8. Suppose you are thinking about buying a car. The figure shows our standard downloads-and-chocolate-bar diagram, except that there are two budget lines. The outer budget line applies in the case where you do not buy the car. You have a preferred point in terms of downloads and chocolate bars (A). If you buy the car, you have less income to spend on everything else. The effect is to shift your budget line inward, in which case you have a new preferred point (B). The thought experiment here is to decrease your income and shift your budget line inward until you are equally happy with the two bundles. The change in your income—the amount by which the budget line must shift—is your valuation of the car. Your valuation, in other words, is the opportunity cost of the car: if you buy the car, you can only consume bundle B rather than bundle A.

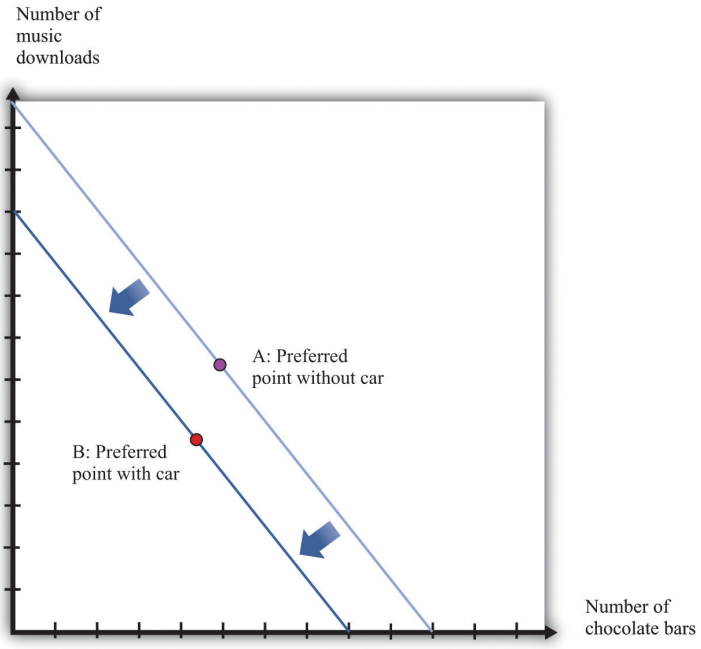


Figure 4.3.8: The Valuation of a Car. If you don't buy a car, your preferred mix of chocolate bars and downloads is at point A. If you buy a car, then you no longer have that income available to spend on chocolate bars and downloads. Your budget line shifts inward, and you consume at your preferred point B. Now imagine that you are equally happy at point A and point B. Then the difference in income is equal to your valuation of the car. Thus the valuation of a car is its opportunity cost in terms of other goods.

Budget Studies

Economists' theories are all well and good. But we do not actually get to see people's preferences or marginal valuations. We observe what people actually do. To see our theory in action, we can look at *household budget studies*. These are surveys where government statisticians interview households and ask them how they spend their income. For example, Table 4.3.3 "Budget Shares in the United States" contains data on US consumer expenditures for the years 2005, 2007, and 2009.

Table 4.3.3: Budget Shares in the United States

Year	2005	2005 (Under 25)	2007	2007 (Under 25)	2009	2009 (Under 25)
Income (before tax)	58,712	27,404	63,091	31,443	62,857	25,695
Spending	46,409	27,770	49,638	29,457	49,067	28,119
Category	**Percentage of Total Spending**					
Food	12.8	14.2	12.4	14.1	13.0	14.9
Alcohol	0.9	1.4	0.9	1.6	0.9	1.2
Housing	32.7	32.2	34.1	32.6	34.4	34.6
Apparel	4.1	5.7	3.8	5.0	3.5	5.0
Transportation	18.0	21.6	17.6	19.4	15.9	19.0
Health care	5.7	2.5	5.7	2.7	6.4	2.4
Entertainment	5.2	5.0	5.4	4.9	5.5	4.4
Personal care products and services	1.2	1.2	1.2	1.1	1.2	1.3
Reading	0.3	0.2	0.2	0.2	0.2	0.2
Education	2.0	4.9	1.9	6.1	2.2	6.8
Tobacco	0.7	1.1	0.7	1.0	0.8	1.2
Personal insurance and pensions	11.2	7.7	10.8	8.3	11.2	7.1
Other	5.3	2.3	5.3	3.0	4.8	1.9

Source: US Department of Labor, "Consumer Expenditures Survey," table 47, http://www.bls.gov/cex/2005/share/age.pdf, http://www.bls.gov/cex/2007/share/age.pdf, and http://www.bls.gov/cex/2009/share/age.pdf, all accessed February 24, 2011.

You can see that, on average, households spend a little more than 45 percent of their income on food and housing. Insurance is also a large category, with about 11 percent of income being spent on it. Chapter 5 discusses why we buy insurance. Interestingly, the budget shares do not change much over the three years despite the differences in income and spending. From this we see that, although individual goods may be inferior or luxury goods, such differences are largely offset when we look at broad categories of goods or services.

The table also contains data for households under age 25.This is based on the age of the reference person in the household, who is the individual who owns or rents the property. We can compare the spending patterns of this group against all households. Not surprisingly, the younger group earns less than the average household. Also, this younger group often spends more than it earns, indicating that younger people are borrowing, on average. The younger group spends more on alcohol, transportation, and education and much less on health care and insurance than the average household. This makes sense given the health status of young individuals as well as their demand for education.

Table 4.3.4 "United Kingdom Budget Study" is a UK budget study for households headed by young people (under the age of 30) in 2009. It shows how these households allocated their expenditures over a week.The British source is as follows: Office for National Statistics, *Family Spending, 2010,* table A.11, accessed January 24, 2011, www.statistics.gov.uk/downloads/theme_social/family-spending-2009/familyspending2010.pdf. We can compare these figures against those for young people in the United States. (We need to be careful in making comparisons because the categories for spending are not exactly the same across surveys. Still, it is useful to explore these differences.) In the United Kingdom, spending on food and housing is much lower for these younger households than in the United States. Health also has a much lower expenditure share.

Table 4.3.4: United Kingdom Budget Study

Category of Expenditure	Spending Share (%)
Food and nonalcoholic drinks	9.2
Alcoholic beverages, tobacco, and narcotics	2.2
Clothing and footwear	4.3
Housing, fuel, and power	22.6
Household goods and services	4.4
Health	0.9
Transport	12.2
Communication	2.7
Recreation and culture	9.0
Education	4.5
Restaurants and hotels	8.0
Miscellaneous goods and services	7.1
Other expenditure items	13.0

Source: Data from Office for National Statistics, *Family Spending, 2010*, table A.11, accessed January 24, 2011, www.statistics.gov.uk/downloads/theme_social/family-spending-2009/familyspending2010.pdf.

Key Takeaways

- The demand curve of an individual shows the quantity of a good or service demanded at different prices, given income and other prices.
- The law of demand—which holds for almost all goods and services—states that the demand curve slopes downward: as the price of a good decreases, the quantity demanded of that good will increase.
- When you are making an optimal choice between two goods, the rate at which you want to trade off the two goods—at the margin—should equal the rate at which the market allows you to trade off the two goods.
- You should buy one more unit of a good whenever your marginal valuation of the good is greater than the price.
- When you are willing to buy at most one unit of a good (unit demand), your valuation and your marginal valuation are identical, so you should purchase the good as long as your valuation of that good is greater than the price.

check your understanding

1. Think about your own preferences. Can you think of a good that—for you—is a substitute for a chocolate bar? Can you think of a good that is a complement?
2. Draw a version of Figure 4.3.1 to show a decrease in income.
3. Create a version of Figure 4.3.2 that shows music downloads as an inferior good. Why can't you draw a version of the figure where both music downloads and chocolate bars are inferior goods?

This page titled 4.3: Individual Demand is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

4.4: Individual Decision Making- How You Spend Your Time 4.4: Individual Decision Making- How You Spend Your TimeEdit Update titlecancel

Learning Objectives

1. What is the time budget constraint of an individual?
2. What is the opportunity cost of spending your time on a particular activity?
3. What is the meaning of real wage?
4. What is the labor supply curve of an individual, and how does it depend on the real wage?

So far we have discussed how you choose to spend your money. There is another decision you make every day: how to spend your time. You have 24 hours each day in which to do all the different things you want to do: work, sleep, eat, study, watch television, surf the Internet, go to the movies, and so on. Time, like money, is scarce. Given that you have only 24 hours to allocate in every day, how do you decide which activities to spend your time on? This problem is very similar to the allocation of your budget, with one key difference: you cannot save or borrow time in the way that you can save or borrow money. There are exactly 24 hours in each day—no more, no less.

Choosing among Different Uses of Your Time

We begin with the most fundamental time allocation problem for all students: choosing between studying and sleeping. As before, we keep things simple by thinking about only two possible uses of your time. You are given 24 hours in the day to study and sleep. How should you allocate your time?

As with the allocation of your income, there are two aspects of this problem. First, there is a budget constraint—only now it is your time that is scarce, not money. Second, you have your own preferences about sleep and study time. Your ability to meet your desires is constrained by the scarcity of your time: you must trade off one activity for another.

The **time budget constraint** is the restriction that there are only 24 hours in the day. It is shown in Figure 4.4.1 and is the counterpart to the budget line in our earlier discussion. Any point in this figure represents a combination of sleep and study time. The sum of sleep and study time must equal 24 hours (remember we are supposing that these are the only ways you spend your time). Thus your allocation of your time must lie somewhere on this line.

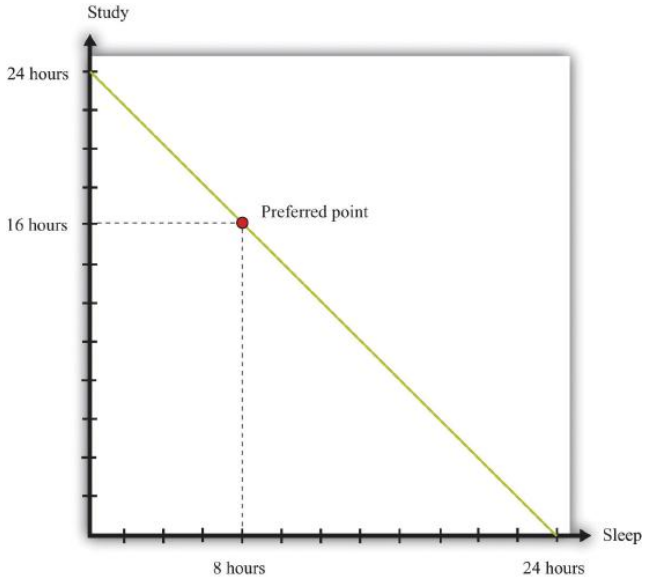


Figure 4.4.1: The Time Budget Constraint. The time allocation line shows your options for dividing your time between study and sleep.

Figure 4.4.1 also shows one possible choice that you might make: allocating 8 hours to sleep and 16 hours to studying. The choice of this point reflects your desires for sleep and study. As with the spending decision, we pass no judgment, as economists, on the actual decision you make. We suppose you typically make the choice that makes you the happiest.

At your preferred point, your choice to sleep for 8 hours means that your study time *must* equal 16 hours; equivalently, your choice to study for 16 hours means that you must sleep for 8 hours. Any increase in one activity must be met by a reduction in time for the other. The opportunity cost of each hour of sleep is an hour of study time, and the opportunity cost of each hour of study time is an hour of sleep. If you choose this point, you reveal that you are willing to "pay" (that is, give up) 8 hours of study time to obtain 8 hours of sleep, and you are willing to pay 16 hours of sleep to obtain 16 hours of study time.

As with consumption choices, it is often enough to look at small changes to evaluate whether or not you are making a good decision. The opportunity cost of a little more sleep is a little less study time. If you are making a good decision about the allocation of your time, then the extra sleep is not worth the extra study time. Suppose you are contemplating a particular point on the time budget line, and you want to know if it is a good choice. If a very small movement away from your chosen point will not make you happier, then—in most circumstances—neither will a big movement. By a very small movement, we mean sleeping a little less and studying a little more *or* studying a little less and sleeping a little more. If you are making a good decision, then your willingness to substitute sleep for study time is exactly the same as that allowed by the time constraint.

Individual Labor Supply

Sleeping and studying are uses of your time that are directly for your own benefit. Most people—perhaps including yourself—also spend time working for money. So let us now look at the choice between spending time working and enjoying leisure. Our goal is to determine how much labor you will choose to supply to the market—which is equivalently a choice about how much leisure time to enjoy, because choosing the number of working hours is the same as choosing the number of leisure hours. Your choice between the two is based on the trade-off between enjoying leisure and working to earn money that allows the purchase of goods.

Once again—to make it easy to draw diagrams—we suppose that these are the only uses of your time. Part (a) of Figure 4.4.2 presents the allocation of time between work and leisure. As with the sleep-study choice, there is a time budget constraint, and you have preferences between these two ways to allocate time. Your best choice satisfies the same property as before: you allocate time such that no other division of your time makes you happier.

What makes this different from the sleep-study choice is the valuation of your time. We can think of sleep as a good thing in that you generally prefer more to less. Likewise, we can think of study as a good thing in that—even if you don't always enjoy it—you perceive a gain to spending time studying. So Figure 4.4.1 is like our earlier diagrams with downloads and chocolate bars: it has a good thing on each axis. Now, people presumably prefer more leisure to less: leisure is a "good," like chocolate bars, blue jeans, or cans of soda. But we have drawn part of (a) Figure 4.4.2 as if work is also a good thing. Most people, however, see work time as a "bad" rather than as a "good." Even people who like their work would almost always prefer to work a little less and have a little more leisure time.

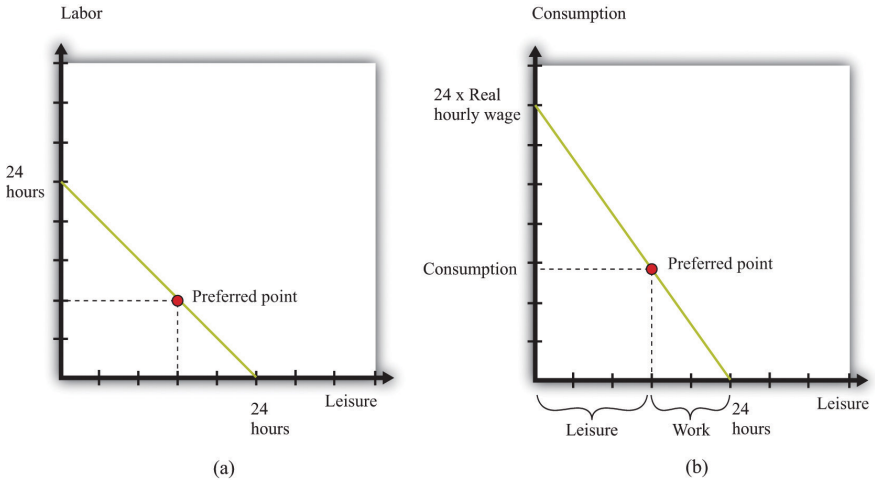


Figure 4.4.2: Choosing between Work and Leisure. (a) The time budget line shows your options for dividing your time between labor and leisure. However, we generally think of labor as a "bad" rather than as a "good." (b) Now the choice is between consumption and leisure. For each hour of your time, you earn the nominal hourly wage. If you divide this by the price level, you get the real wage. The real wage tells you how many goods and services you can enjoy for one hour of work.

The gain from working, of course, is that you earn income, allowing you to purchase goods and services. Each extra hour of your work allows you to buy more goods and services. Conversely, if you want more leisure time, you must give up some goods and services. Thus the choice between labor and leisure is linked to the choice about how many chocolate bars and other goods you

buy. The income we take as given in describing your budget set typically comes from your decision to supply labor time. (Of course, you may have other sources of income as well, such as loans or grants.)

Part (b) of Figure 4.4.2 takes the labor-leisure choice and converts it into a choice between leisure and consumption. Here, consumption refers to all the goods and services you consume. We lump together all the products you consume, just as we lump together all your different forms of leisure (sleep, study, watching television, and so on). As before, time is measured on the horizontal axis: there are 24 hours to the day, which must be split between leisure time and labor time. On the vertical axis, we measure consumption.

To get the budget constraint for this picture, we begin with the time budget constraint:

$$leisure\ hours\ +\ labor\ hours\ =\ 24.$$

The value of an hour of time in dollars is given by the wages at which you can sell your time. Multiplying the time budget constraint by the wage gives us a budget constraint in dollars:

$$(leisure\ hours\ \times\ wage)\ +\ wage\ income\ =\ 24\ \times\ wage.$$

Wage income is equal to the number of hours worked times the hourly wage. Because wage income is used to buy goods, we can replace it by total spending on consumption, which is the price level times the quantity of consumption goods purchased:

$$(leisure\ hours\ \times\ wage)\ +\ (price\ level\ \times\ consumption)\ =\ 24\ \times\ wage.$$

This is the budget constraint faced by an individual choosing between leisure and consumption. Think of it as follows: The individual first sells all her labor at the going wage, yielding the income on the right-hand side. With this income, she then "buys" back leisure and also buys consumption goods. The price of an hour of leisure represents the wage rate, and the price of a unit of consumption goods represents the price level.

Toolkit: Section 31.3

The real wage is the *relative price* of labor in terms of consumption goods:

Dividing the time budget constraint by the price level, we get the budget in line in part (b) of Figure 4.4.2.

$$leisure\ hours\ \times\ real\ wage\ +\ consumption\ =\ 24\ \times\ real\ wage.$$

As you move along the budget line, you trade hours of leisure for consumption goods. The slope of the budget line is the negative of the **real wage**. If you give up an hour of leisure, you obtain extra consumption equal to the real wage. Put differently, the opportunity cost of an hour of leisure is the amount of consumption you give up by not working. Once we have worked out how much leisure you consume, we have equivalently worked out how much labor you supply:

$$labor\ hours\ =\ 24\ -\ leisure\ hours.$$

Part (a) of Figure 4.4.3 shows what happens when the real wage changes. When the real wage increases, the vertical intercept of the budget line is higher because the vertical intercept tells us how much consumption an individual could obtain if she worked for all 24 hours in the day. The horizontal intercept does not change as the real wage changes: if an individual does not work, then the level of consumption is zero regardless of wages. It follows that the budget line is steeper as the real wage increases. If an individual gives up an hour of leisure time, he or she gets more additional consumption when the real wage is higher. The opportunity cost of leisure in terms of forgone consumption is higher.

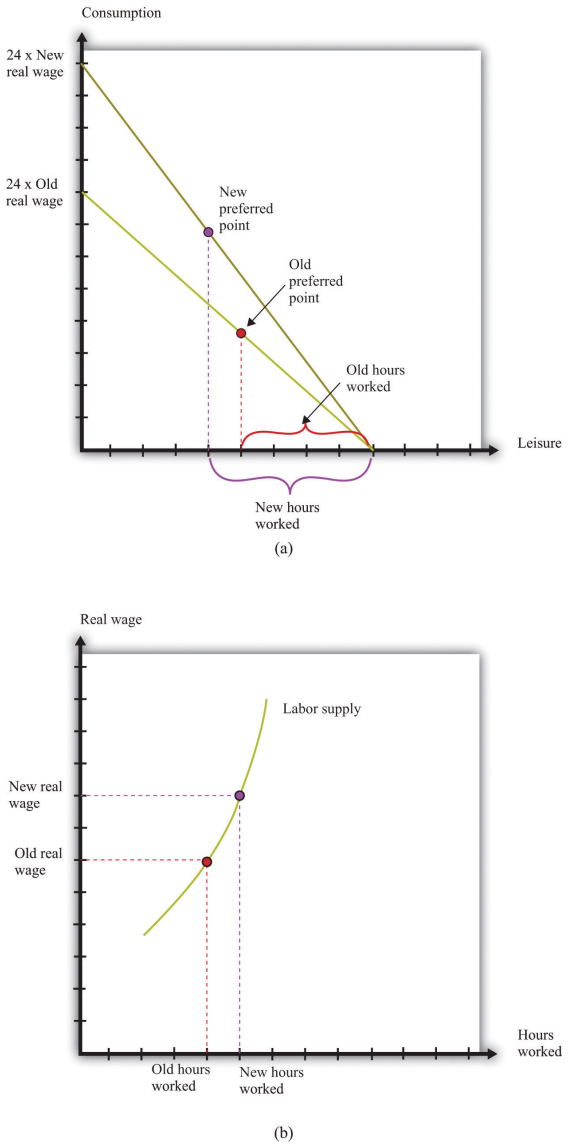


Figure 4.4.3: The Effect of a Real Wage Increase on the Quantity of Labor Supplied. (a) An increase in the real wage causes the budget line to rotate. Income and substitution effects are both at work: the income effect encourages more leisure (less work), while the substitution effect encourages more work. The substitution effect generally dominates, so higher real wages lead to more work. This means that the labor supply curve slopes upward, as shown in (b).

Part (b) of Figure 4.4.3 shows the **individual labor supply curve** that emerges from the labor-leisure choice.

Toolkit: Section 31.3

The individual labor supply curve shows the number of hours that an individual chooses to work at each value of the real wage.

In fact, there are conflicting incentives at work here. As the real wage increases, the opportunity cost of leisure is higher, so you are tempted to work more. But at a higher real wage, you can enjoy the same amount of consumption with fewer hours of work, so this tempts you to work less. This is another example of substitution and income effects. The substitution effect says that when something gets more expensive, we buy less of it. When the real wage increases, leisure is more expensive. The income effect says that as the real wage increases, you can buy more of the things you like, including leisure. We know from our study of demand that, for normal goods, the income and substitution effects act in the same direction. In the case of supply, however, income and substitution effects point in *different* directions. Consistent with this, most economic studies find that, though the labor supply curve slopes upward, hours worked are not very responsive to changes in the real wage.

Some jobs do not give you any control over the number of hours that you must work. Labor supply then becomes a “unit supply” decision, analogous to the unit demand decisions we considered previously. Should you take a job at all and, if so, what job? To the extent that you can choose among different jobs that offer different hours of work, your decision about whether or not to work will still reflect a trade-off between leisure and consumption.

Individuals and Households

We have discussed almost everything in this chapter in terms of an individual’s decision making. However, economists often think in terms of households rather than individuals. In part this is because—as we saw with the budget studies—much more economic data are collected for households than for individuals. Also, many of the decisions we have discussed are really made by a household as a whole, rather than by the individual members of that household.

For example, many households have two working adults. Their decisions about how much to work will usually be made jointly on the basis of the real wages they both face. To see some implications of this, consider a two-person household in which both are working. Now suppose that the real wage increases for one person in the household. One person will probably respond by increasing the number of hours worked. However, the other person may choose to work less. Imagine, for example, that there are household chores that either could do. By working less, one person can do more of these chores and thus compensate the other person for the extra hours worked.This is an application of an important economic idea called *comparative advantage*, which we discuss in more detail in Chapter 6. Most of the time, though, we do not need to worry about the distinction between the individual and the household, and we often use the terms interchangeably.

Time Studies

Table 4.4.1 "Allocation of Hours in a Day" shows the allocation of time to certain activities for individuals in three countries: the United States, the United Kingdom, and Mexico. It shows the time allocated on average per day for each of four activities: work, study, personal care, and leisure.

Table 4.4.1: Allocation of Hours in a Day

Country	Age	Work	Study	Personal Care	Leisure
United States	15–24	2.65	2.2	9.95	5.46
United Kingdom	16–24	3.00	1.39	9.96	5.13
Mexico	20–29	4.49	0.72	10.37	3.1

For the United States and the United Kingdom, the average number of hours worked is between 2.65 and 3.00. This is an average: some people in this age group may work a full-time job, while others may be students who are not working for pay at all. The sample from Mexico differs from the US and UK samples. First, the group is slightly older. Second, Mexico is notably poorer than the United States and the United Kingdom. For these two reasons, individuals sampled in Mexico are more likely to be working and less likely to be studying and enjoying leisure—which is indeed what we see.

Combining Your Time and Spending Choices

So far we have looked at the allocation of your income separately from the allocation of your time. Yet these choices are linked. The allocation of your time influences the income you have to spend on goods and services. So a change in the wages you are paid will affect how you allocate your time *and* the goods and services you choose to buy.

In a similar fashion, the prices of goods and services you purchase will have an influence on your allocation of time. For example, if the price of a computer you want to buy decreases, you may respond by working a little more to earn extra income to purchase the computer. The reduction in the price of the computer raises your real wage, so you respond by working more.

Effects of Real Wages on Household Demand

If the real wage changes, there are changes in both consumption decisions and work choices. Figure 4.4.3 shows that an increase in the real wage means you can obtain more consumption for a given amount of work time. Further, as in Figure 4.4.1, the budget set

expands as income increases. Because an increase in the real wage will lead to an increase in hours worked (see Figure 4.4.3), labor income will increase. So we can interpret the shift in the budget line in Figure 4.4.1 as coming from this increase in labor income.

When income increases, you will generally consume more of all goods and services. An increase in the real wage leads to an outward shift in your demand curves for chocolate bars, downloads, and all other normal goods. Combining these figures, we can make the following predictions about the effects of an increase in the real wage:

- You will work more hours.
- You will have more income.
- You will consume more goods and services.
- You will be happier.

The last item in the list draws on Figure 4.4.1 but is less direct than the other implications. As the real wage—and thus your income—increases, the set of bundles you can afford is larger. Moreover, every bundle you could afford when you had less income is still affordable now that you have more income. Thus we conclude that you will be happier. After all, you can always purchase the bundle you bought with lower income and still have some extra income to spend.

Effects of Prices on Time Allocation

Suppose the price of a chocolate bar increases. We saw from Figure 4.4.3 that when this price increases, the budget set shrinks. We also saw from Figure 4.4.4 that the demand for chocolate bars decreases when the price of a chocolate bar increases. But there are also implications for labor supply. Remember that the real wage is the nominal wage dividing by a price index representing a household's cost of purchasing a bundle of goods and services. So when the price of a chocolate bar increases, the cost of purchasing the bundle will increase and the real wage will decrease. From Figure 4.4.3, labor supply will decrease as the real wage decreases. This is a *movement along* the labor supply curve.

Key Takeaways

- The time budget constraint states that the sum of the time spent on all activities each day must equal 24 hours.
- The opportunity cost of time spent on one activity is the time taken away from another.
- Decisions about how much to work depend on how much more you can purchase if you work a little more: that is, they depend on the real wage.
- The individual labor supply curve shows how much an individual will choose to work given the real wage.
- As the real wage increases, an individual will supply more labor if the substitution effect dominates the income effect.

Check your understanding

1. If you must sleep a minimum of five hours each day, how would you modify Figure 4.4.1 to indicate this necessity of life?
2. If both the nominal wage and the price level double, what will happen to your allocation of time and consumption?

This page titled 4.4: Individual Decision Making- How You Spend Your Time is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

4.5: End-of-Chapter Material 4.5: End-of-Chapter MaterialEdit Update titlecancel

In Conclusion

We all make economic decisions every day, often without giving them very much thought. In this chapter, we highlighted two fundamental decisions: allocating income and allocating time. You (and everyone else) generally make choices over income and time allocations in a manner that makes you happy. Thus we predict that you will not throw income away. Further, whatever combinations of goods and services and time allocations you choose, economic theory presumes that these are the best ones available to you. Remember also that even though it is usually easier to focus on one decision at a time, your spending and time allocation decisions are interconnected. Changes in the prices of goods and services affect how you spend your time, and changes in the real wage affect your consumption choices. Chapter 8 has more to say on the connections among different markets in the economy.

Economics is often defined as the study of how we allocate scarce resources that have alternative uses. In this chapter, we saw this idea at work at the individual level. Your income is a scarce resource; you don't have enough income to buy everything that you would like. Your income has alternative uses because there are lots of things you might want to buy.

Perhaps the most fundamental idea of this chapter is that of opportunity cost. Given that you have limited income to allocate across goods and services, the opportunity cost of consuming one good or service is the amount of another good or service you give up. Given that you have limited time to allocate across activities, the opportunity cost of spending time on one activity is the value of the time you could have spent on another. The budget and time budget constraints are graphical representations of this central economic principle. And the interaction between these budget constraints and people's wants and desires is at the heart of the economic analysis of decision making.

Key Link

- *New York Times* Time Allocation Study: http://www.nytimes.com/interactive/2009/07/31/business/20080801-metrics-graphic.html

exercises

1. Which statements are prescriptive? Which statements are descriptive?
   - The government should take care of the poor.
   - If the real wage increases, households will be willing to supply more labor.
   - When people's incomes decrease, they consume more cheap cuts of meat.
   - We ought to consume fewer resources to protect the planet.
   - Young people should purchase medical insurance because it is cheaper for young people than for old people.
2. Draw a budget line assuming disposable income equals $100, the price of a music download is $1, and the price of a chocolate bar is $5. On the same graph, draw another budget line assuming the same level of income but with the price of a music download equal to $2 and the price of a chocolate bar equal to $1. Explain how the budget sets differ. If you liked downloads but hated chocolate bars, which budget set would you prefer?
3. Consider a bundle consisting of 4 chocolate bars and 30 downloads (call it bundle A). Using the assumption that "more is better," what bundles can we say are definitely preferred to bundle A? What bundles are definitely worse than bundle A? Show your answers on our usual kind of diagram (that is, a budget set diagram with chocolate bars and downloads on the axes).
4. If we observe that a household buys bundle A but not bundle B and we know that bundle B has more of every good than does bundle A, what can we say about the household's preferences for bundle A and bundle B? What can we say about the household's budget set?
5. (Advanced) Look at Table 4.5.1 "Preferences over Downloads and Chocolate Bars". The top part of the table lists four different bundles of downloads and chocolate bars. The bottom part of the table shows which bundle is preferred when we compare any two bundles. Look at bundle 1. The top part of the table tells us it contains 0 downloads and 20 chocolate bars. The first row of the bottom part of the table shows how this bundle compares to the other bundles. So this individual prefers bundle 1 to bundle 2 but also prefers both bundle 3 and bundle 4 to bundle 1. Do these preferences satisfy "more is better"? Are they consistent or can you find some contradictions?

Bundle	Downloads Consumption	Chocolate Bar Consumption
1	0	20
2	100	0
3	50	10
4	110	30

Bundle	Which Bundle Is Preferred When Comparing Bundles?			
	Bundle 1	Bundle 2	Bundle 3	Bundle 4
1	—	1	3	4
2	1	—	2	4
3	3	2	—	3
4	4	4	3	—

TABLE 4.5.1: PREFERENCES OVER DOWNLOADS AND CHOCOLATE BARS

6. Suppose income increases by 10 percent, but the price of a chocolate bar and the price of downloads both increase by 5 percent. Will the budget line shift inward or outward? Will the slope of the budget line change?
7. We explained the household demand curve and the law of demand by focusing on how a change in the price of a chocolate bar influences the quantity of chocolate bars demanded. Redo this discussion and the figures to illustrate how a change in the price of downloads will affect the demand for downloads and the demand for chocolate bars.
8. In our example, we noted that it was not possible for both chocolate bars and music downloads to be inferior goods. Suppose there were three goods: chocolate bars, music downloads, and tuna sushi. Is it possible now that chocolate bars and music downloads are both inferior goods? Could all three be inferior goods?
9. (Advanced) Suppose the government imposes a tax on chocolate bars. Draw a diagram that shows what happens to the budget set. If chocolate bars and downloads are both normal goods, can you say whether the consumption of chocolate bars will increase or decrease? What about the consumption of downloads?
10. Explain why a price increase in movie tickets causes the demand curve for chocolate bars to shift.
11. Suppose you are thinking of buying chocolate bars. Your marginal valuation of the seventh chocolate bar is $3. The price of a chocolate bar is $4. Should you buy more or fewer than seven bars?
12. Explain how the law of demand works in the case of a unit demand curve.
13. Can preferences include altruism or a regard for fairness and still exhibit rationality?
14. Using the data in Table 4.3.4 "United Kingdom Budget Study", create a pie graph of expenditure shares. How might you explain the differences in spending between younger households in the United States and the United Kingdom? How might you explain the differences in spending in 2005 between younger households and average households?
15. (Advanced) In discussing labor supply, we did not allow an individual to decide not to work. Yet we observe many individuals who could work but choose not to. How would you have to amend the discussion to include the choice of working or not working?
16. If you face a big exam this week, how might this influence your time allocation choice in Figure 4.5.1?
17. (Advanced) If there is a reduction in the price of a chocolate bar, what does our theory predict will happen to labor supply?
18. (Advanced) Suppose the government imposes a tax on labor. What will that tax do to the labor supply of a household and its demands for downloads and chocolate bars?
19. (Advanced) If one member of a two-person household gets a raise, what will that do to the hours worked by that person and to the other household member? Explain this using income and substitution effects. Could this raise cause the other household member not to work at all?

Economics Detective

1. Search the Internet to find the level of spending by Japanese households on food in a recent year. Convert this figure to dollars.
2. The data in Table 4.3.3 "Budget Shares in the United States" come from a survey. Who was surveyed? How frequently?

3. Go to the web page of the Office for National Statistics in the United Kingdom (www.statistics.gov.uk/CCI/Nscl.asp?ID=5407&Pos=1&ColRank=1&Rank=16UK) and create a version of Table 4.3.4 "United Kingdom Budget Study" for different income groups. What differences do you see in spending patterns across income groups? How would you explain the differences in spending patterns as income changes?
4. Go to http://www.bls.gov/tus. Pick two years. Prepare a table to illustrate how the allocation of time has changed for one of two age groups over these two years. How might you explain these changes?

Spreadsheet Exercise

1. (Advanced) Create a spreadsheet to reproduce the graph of a budget constraint with two goods (chocolate bars and downloads) in Figure 4.5.4. In column A, put the quantity of chocolate bars (from 0 to 20). In column B, put the price of a chocolate bar (that is, each cell should contain a 5).This is the simplest but not the most elegant way to create this spreadsheet. If you are an experienced user of spreadsheets, you may know tricks that will allow you to create the spreadsheet in a more compact way. In column C, put the price of downloads. In column D, put income. Then write an equation to enter in each cell of column E, based on the budget line. This equation should calculate the quantity of downloads in terms of the prices, income, and the quantity of chocolate bars. Make sure that you allow only nonnegative quantities of the goods. Use this to graph the budget line. Now try changing the prices and the level of income and make sure you can explain how the budget line shifts as income and prices change.

This page titled 4.5: End-of-Chapter Material is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

5: Life Decisions

5.1: Life Choices
5.2: Consumption and Saving
5.3: Using Discounted Present Values
5.4: Avoiding Risk
5.5: Embracing Risk
5.6: End-of-Chapter Material

This page titled 5: Life Decisions is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

5.1: Life Choices 5.1: Life ChoicesEdit Update titlecancel

Some economic decisions, like how to spend your money and your time, are everyday decisions.See Chapter 4 for more discussion. There are also bigger and more difficult economic decisions that you confront only occasionally. In the months and years after graduation, you will face major life choices, such as the following:

- Upon graduation, which occupation should you choose?
- Should you go to graduate school?
- Should you purchase a new car?
- Should you purchase a house?
- How much of your income should you save?
- Should you purchase health insurance?
- Should you purchase insurance for your home or apartment?
- What should you do with the money you save?

These economic choices are more complicated than choosing how many chocolate bars to buy or how much time you should spend watching television today.

Two things make these decisions hard. First, there is the element of time—not the 24 hours in a day, but the fact that you must make decisions whose consequences will unfold over time. In choosing an occupation, deciding on graduate school, or picking a portfolio of financial assets, you must look ahead. Second, there is the element of uncertainty. Will you be healthy? Will you live to an old age? Will you succeed as a rock musician? The future is unknown, yet we cannot ignore it. The future is coming whether we like it or not.

We cannot tell you whether you should buy a new car or if you will be a rock star. But we can give you some tools that will help you when you are making decisions that involve time and uncertainty. In this chapter we tackle the following questions:

- How do we make decisions over time?
- How do we make plans for an uncertain future?

Road Map

The chapter is organized around the two themes of time and uncertainty. We begin with a brief review of the choice between two goods at a given time.This decision is analyzed at length in Chapter 4. Then we look at choices over time. Economists typically assume that individuals are capable of choosing consistently among the bundles of goods and services they might wish to consume. The ability to make such a choice is perhaps not too onerous in the case of simple choices at a given time (such as whether to go to a movie or go to dinner). It is more difficult when we consider choices over a broad range of goods from now into the future.

- There are goods and services that will be available to you in the future that you cannot imagine today. When people chose among different types of handheld calculators 30 years ago, they could not imagine that today they would be choosing among different types of tablet computers. Many products that we now consume simply did not exist in any form until comparatively recently, and when we make choices now, we do so in ignorance of future consumption possibilities.
- Your tastes may change. When you are 20 years old, it is difficult to predict what goods and services you will want to buy when you are 30, 40, or 50 years old. Your future self might regret past decisions.

We tackle time and uncertainty separately. To begin with, we will suppose that the future is known with certainty. This allows us to focus on including time in our analysis of economic decision making. We begin with a discussion of the choice between consumption and saving and explain how this decision is affected by changes in interest rates. We then look at problems such as how to choose an occupation. A major part of this analysis is an explanation of how to compare income that we receive in different years.

We then turn to uncertainty. We explain the idea of risk and then discuss the kinds of risks you cannot avoid in life. We explain how insurance is a way to cope with these risks. We also discuss uncertainties that we *create* in our lives—through occupational choice, portfolio choice, and gambling.

This page titled 5.1: Life Choices is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

5.2: Consumption and Saving 5.2: Consumption and SavingEdit Update titlecancel

Learning Objectives

1. What is your lifetime budget constraint?
2. What factors influence your choice between consumption today and saving for the future?
3. What is the difference between real and nominal interest rates?
4. What are the effects of a change in the interest rate on consumption and saving?

Your choice at any given time between two goods—say, chocolate bars and music downloads—reflects the tension between your desires for chocolate bars and downloads and your income, as summarized by your **budget line**. The budget line shows us the bundles of goods and services that you can afford, given prices and your income, under the presumption that you do not throw any money away. For an individual choosing between two goods only (chocolate bars and music downloads), the budget line states that total spending is equal to spending on chocolate bars plus spending on downloads:

$$(number\ of\ chocolate\ bars\ \times\ price\ of\ a\ chocolate\ bar)\ +\ (number\ of\ downloads\ \times\ price\ of\ download)\ =\ disposable\ income.$$

Figure 5.2.1 displays the budget line for the choice between downloads and chocolate bars.

Toolkit: Section 31.1

You can review the derivation and meaning of the budget line in the toolkit.

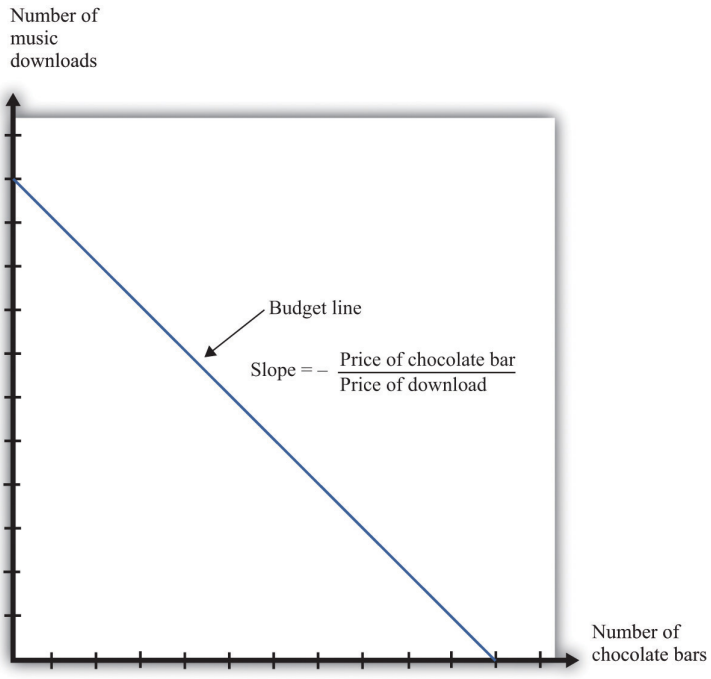


Figure 5.2.1: The Budget Line with Two Goods

This diagram shows the budget line for an individual choosing between chocolate bars and music downloads. The slope of the budget line reflects the rate at which an individual can trade off the two goods.

The slope of the budget line is

As you move along the budget line, you are giving up downloads to get chocolate bars. If you give up one chocolate bar, you get an amount of money equal to the price of a chocolate bar. You can take this money and use it to buy music downloads. You have to divide this amount of money by the price of a download to determine how many downloads you can buy. The slope of the budget line reflects the opportunity cost of chocolate bars in terms of music downloads.

The Budget Line with Two Periods

Now think about what the budget line looks like when we are choosing between now and the future. Just as we find it easier to think about the choice between two goods rather than among 2,000 goods, so too is it easier to think about the choice over only two periods of time. We call these two periods "this year" and "next year." There is nothing special about the two-period example beyond the fact that it makes it easy to draw diagrams and see what is going on. The principles that we uncover for this case also apply to decisions made with more than two time periods in mind.

We also do not worry about all the different goods and services that are available, preferring instead to talk in general about consumption this year and consumption next year. We *aggregate* together all the different products that we consume. Thus "consumption" means the bundle of goods and services people consume. This consumption also has an associated price, which we call the **price level**. Think of this as the average price of goods and services in the economy. If you find it easier, imagine there is a single good, like chocolate bars, that you consume, and think of the price level as simply being the price of a chocolate bar.

Saving and the Nominal Interest Rate

If you choose not to spend all your income on consumption this year, you are saving. When you save, you can put your money into a financial institution and earn interest on it. Suppose you have $100 this year that you save by putting it in a bank. You are then *lending* to the bank—saving and lending are really the same thing. The bank acts as an intermediary, taking your $100 and giving it to someone else who borrows from the bank.

The bank offers you interest on this loan—for example, it may pay a **nominal interest rate** of 5 percent per year. After a year, your bank account will contain your original $100, plus an extra 5 percent. Because 5 percent of $100 is $5, you earn $5 worth of interest. We talk about interest rates in percentage terms, but you should remember that a percentage is simply a number. For example, 5 percent is 0.05, and 20 percent is 0.2. The nominal interest rate is the interest rate at which individuals and firms in the economy can save or borrow.There are actually many different interest rates in an economy. Chapter 10 looks at some of these. Here, we simplify the process by supposing there is only one interest rate. It is called a *nominal* interest rate because it is measured in monetary terms. Most interest rates are quoted on an annual basis, meaning that they specify the amount earned per year.

Of course, if you put a $100 in the bank for one year, then next year you will still have the original $100 as well as the interest you earned. At an interest rate of 5 percent, $100 this year is worth $105 next year. To calculate the *total* amount of money that you can earn, we simply add one to the nominal interest rate, giving us the **nominal interest factor**:

$$nominal\ interest\ factor\ =\ 1\ +\ nominal\ interest\ rate.$$

The nominal interest factor is used to convert dollars today into dollars next year.

Toolkit: Section 31.6

The nominal interest rate is the rate at which individuals and firms in the economy can save or borrow. The nominal interest factor is 1 + the nominal interest rate.

In this chapter, we typically use the nominal interest factor rather than the nominal interest rate because it makes the equations easier to understand. Just keep in mind that it is easy to move back and forth between the interest rate and the interest factor by adding or subtracting one.

If you have $100 today, then tomorrow it will be worth $100 × the nominal interest factor. In general,

$$z\ this\ year\ will\ be\ worth\ z\ timesthe\ nominal\ interest\ factor\ next\ year.$$

Armed with this idea of the nominal interest factor, we can graph the budget line for a two-period consumption-saving problem. Figure 5.2.2 shows consumption this year on the horizontal axis and consumption next year on the vertical axis. To discover what the budget line looks like, we first determine its slope and then its position.

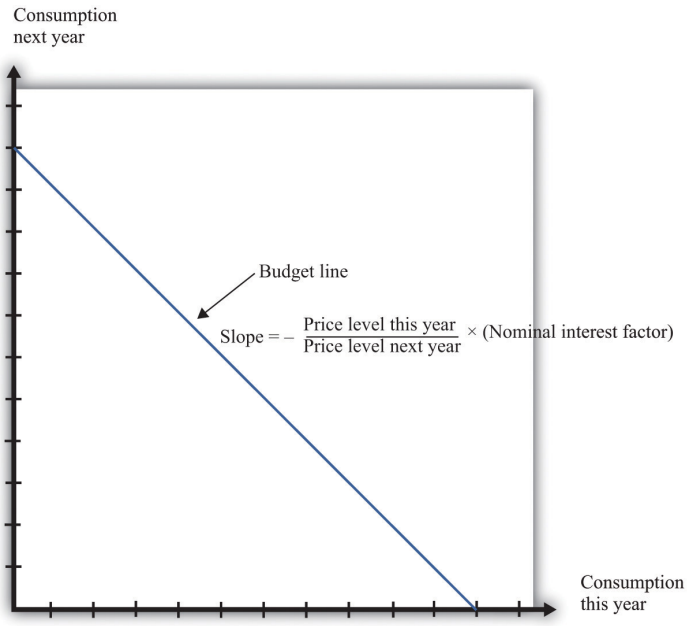


Figure \(\PageIndex{2}\): *The Budget Line with Two Periods*. This diagram shows the budget line for an individual choosing consumption over time. The slope of the budget line depends on the price level this year, the price level next year, and the nominal interest factor (= 1 + nominal interest rate). Suppose the price level this year is $9, the price level next year is $10, and the nominal interest rate is 20 percent. Then the slope of the budget line is −(9/10) × 1.2 = −1.08. This means that if you give up 1 unit of consumption this year, you can get 1.08 units next year.

The Slope of the Budget Line

The slope of the budget line tells you how much extra consumption you will get next year if you give up a unit of consumption this year. So to determine the slope of the budget line, we use the following thought experiment.

1. If you give up one unit of consumption this year, you get an amount of dollars equal to the price of consumption this year (the price level).
2. You can then save those dollars, so next year you will have a number of dollars equal to the price level this year × the nominal interest factor.
3. Next year, you can take these funds and purchase

   units of consumption.

So the slope of the budget line is as follows:

We show the budget line in Figure 5.2.2.

The budget line has a negative slope because—as with the choice between chocolate bars and downloads—you must give one thing up to get another. If you want to consume more in the future, you must be willing to consume less right now. If you want to consume more now, you will have to sacrifice consumption in the future.

The slope of the budget line depends not only on the nominal interest factor but also on prices this year and next year. Suppose the price of a unit of consumption this year is $100 and next year it is $110. Economists call the percentage increase in the price level the **inflation rate**; it is calculated as follows:

Put differently, it is the rate of growth of the price level. In our example, the inflation rate is 10 percent.

Now suppose the nominal interest rate is also 10 percent, which means that the nominal interest factor is 1.1. Then

In this case, the price level increased by 10 percent, from 100 to 110. But the nominal interest rate also increased by 10 percent, which offset the increase in prices. We see that the slope of the budget line depends on both the nominal interest factor and the rate of inflation. In fact, it depends on the **real interest factor**:

$$slope\ of\ budget\ line\ =\ -(1\ +\ real\ interest\ rate)\ =\ -real\ interest\ factor.$$

Toolkit: Section 31.6

The real interest rate is the rate of interest adjusted for inflation. It tells you how much you will get next year, *in terms of goods and services*, if you give up a unit of goods and services this year. The real interest factor allows you to convert units of goods and services this year into units of goods and services next year. The real interest factor is 1 + the real interest rate.

As you move along the budget line in Figure 5.2.2, you are giving up chocolate (consumption) this year for chocolate next year. So the slope of the budget line must be a number, not a dollar amount. Because this year's price and next year's price are both denoted in dollars, their ratio is a number. Likewise, the interest rate is a number, so the slope of the budget line is indeed a number.

An example may help you understand the difference between the nominal interest rate and the **real interest rate**. Suppose you go to your bank and get a one-year, $20,000 loan to buy a car, with a nominal interest rate of 5 percent. Your contract with the bank thus stipulates that you must pay the bank $21,000 at the end of the year. If the inflation rate is zero, then the cost of borrowing measured in terms of real goods and services is $1,000, which is 5 percent of the amount that you borrowed. But if the inflation rate is 5 percent, then the $21,000 you pay to the bank at the end of the loan buys the same amount of goods and services that the $20,000 the bank lent to you. In this case, you are effectively able to borrow for free.

Good decisions about borrowing and lending are based on real interest rates rather than nominal interest rates. Your cost of borrowing to buy the car is not the monetary payments you make on the loan but rather the value of the goods and services you could have purchased with that money. So we need a way to convert from the commonly observed nominal interest rate to a measure of real interest rates. We do this by using a formula for the real interest rate that was discovered by a famous economist named Irving Fisher.

Toolkit: Section 31.8

The Fisher equation is a formula for converting from nominal interest rates to real interest rates, which is as follows:

$$real\ interest\ rate\ \approx\ nominal\ interest\ rate - inflation\ rate.$$

Equivalently,

$$real\ interest\ factor\ \approx\ nominal\ interest\ factor - inflation\ rate.$$

For example, suppose the nominal interest rate is 5 percent. If the rate of inflation is zero, then the real interest rate is 5 percent. But if the rate of inflation is 4 percent, the real interest rate is only 1 percent. The **Fisher equation** is a tool that tells us how to convert nominal interest rates—the interest rates you see in the newspapers and on television—into real interest rates, which are key for decision making.The precise formula is as follows:This equation is, to a very good approximation, the same as the one in the text.

The Position of the Budget Line

Whereas the slope of your budget line depends on the real interest rate, the *position* of your budget line depends on how much income you have. When you have more income, the budget line is further away from the origin. One way to determine the position of the budget line is by looking at its intercepts. The horizontal intercept is the amount you can consume this year if you spend all of this year's income *and* borrow against your entire future income. The vertical intercept is the amount you can consume next year if you choose to consume nothing this year *and* save all of your current income.

It is easier and more instructive, however, to look at a different point on the budget line. Remember that the budget line is the bundles of consumption you can just afford. One bundle you can certainly afford is the bundle where you spend all of this year's income on consumption this year and all of next year's income on consumption next year. In other words, one available option is that you neither save nor borrow. In this case,

and

On the right-hand side of these equations, we divided dollar income by the price level to give us real income (that is, income measured in terms of purchasing power). We must do this to find out how much you can consume in terms of goods and services.

For example, suppose your nominal income this year is $23,000, and your nominal income next year is $24,200. Suppose the price level this year is $10 and the price level next year is $11. This means that

and

So one possible consumption choice, as shown in Figure 5.2.3, is 2,300 units of consumption this year and 2,200 units of consumption next year. In this case, you are neither borrowing nor saving. Of course, you might choose some different point on the budget line. Figure 5.2.3 shows that your real income this year and next year does indeed pick out a point on the budget line. And because we already know the slope of the budget line, we are done; we can now draw the budget line.

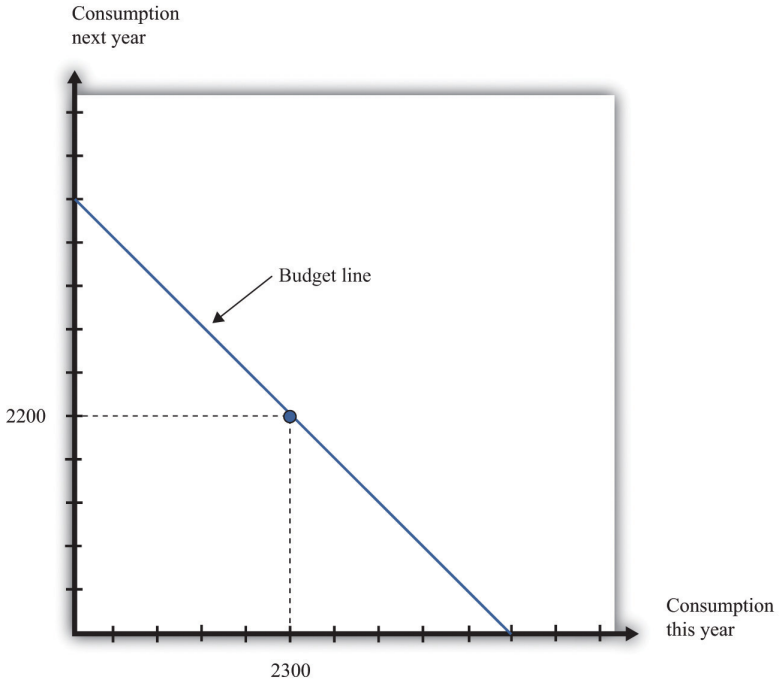


Figure 5.2.3: Determining the Position of the Budget Line

The position of the budget line depends on income this year and next year. We know that one possible choice of consumption is where the consumer neither saves nor borrows. This means that the budget line must pass through this point.

Adding Income and Consumption Spending over Two Periods

Your budget line describes the condition that total spending equals total income. This is true for the choice about the consumption of downloads and chocolate bars, and it is also true for the choice over time. But once we move to two periods, we must be careful about measuring both total income and total spending.

Adding Nominal Income over Two Periods

Suppose you earn some income this year (say, $23,000) and will earn some more next year (say $24,200). What is your income for these two years together? Your first instinct is probably to add the income in the two years and say $47,200. Superficially this makes sense—after all, income is measured in dollars in both years. Unfortunately, this is not a very good way to add money over time. It is flawed because it views income in two different years as if they are the same thing. In fact, money this year and money next year are not the same.

Imagine that a friend asks to borrow $1,000 from you today, promising to pay you back the $1,000 twenty years from now. Would you be likely to agree to this? Even if you trust your friend completely, the answer is surely no. After all, you could take your $1,000 and put it in the bank for twenty years, and the bank will pay you interest on your money—that is, the bank is willing to pay you for the privilege of using your money. Over twenty years, you could earn quite a bit of interest. By contrast, your friend is asking for a *zero-interest loan* in which no interest is paid on the money that you lend.

Positive interest rates mean that a dollar today and a dollar in the future are not worth the same. Adding dollars in one year to dollars in another year makes no more sense than adding apples and oranges. We need to convert dollars next year into their value right now. Remember that

$$z \; this \; year \; will \; be \; worth \; z \; \times \; the \; nominal \; interest \; factor \; next \; year.$$

We can turn this around. If the interest rate is 5 percent, then $105 next year will be worth only $100 this year. A dollar next year is worth dollars this year. This is the most you would be willing to give someone this year if he or she promised to give you a dollar

next year. You would not give them more than this because you would lose money relative to the alternative of putting the dollar in the bank and earning interest. More generally,

> **Toolkit: Section 31.5**
>
> Unless the interest rate is zero, a dollar this year is not the same as a dollar a year from now. To avoid this problem, economists use discounted present value as a device for measuring flows that occur over time. Discounted present value tells you the value of something you will receive in the future, discounted back to the present.

For example, if we want to add income in dollars over two years, the **discounted present value** of such a two-year *flow of income* is given by the following formula:

This is the income term that we need for the budget line in our two-year example. Go back to our earlier example, where income this year is $23,000 and income next year is $24,200, and the nominal interest factor is 1.1. Then

Even though income next year is higher in dollar terms, it is lower in terms of present value: $24,200 next year is worth only $22,000 today. Notice that when we measure the discounted value of a flow of nominal income, we still end up with a nominal value—the value of the income flow in terms of this year's dollars.

Table 5.2.1 "Discounted Present Value of Income" provides another illustration: it shows the calculation of the discounted present value of income when this year's income is $100 and next year's income is $200. You can see that, as the interest rate increases, the discounted present value of income decreases.

Table 5.2.1: Discounted Present Value of Income

Nominal Income This Year ($)	Nominal Income Next Year ($)	Discounted Present Value of Nominal Income Flow ($)		
		Nominal Interest Rate		
		0%	**5%**	**10%**
100	200	300.00	290.47	281.82
250	500			

Adding Nominal Consumption over Two Periods

When we want to add consumption spending this year and next year, measured in dollars, we use exactly the same logic as we did when adding income. Nominal consumption this year and next are given as follows:

$$nominal\ consumption\ this\ year\ =\ price\ level\ this\ year\ \times\ consumption\ this\ year$$

and

$$nominal\ consumption\ next\ year\ =\ price\ level\ next\ year\ \times\ consumption\ next\ year.$$

(Again, if you find it easier, just think of this as chocolate: total spending is the number of chocolate bars purchased times the price per bar. When we talk about "consumption," we mean something measured in real units, such as chocolate bars. When we talk about "nominal consumption," we are referring to a value measured in dollars.) Just as it is incorrect to add this year's and next year's income, so too should we not add together nominal consumption. Instead, we must calculate a discounted present value, exactly as we did before.

As with income, the discounted present value of nominal consumption is measured in this year's dollars.

Which Interest Factor Should You Use?

Earlier, we emphasized that people think about the real interest factor when they are comparing this year and next year. Yet in calculating the discounted present value of income and consumption spending, we are using the nominal interest factor. What is going on?

The rule for determining which interest factor to use in a discounted present value calculation is simple. *If you are converting nominal values, then you should use the nominal interest factor. If you are converting real values, then you should use the real*

interest factor. So if you want to know how much a given number of dollars in the future will be worth in dollars today, you should use the nominal interest factor. This is the normal case for most calculations that you would do. However, if you want to calculate a discounted present value for variables that have already been corrected for inflation, you must use the real interest factor. In this case, the answer you get is also a real quantity.

The Two-Period Budget Line Revisited

The tool of discounted present value gives us another way of thinking about the two-period budget line—the condition that

discounted present value of two-year flow of nominal consumption = discounted present value of two-year flow of nominal income.

Remember that both sides of this equation are measured in terms of this year's dollars. If we were to divide both sides of this equation by this year's price level, then we would get the equivalent expression in real terms:

$$discounted\ present\ value\ of\ two-year\ flow\ of\ consumption = discounted\ present\ value\ of\ two-year\ flow\ of\ real\ income.$$

In this case, as we just explained, the discounting must be done using the real interest factor instead of the nominal interest factor.

Income, Consumption, and Saving

Given your budget line, we suppose you choose a combination of consumption this year and next year that makes you as well off as possible. An example of such a preferred point is indicated in Figure 5.2.4.

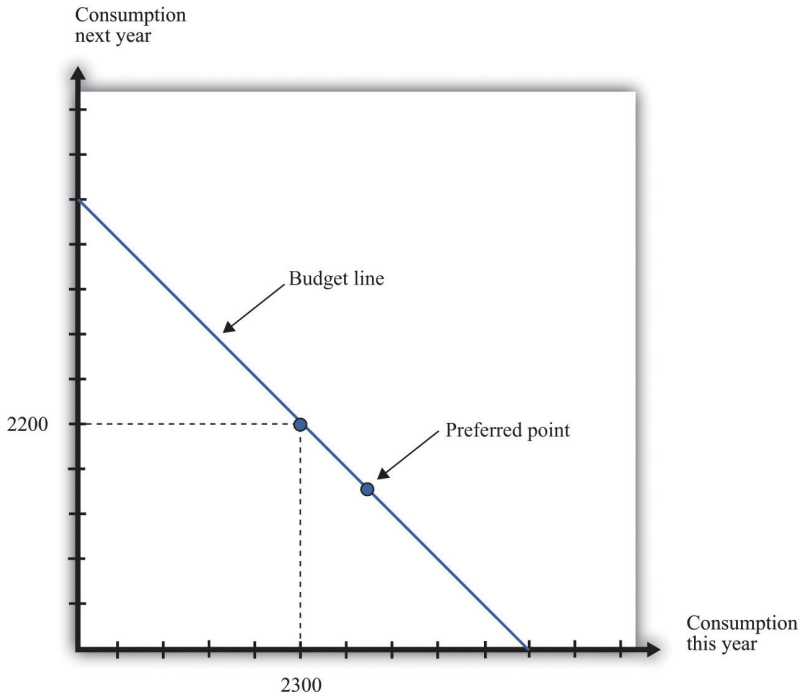


Figure 5.2.4: The Preferred Point

The consumer's preferred point must lie somewhere on the budget line. In this example, the consumer is choosing to consume in excess of his income this year. The consumer must borrow against future income, which means that consumption next year will be below next year's income.

The choice of a preferred point reflects two ideas. Whatever your tastes between consumption in the two years, we presume that you will not throw any income away. As a result, your choice will be on, not inside, the budget line. Further, if you choose well, according to your preferences, then you will pick the best combination of consumption; there is no other point on the budget line that you prefer.

Saving and Borrowing

Your preferred point implies a choice about how much saving or borrowing you do. Figure 5.2.5 shows two possible cases. In part (a) of Figure 5.2.5, you are a saver: you are consuming less than your income this year. The difference between your income and

your consumption is the amount of your savings. Those savings, plus interest, are available to you next year, so next year you can consume in excess of your income. In part (b) of Figure 5.2.5, you are a borrower: you are consuming more than your income this year. When you borrow this year, you must repay the loan with interest next year, so your consumption next year is less than your income.

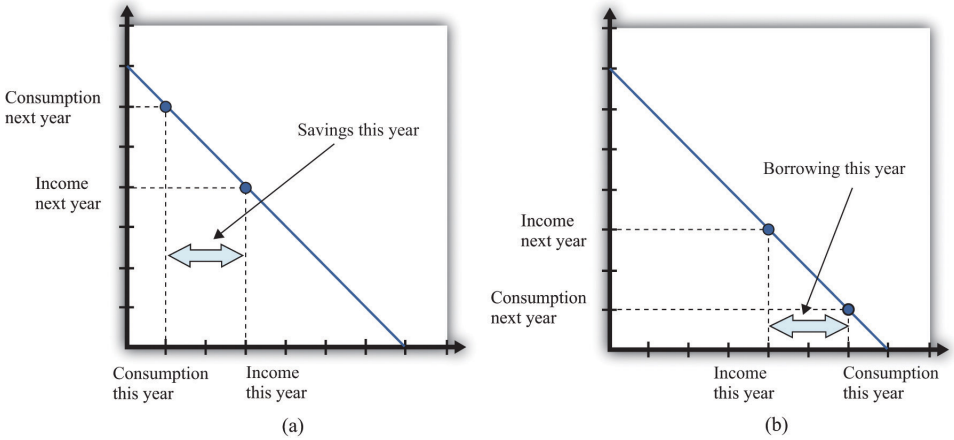


Figure 5.2.5: Consumption and Saving

(a) The individual is a saver this year. (b) The individual is a borrower this year.

The budget line tells you the rate at which the *market* allows you to substitute goods between this year and next year. This is distinct from your personal tastes about consuming this year or next year. The saver in part (a) of Figure 5.2.5 is a relatively patient person: she is willing to give up a lot of consumption this year to be able to consume more next year. The borrower in part (b) of Figure 5.2.5 is a relatively impatient person: he wants to consume a lot this year and is willing to sacrifice a great deal of future consumption.

You will sometimes hear discussions of how much individuals like to "discount the future." This is a statement about their tastes. Someone who discounts the future a great deal is impatient. Such a person wants to consume right away, so he will give up a lot of future consumption to have more today. Someone who discounts the future only a little is patient. Such a person is willing to give up consumption today even if she gets only a little extra consumption in the future. Economists pass no judgment on whether it is better to be impatient or thrifty. These are matters of personal preference.

The Timing of Income

Interestingly, the timing of your income turns out not to matter for your choice of consumption, which is illustrated in Figure 5.2.6.

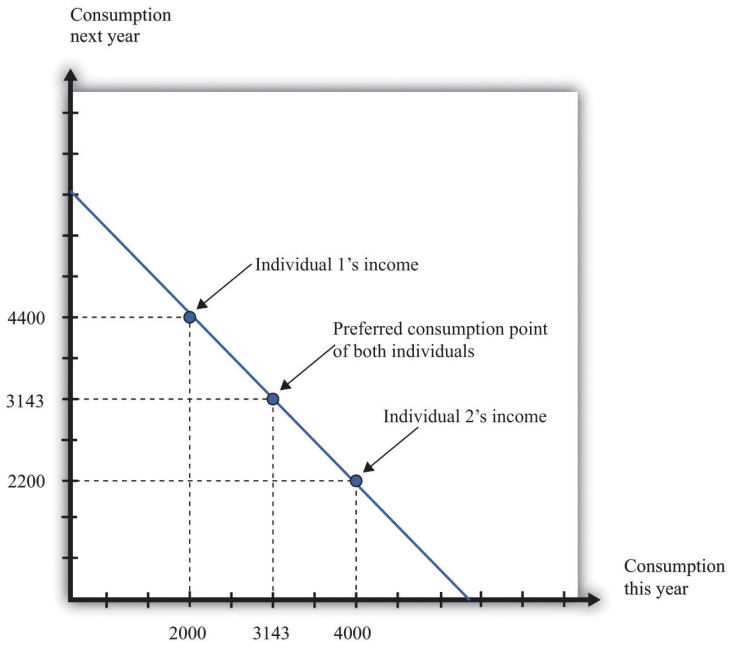


Figure 5.2.6: The Timing of Income

The timing of income is irrelevant to the consumption choice. Here, one individual has low income this year and high income next year, while the opposite is true for the other individual. However, the discounted present value of income is the same in both cases. If they both have the same tastes, they will choose the same consumption point.

Suppose that the nominal interest rate is 10 percent and that the price level is $10 in both periods. This means that the inflation rate is zero, so—from the Fisher equation—the real interest rate is also 10 percent. Now imagine that there are two individuals who have identical tastes. One of them earns income of $40,000 this year and $22,000 next year, so real income (nominal income divided by the price level) is therefore 4,000 this year and 2,200 next year. The other person earns $20,000 this year and $44,000 next year, yielding a real income of 2,000 this year and 4,400 next year. Both of these individuals share the same budget line (see Figure 5.2.6). This is because the discounted present value of their nominal income is the same: $60,000. (Check to make sure you understand why this is true.) For example, suppose that the preferred point of both individuals is to consume the same amount in each year. Then they can both consume approximately 3,143 units of consumption in each period.

Let us see how this works. Because the price level is $10, this amount of consumption costs $31,430 in each period. The first individual takes her income of $40,000 and saves $8,570 by putting it in the bank. This saving earns 10 percent interest, so she gets an additional $857. She thus has income in the following year equal to $22,000 + $8,570 + $857 = $31,427. This allows her to buy 3,142.7 units of consumption goods. The second individual needs to borrow $11,430 to add to his income this year. Next year, he must repay this amount plus 10 percent interest (that is, another $1,143). So his income next year is $44,000 − $11,430 − $1,143 = $31,427. So one individual must save to reach her preferred consumption bundle, while the other must borrow to reach his. Yet because they have the same discounted present value of income and the same tastes, they will consume the same bundle of goods.

Keep in mind that our discussion so far ignores uncertainty. We assumed that both individuals know their current and future income with certainty. Just as importantly, we have supposed that a bank is confident that the borrower will have sufficient income next year to repay the loan. In a world of uncertainty, we do not know for sure how much money we will have next year, and lenders worry about the possibility that people might not make good on their loans. Later in the chapter, we explain more about decision making in an uncertain world.

Lifetime Budget Constraint

So far, we have worked everything out in terms of a two-period example. The two-period budget constraint tells us how income and consumption are linked over time. In reality, of course, you make these decisions with longer time horizons, and you can save or borrow for multiple years. But the same fundamental insight holds. If you save this year, then you will have extra resources to spend at some future date. If you borrow this year, then you will have to repay that loan sometime in the future, at which time you will have fewer resources to spend.

Toolkit: Section 31.4

Individuals face a lifetime budget constraint. They can save in some periods of their lives and borrow (not save) in other periods. Over the course of any individual's lifetime, however, income and consumption spending must balance. (If you begin life with some assets [for example, a bequest], we count this as part of income. If you leave a bequest when you die, we count this as part of consumption.) The lifetime budget constraint is as follows:

$$discounted\ present\ value\ of\ lifetime\ consumption\ =\ discounted\ present\ value\ of\ lifetime\ income.$$

Again, it is important to be consistent in calculating the discounted present values in this expression. We have written the equation in terms of (real) consumption and (real) income, which means that the real interest factor must be used for discounting. An alternative is to measure both consumption and income in nominal terms and then use the nominal interest factor for discounting. There is a useful special case where real interest rates are zero, in which case it *is* legitimate simply to add income and consumption in different years. Thus the **lifetime budget constraint** becomes

$$total\ lifetime\ consumption\ =\ total\ lifetime\ income.$$

Although the principles of decision making are the same whether we are thinking about 2 months, 2 years, or an entire lifetime, it is obviously harder to make decisions over a 30-year horizon than over a 30-day horizon. One reason is that, over longer time horizons, we are more likely to face uncertainty. We don't know what our income will be 30 years from now, and we don't know our tastes. But even without that uncertainty, we may not always make good decisions.

In particular, economists and psychologists have discovered that we do not view choices involving the near future the same way as we view distant choices. For example, suppose an individual is given a choice between 1 cookie today or 2 cookies tomorrow. If he is impatient (or hungry), he is likely to choose the single cookie today. But if the same individual is given a choice between 1 cookie in 30 days or 2 cookies in 31 days, he or she may very well choose the 2-cookie option. Yet after 30 days have gone by, that person will be confronting the earlier decision, wishing that he or she could have the 2 cookies today.

Another way of saying this is that our decisions are not always consistent over time: our future selves may wish that our current selves had displayed more self-control. For example, we may choose to consume a lot today—instead of saving—and then regret that decision when we are older. Indeed, people often engage in tricks to get around their lack of self-control. For example, some people have a separate bank account for their savings, so they are less tempted to spend that money. Governments also take actions that compensate for our lack of self-control. Social security is in some ways a "forced saving" scheme: the government takes money from us when we are working but pays us money when we are retired.

There is a new and exciting field of economics called "neuroeconomics" that tries to understand the processes in the brain that underlie economic decision making. This field, while still very much in its infancy, promises to help us understand why the economic theory of how we make choices often works well, and why it sometimes does not. Some recent research suggests that different brain processes may deliver conflicting messages when making choices over time. Some processes are deliberative, in line with the economic model, while others are more impulsive. It is likely that the next two decades will bring a much deeper understanding of how the brain makes decisions, perhaps leading to a richer theory of economic decision making.

Changes in Interest Rates

Whenever the real interest rate changes, then the *relative price* of consumption this year and next year changes. As we already know, changes in the real interest rate can come from two different sources: changes in the nominal interest rate and changes in the inflation rate. (Look back at the Fisher equation for a reminder of this.) Figure 5.2.7 shows the effect of an increase in the real interest rate on your budget line. The budget line becomes steeper because the opportunity cost of consumption this year increases. Notice, though, that the point at which you just consume your income in each period is still on the budget line. This is the point at which you are neither saving nor borrowing. Thus no matter what the interest rate, this point is always available to you.

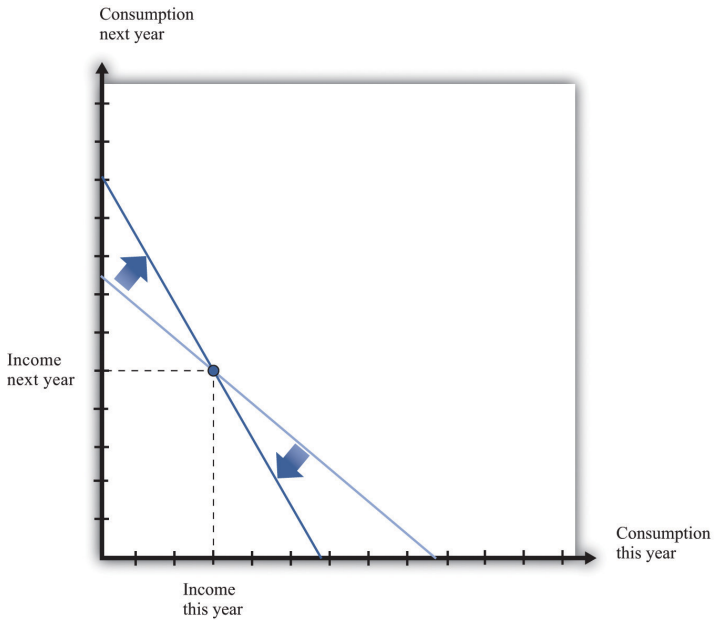


Figure 5.2.7: An Increase in the Real Interest Rate

A change in the real interest rate changes the slope of the budget line. At any real interest rate, however, it is possible to consume exactly one's income. So the point corresponding to no saving or borrowing is always available, no matter what the real interest rate. An increase in the real interest rate therefore causes the budget line to become steeper and rotate through the income point.

Changes in relative prices lead to income and substitution effects. To understand the effect of an increase in the real interest rate, we must look at both effects.

- **Substitution effect.** An increase in the real interest rate makes consumption next year look more attractive relative to consumption this year. This encourages saving and discourages borrowing.
- **Income effect.** The income effect is somewhat more complicated. Look again at Figure 5.2.7. *Savers* are made better off by an increase in interest rates because, to the left of the income point, the new budget line lies outside the old budget set. This encourages savers to increase consumption this year and next year. Because income this year hasn't changed, but there is an incentive to consume more this year, we can see that the income effect discourages saving. *Borrowers*, meanwhile, are worse off by the increase in interest rates. To the right of the income point, the new budget line lies inside the old budget set. This encourages borrowers to decrease consumption this year and next year. This incentive to consume less tells us that the income effect discourages borrowing. In sum, the income effect gives an incentive for savers to save less and for borrowers to borrow less.

Combining the income and substitution effects and following an increase in the interest rate, borrowers have an incentive to borrow less. The substitution effect encourages saving, while the income effect discourages saving. The overall effect is ambiguous.

The evidence suggests that most people are like the individual in Figure 5.2.8. For this person, the substitution effect dominates: the amount of saving increases as the real interest rate increases. Because an individual's savings represent funds that can be lent out to others in the economy, we call them the **individual loan supply**.

Toolkit: Section 31.6

Individual loan supply is the amount of saving carried out by an individual at different values of the real interest rate. It is illustrated in a diagram with the real interest rate on the vertical axis and the supply of loans on the horizontal axis.

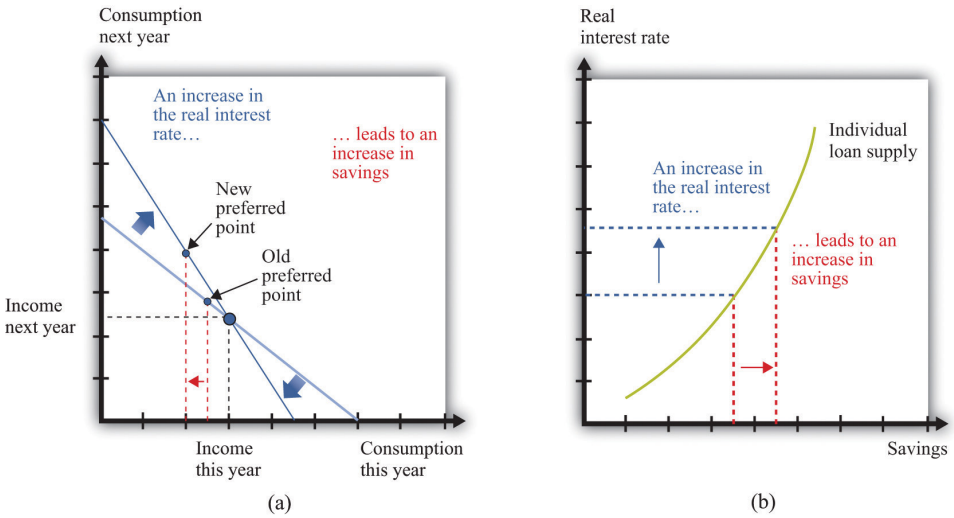


Figure 5.2.8: Individual Loan Supply

For savers in the economy, the effects of an increase in the real interest rate are ambiguous. The substitution effect encourages saving, but the income effect discourages saving. The evidence suggests that, on balance, the substitution effect dominates, so that savings increase. (a) In this two-period diagram, an increase in interest rates causes consumption this year to decrease. Because income this year is unchanged, savings increases. (b) The same diagram is applied to an individual supply of loans.

As the real interest rate changes, the response of individual saving is a *movement along* the loan supply curve. What might cause the whole curve to shift? If an individual has a higher income in the current year, this will cause the budget line to shift outward, and the person will consume more goods in the current year *and* more goods in the future. To consume more in the future, the person will have to save more. In this case, the supply of savings shifts outward as current income increases. This is shown in Figure 5.2.9.

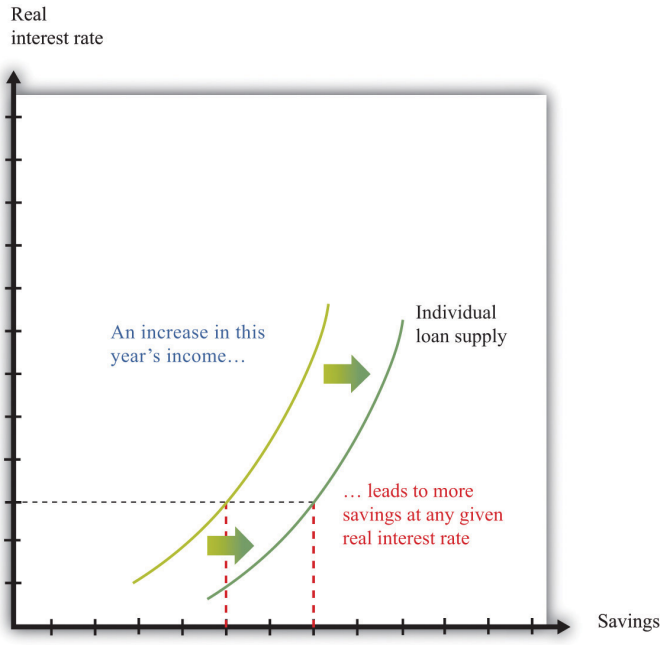


Figure 5.2.9: A Shift in an Individual's Supply of Savings. An increase in this year's income means that an individual will save more at any given interest rate. This means that the loan supply curve for the individual shifts outward.

Key Takeaways

- Over the course of an individual's lifetime, the discounted present value of spending equals the discounted present value of income.
- Households save to consume more in the future.

- Unless the interest rate is zero, a dollar today does not have the same value as a dollar tomorrow.
- The nominal interest rate is expressed in dollar terms, while the real interest rate is expressed in terms of goods and services. Economists think that households and firms make decisions on the basis of real interest rates.

check your understanding

1. Fill in the missing values in Table 5.2.1 "Discounted Present Value of Income".
2. If the interest rate increases, what will happen to the amount saved by a household? How does this answer depend on whether the household is a lender (saving is positive) or a borrower (saving is negative)?

This page titled 5.2: Consumption and Saving is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

5.3: Using Discounted Present Values 5.3: Using Discounted Present ValuesEdit Update titlecancel

Learning Objectives

1. When should you use the tool of discounted present value?
2. How does an increase in the interest rate affect the discounted present value of a flow of income?

Section 5.2 introduced a valuable technique called discounted present value. You can use this technique whenever you need to compare flows of goods, services, or currencies (such as dollars) in different periods of time. In this section, we look at some of the big decisions you make during your life, both to illustrate discounted present value in action and to show how a good understanding of this idea can help you make better decisions.

Choosing a Career

A decision you typically make around the time that you graduate is your choice of a career. What makes the choice of a career so consequential is the fact that it can be very costly to switch from one career to another. For example, if you have trained as an engineer and then decide you want to be a lawyer, you will have to give up your engineering job (and give up your salary as well) and go to law school instead.

Suppose you are choosing among three careers: a lawyer, an insurance salesperson, or a barista. To make matters simple, we will work out an example with only two years. Table 5.3.1 "Which Career Should You Choose?" shows your earnings in each year at each occupation. In the first year in your career as a lawyer, we suppose that you work as a clerk, not earning very much. In the second year, you join a law firm and enjoy much higher pay. Selling insurance pays better than the legal career in the first year but worse in the second year. Working as a barista pays less than selling insurance in both years.

Table 5.3.1: Which Career Should You Choose?

Career	First-Year Income ($)	Second-Year Income ($)
Lawyer	5,000	60,000
Insurance salesperson	27,000	36,000
Barista	18,000	20,000

It is obvious that, if you care about the financial aspect of your career, you should not be a barista. (You would choose that career only if it had other benefits—such as flexible working hours and lack of stress—that outweighed the financial penalty.) It is less obvious whether it is better financially to work as a lawyer or as an insurance salesperson. Over the two years, you earn $65,000 as a lawyer and $63,000 as an insurance seller. But as we have already explained, simply adding your income for the two years is incorrect. The high salary you earn as a lawyer comes mostly in the second year and must be discounted back to the present.

To properly compare these careers, you should use the tool of discounted present value. With this tool, you can compare the income flows from the different occupations. Table 5.3.2 "Comparing Discounted Present Values of Different Income Streams" shows the discounted present value of the two-year flow of income for each career, assuming a 5 percent interest rate (that is, an interest factor equal to 1.05). Look, for example, at the lawyer’s income stream:

Similarly, the discounted present value of the income stream is $61,286 for the insurance salesperson and $37,048 for the barista. So if you are choosing your career on the basis of the discounted present value of your income stream, you should pick a career as a lawyer.

Table 5.3.2: Comparing Discounted Present Values of Different Income Streams

Career	First-Year Income ($)	Second-Year Income ($)	Discounted Present Value at 5% Interest Rate ($)
Lawyer	5,000	60,000	62,143
Insurance salesperson	27,000	36,000	61,286

Career	First-Year Income ($)	Second-Year Income ($)	Discounted Present Value at 5% Interest Rate ($)
Barista	18,000	20,000	37,048

This conclusion, however, depends on the interest rate used for discounting. Table 5.3.3 "Discounted Present Values with Different Interest Rates" adds another column, showing the discounted present values when the interest rate is 10 percent. You can see two things from this table: (1) The higher interest rate reduces the discounted present value for all three professions. If the interest rate increases, then future income is less valuable in present value terms. (2) The higher interest rate reverses our conclusion about which career is better. Selling insurance now looks better than being a lawyer because most of the lawyer's earnings come in the future, so the discounting has a bigger effect.

Table 5.3.3: Discounted Present Values with Different Interest Rates

Career	First-Year Income ($)	Second-Year Income($)	Discounted Present Value at 5% Interest Rate ($)	Discounted Present Value at 10% Interest Rate ($)
Lawyer	5,000	60,000	62,143	59,545
Insurance salesperson	27,000	36,000	61,286	59,727
Barista	18,000	20,000	37,048	36,182

Of course, what you might really like to do is to sell insurance for the first year and work as a lawyer in the second year. This evidently would have higher income. Sadly, it is not possible: it is almost impossible to qualify for a high-paying lawyer's job without investing a year as a law clerk first. Changing occupation can be very costly or even impossible if you don't have the right skills. So choosing a career path means you must look ahead.

Going to College

If you are like most readers of this book, you have already made at least one very important decision in your life. You have chosen to go to college rather than taking a job immediately after graduating from high school. Ignoring the pleasures of going to college—and there are many—there are direct financial costs and benefits of a college education.

Think back to when you were deciding whether to go to college or to start work immediately. To keep our example from being too complicated, we again look at a two-year decision. What if you could obtain a college degree in one year, at a tuition cost of $13,000, and the interest rate is 5 percent annually? Your earnings are presented in Table 5.3.4 "Income from Going to College versus Taking a Job". In your year at college, you would earn no income, and you have to pay the tuition fee. In the following year, imagine that you can earn $62,143 working as a lawyer. Alternatively, you could bypass college and go to work as a barista, earning $10,000 in the first year and $37,048 in the second year. (We are assuming, as before, that you know these figures with certainty when you are making your decision.)

Table 5.3.4: Income from Going to College versus Taking a Job

Career	Income in the Year at College ($)	Income in the Year after College ($)
College	−13,000	62,143
Barista	10,000	37,048

Going to college is an example of an *investment decision*. You incur a cost in the year when you go to college, and then you get a benefit in the future. There are two costs of going to college: (1) the $13,000 tuition you must pay (this is what you probably think of first when considering the cost of going to college) and (2) the *opportunity cost* of the income you could have earned while working. In our example, this is $10,000. The explicit cost and the opportunity cost together total $23,000, which is what it costs you to go to college instead of working in the first year.

By the way, we do *not* think about living expenses as a cost of going to college. You have to pay for food and accommodation whether you are at college or working. Of course, if these living expenses are different under the two scenarios, then you should

take this into account. For example, if your prospective college is in New York City and has higher rental costs than in the city where you would work, then the *difference* in the rent should be counted as another cost of college.

The benefit of going to college is the higher future income that you enjoy. In our example, you will earn $62,143 in the following year if you go to college, and $37,048 if you do not. The difference between these is the benefit of going to college: $62,143 − $37,048 = $25,095. Even though this is greater than the $23,000 cost of going to college, we cannot yet conclude that going to college is a good idea. We have to calculate the discounted present value of this benefit. Suppose, as before, that the interest rate is 5 percent. Then

We can conclude that, with these numbers, going to college is a good investment. It is worth $900 more in discounted present value terms.

We could obtain this same conclusion another way. We could calculate the discounted value of the two-year income stream for the case of college versus barista, as in Table 5.3.5 "Income Streams from Going to College versus Taking a Job". We see that the discounted value of the income stream if you go to college is $46,184, compared to $45,284 if you work as a barista. The difference between these two is $900, just as before.

Table 5.3.5: Income Streams from Going to College versus Taking a Job

Career	Income in the Year at College ($)	Income in the Year after College ($)	Discounted Present Value at 5% Interest Rate ($)
College	−13,000	62,143	46,184
Barista	10,000	37,048	45,284

You might have noticed that the figures we chose as “income in the year after college” in Table 5.3.4 "Income from Going to College versus Taking a Job" are the same as the numbers that we calculated in Table 5.3.2 "Comparing Discounted Present Values of Different Income Streams". The numbers in Table 5.3.2 "Comparing Discounted Present Values of Different Income Streams" were themselves the result of a discounted present value calculation: they were the discounted present value of a two-year income stream. When we compare going to college with being a barista, we are therefore calculating a discounted present value of something that is already a discounted present value. What is going on?

To understand this, suppose you are deciding about whether to go to college in 2012. If you do go to college, then in 2013 you will decide whether to be a lawyer, an insurance salesman, or a barista. If you decide on the legal career, then you will be a law clerk in 2013, and you will earn the high legal salary in 2014. Our analysis in Table 5.3.2 "Comparing Discounted Present Values of Different Income Streams" is therefore about the choice you make in 2013, thinking about your income in 2013 and 2014. Table 5.3.2 "Comparing Discounted Present Values of Different Income Streams" gives us the discounted present value *in 2013* for each choice. If we then take those discounted present values and use them as “income in the year after college,” as in Table 5.3.5 "Income Streams from Going to College versus Taking a Job", we are in fact calculating the discounted present value, in 2012, of the flow of income you receive in 2012, 2013, and 2014.

If you think carefully about this, you will realize that

This is the same answer that we got before. As this example suggests, you can calculate discounted present values of long streams of income, including income you will receive many years in the future.The toolkit gives a more general formula for calculating the discounted present value. (See the more formal presentation of discounted present value at the end of Chapter 5, Section 5.4.)

Economists have worked hard to measure the return on investment from schooling: “Alan B. Krueger, an economics professor at Princeton, says the evidence suggests that, up to a point, an additional year of schooling is likely to raise an individual’s earnings about 10 percent. For someone earning the national median household income of $42,000, an extra year of training could provide an additional $4,200 a year. Over the span of a career, that could easily add up to $30,000 or $40,000 of present value. If the year’s education costs less than that, there is a net gain.”Anna Bernasek, “What’s the Return on Education?” Economic View, Business Section, *New York Times*, December 11, 2005, accessed February 24, 2011, http://www.nytimes.com/2005/12/11/business/yourmoney/11view.html. Notice several things from this passage. First, the gains from education appear as an increase in earnings *each year*. So even if a 10 percent increase in earnings does not seem like a lot, it can be substantial once these gains are added over one’s lifetime. Second, Krueger is careful to use the term *present value*. Third, the number given is an average. Some people will benefit more; others will benefit less. Equally, some forms of schooling will

generate larger income gains than others. Fourth, Krueger correctly notes that the present value must be compared with the cost of education, but you should remember that the cost of education includes the opportunity cost of lost income.

Table 5.3.6 "Return on Education" provides some more information on the financial benefits of schooling.US Census Bureau, “Income, Earnings, and Poverty from the American Community Survey,” 2010, table 5, accessed February 24, 2011, www.census.gov/hhes/www/poverty/publications/acs-01.pdf. The table shows average income in 2004. There is again evidence of a substantial benefit from schooling. Male college graduates, on average, earned more than $21,000 (68 percent) more than high school graduates, and female college graduates earned more than $16,000 (78 percent) more than high school graduates. The table shows that women are paid considerably less than men and also that the return on education is higher for women.

Table 5.3.6: Return on Education

Schooling	Median Annual Income (Men)	Median Annual Income (Women)
High school graduate	31,183	19,821
Some college	37,883	25,235
College graduate	52,242	35,185

Source: US Census Bureau, “Income, Earnings, and Poverty From the 2004 American Community Survey,” August 2005, table 5, page 10, accessed March 14, 2011, www.census.gov/hhes/www/poverty/publications/acs-01.pdf.

The presence of such apparently large gains from education helps explain why economists often suggest that education is one of the most important ingredients for the development of poorer countries. (In poorer countries, we are often talking not about the benefit of going to college but about the benefit of more years of high school education.) Moreover, the benefits from education typically go beyond the benefits to the individuals who go to school or college. There are benefits to society as a whole as well.

However, you should be careful when interpreting numbers such as these. We cannot conclude that if you randomly selected some high school graduates and sent them to college, then their income would increase by $17,000. As we all understand, individuals *decide* whether to go to college. These decisions reflect many things, including general intelligence, the ability to apply oneself to a task, and so on. People who have more of those abilities are more likely to attend—and complete—college.

One last point: we conducted this entire discussion “ignoring the pleasures of going to college.” But those pleasures belong in the calculations. Economics is about not only money but also all the things that make us happy. This is why we occasionally see people 60 years old or older in college. They attend not as an investment but simply because of the pleasure of learning. This is not inconsistent with economic reasoning or our discussion here. It is simply a reminder that your calculations should not only be financial but also include the all the nonmonetary things you care about.

The Effects of Interest Rates on Labor Supply

After you have decided whether or not to go to college and have chosen your career, you will still have plenty of decisions to make involving discounted present value. Remember that you have a lifetime budget constraint, in which the discounted present value of your income, including labor income, equals the discounted present value of consumption.

Your labor income is partly under your control. If you have some choice about how many hours to work, then your **individual labor supply curve** depends on the real wage, as shown in Figure Runaway argument for \eatSpaces? Chapter 4 explains the individual supply of labor. The labor supply curve illustrates the fact that as the real wage increases, you are likely to work more. Labor supply, like the supply of loans that we considered previously, is driven by substitution and income effects. As the real wage increases,

- the opportunity cost of leisure increases, so you have an incentive to work more (substitution effect); and
- you have more income to spend on everything, including leisure, so you have an incentive to consume more leisure and work less (income effect).

Toolkit: Section 31.3

The toolkit contains more information if you want to review the labor supply curve.

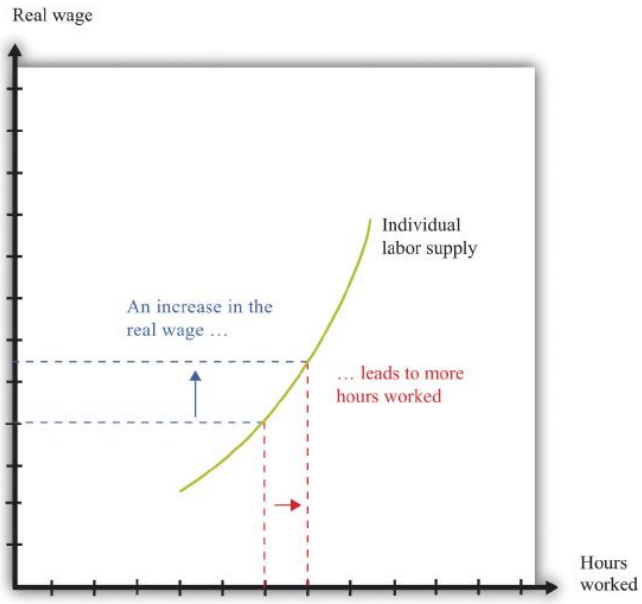


Figure 5.3.1: Individual Labor Supply

An increase in the real wage encourages an individual to work more. The labor supply curve slopes upward.

When you are making a decision about how much to work over many periods of time, your choice is more complicated. How much you choose to work right now depends not only on the real wage today but also on the real wage in the future and on the real interest rate. This is because you work both today and in the future to earn the income that goes into your lifetime budget constraint. If you think it is likely to be harder to earn money in the future, then you will probably decide to work more today. If you think it will be easier to earn money in the future, then you might well decide to work less today.

It is easiest to see how this works with an example. Suppose you are a freelance construction worker in Florida in the aftermath of a hurricane. There is lots of work available, and construction firms are paying higher than usual wages. You realize that you can earn much more per hour of work right now compared to your likely wage a few months in the future. A natural response is to work harder now to take advantage of the unusually high wages.

We can understand your decision in terms of income and substitution effects. The higher wage leads to the usual substitution effect. But because the change in the wage is only temporary, and because you are thinking about your wages over your lifetime, it does not have a large income effect. In this case, therefore, we expect the substitution effect to strongly outweigh the income effect.

Interest rates may also influence your decision about how hard to work. If interest rates increase, then the gains from working today increase as well. If you save money at high interest rates, you can enjoy more consumption in the future. High interest rates, like temporarily high current wages, increase the return to working today compared to working in the future, so you are likely to work more today.

The Demand for Durable Goods

Some of the products we purchase, such as milk or a ticket to a football game, disappear as soon as they are consumed. Other goods last for a long time and are, in effect, consumed over and over. Some examples are a bicycle, a car, and a microwave oven.

Goods that last over many uses are called **durable goods**, while those that do not last very long are called **nondurable goods**. There is no hard-and-fast distinction between durable and nondurable goods. Many everyday items, such as plates, books, T-shirts, and downloaded music, are used multiple times. In economic statistics, however, the term *durable* is reserved for larger items that are bought only occasionally and that typically last for many years. Cars and kitchen appliances are classified as durable goods, but blue jeans and haircuts are not, even though they are not consumed all at once.

Because durable goods last a long time, making decisions about purchasing a durable good requires thinking about the future as well as the present. You must compare the benefits of the durable good *over its entire lifetime* relative to the cost you incur to pay for the durable now. A durable good purchase is typically a **unit demand** decision—you buy either a single unit or nothing. For unit demand, your decision rule is simple: buy if your **valuation** of the good exceeds the price of the good. (Remember that your

valuation is the maximum you would be willing to pay for the good.) In the case of durable goods, there is an extra twist: your valuation needs to be a discounted present value.

Toolkit: Section 31.1

You can review valuation and unit demand in the toolkit.

The idea is that you obtain a *flow of services* from a durable good. You need to place a valuation on that flow for the entire lifetime of the durable good. Then you need to calculate the discounted present value of that flow of services. If this discounted present value exceeds the price of the good, you should purchase it.

Suppose you are thinking of buying a new car that you expect to last for 10 years. You need to place a valuation on the flow of services that you get from the car each year: for example, you might decide that you are willing to pay $3,000 each year for the benefit of owning the car. To keep life really simple, let us think about a situation where the real interest rate is zero; this is the special case where it *is* legitimate to add these flows. So the car is worth $30,000 (= $3,000 per year × 10 years) to you now. This means you should be willing to buy the car if it costs less than $30,000, and you should not buy it otherwise.What if you think you might sell the car before it wears out? In this case, the value of the car has two components: (1) the flow of services you obtain while you own it, and (2) the price you can expect to obtain when you sell it. When you buy a durable good, you are purchasing an **asset**: something that yields some flow of benefits over time and that you can buy and sell. In general, the value of an asset depends on both the benefits that it provides and the price at which it can be traded. We examine these ideas in much more detail in Chapter 10.

If real interest rates are not zero, then spending on durable goods will depend on interest rates. As interest rates increase, the future benefits of the durable good become smaller, in terms of discounted present value. This means that durable goods become more expensive relative to nondurable goods. Thus the demand for durable goods decreases as interest rates increase.

One way to understand this is to realize that it is often easy to defer the purchase of a durable good. New durable goods are frequently bought to replace old goods that are wearing out. People buy new cars to replace their old cars or new washing machines to replace their old ones. If interest rates are high, you can often postpone such replacement purchases; you decide whether you can manage another year with your old car or leaking washing machine. As a result, spending on durable goods tends to be very sensitive to changes in interest rates.

These examples of discounted present value illustrate one key point: whenever you are making economics decisions about the future—be it what career to follow, when it is best to work hard, or if you should buy a new car—your decisions depend on the rate of interest. Whenever the rate of interest is high, future costs and benefits are substantially discounted and are therefore worth less in present value terms. High interest rates, in other words, mean that you put a lot of weight on the present relative to the future. When the rate of interest decreases, the future should play a larger role in your decisions.

Key Takeaways

- You should use discounted present value whenever you need to compare flows of income and expenses in different periods of time.
- The higher the interest rate, the lower the discounted present value of a flow.

check your understanding

1. List five goods that are durable and five services that are nondurable. Is it possible to have a service that is durable?
2. Calculate the appropriate values in Table 5.3.2 "Comparing Discounted Present Values of Different Income Streams" assuming an interest rate of 8 percent.
3. If interest rates are lower, would you expect people to invest more or less in their health?

This page titled 5.3: Using Discounted Present Values is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

5.4: Avoiding Risk 5.4: Avoiding RiskEdit Update titlecancel

Learning Objectives

1. What is the definition of probability?
2. How can I calculate an expected value?
3. What is risk aversion?
4. How do individuals and firms deal with risk?
5. How does insurance work?

In life, there are many uncertainties. So far, we have ignored them all, but you will have to face them. In our various discussions of discounted present value, we pretended that you knew your future income—and your future tastes—with certainty. In real life, we must decide how much to save without knowing for sure what our future income will be. We must pick a career without knowing how much we will enjoy different jobs or how much they will pay. We must decide whether or not to go to college without knowing what kind of job we will be able to get, and so on. How can we deal with all these uncertainties?

Some of the uncertainties we face are forced on us with no choice of our own, such as the following:

- Accidents involving you, your automobile, your house, and so on
- Layoffs resulting in spells of unemployment
- Your health

As you know, one way to deal with these uncertain events is through insurance. Insurance is a way of trying to remove some of the risk that we face. We explain how it works later in this section.

Other risks are more under our control. We accept jobs that entail certain risks. We drive our cars even though we know that there is a risk of accident. We put our savings into risky stocks rather than safe assets. In these cases, we trade off these risks against other benefits. We drive faster, accepting the greater risk of accident to save time. Or we take a risky job because it pays well.

There are yet other kinds of risk that we actually seek out rather than avoid. We play poker or bet on sporting events. We climb mountains, go skydiving, and engage in extreme sports. In these cases, the risks are apparently something good that we seek out, rather than something bad that we avoid.

Risk and Uncertainty

Let us begin by making sure we understand what risk and uncertainty mean. (Here we will use the terms more or less interchangeably, although people sometimes reserve the term *uncertainty* for cases where it is hard to quantify the risks that we face.) Probably the simplest example of risk is familiar to us all: the toss of a coin. Imagine flipping a coin five times. Each time, the outcome will be either a head or a tail. Table 5.4.1 "Coin-Flipping Experiment" shows an example of such an experiment. In this experiment, the outcome was three heads and two tails. For each flip of the coin, there was uncertainty about the outcome. We did not know ahead of time whether there would be heads or tails. The outcome reported in Table 5.4.1 "Coin-Flipping Experiment" is only one example. If you were to carry out this experiment right now, you would almost certainly end up with a different outcome.

Table 5.4.1: Coin-Flipping Experiment

Flip	1	2	3	4	5
Outcome	Heads	Heads	Tails	Heads	Tails

Coin tosses are special because the flips of the coin are *independent* of each other (that is, the history of previous tosses has no effect on the current toss of a coin). In Table 5.4.1 "Coin-Flipping Experiment", the coin was not more likely to come up tails on the third toss because the previous tosses were both heads. Even if you have 100 heads in a row, this does not affect the outcome of the 101st toss of the coin. If you think that the coin is “fair,” meaning that heads and tails are equally likely, then the 101st toss is still just as likely to be heads as tails. By contrast, the likelihood that it will be raining an hour from now is not independent of whether or not it is raining at this moment.

Financial Risk and Expected Value

Some of the risks that we confront are nonfinancial. An example of nonfinancial uncertainty is the risk that you might break your ankle playing basketball or the possibility that your favorite sporting team will win a big game and make you happy. Here, we will focus on financial uncertainty, by which we mean situations where there is money at stake. In other words, we are thinking about risks where you can measure the implications in monetary terms. An obvious example is the money you could win or lose from buying a lottery ticket or playing poker. Another is the money you would have to pay for repairs or medical expenses following a car accident. Another is the gains or losses from buying stocks, government bonds, or other financial assets. Another is the income you would lose if you were laid off from your job.

When we evaluate risky situations, we must have a way of describing the kinds of gambles that we confront. In general, we do this by listing all the possible outcomes together with the likelihood of each outcome. For example, Table 5.4.2 "Outcomes and Probabilities from a Coin Toss" lists the outcomes and the **probability** (that is, the likelihood of each outcome) for the experiment of tossing a coin one time.

Table 5.4.2: Outcomes and Probabilities from a Coin Toss

Outcome	Probability
Heads	
Tails	

Toolkit: Section 31.7

Probability is the percentage chance that something will occur. For example, there is a 50 percent chance that a tossed coin will come up heads. We say that the probability of getting the *outcome* “heads” is 0.5.

There are five things to know about probability:

1. The list of outcomes must be complete.
2. The list of outcomes must not overlap.
3. If an outcome is certain to occur, it has probability of 1.
4. If an outcome is certain *not* to occur, it has probability of 0.
5. If we add the probabilities for all the possible outcomes, the total must equal 1.

Think about rolling a normal six-sided die one time and describing outcomes and probabilities.

- We must make sure that we include every outcome. We cannot list as possible outcomes “less than or equal to 2” and “greater than or equal to 4.” Such a list ignores the possibility of rolling a 3.
- We cannot list as possible outcomes “less than or equal to 4” and “greater than or equal to 3.” These categories overlap because a roll of a 3 or a 4 would show up in both categories.
- The outcome “less than or equal to 6” has a probability of 1 because it is certain.
- The outcome “9” has a probability of 0.
- Provided we have a complete list of outcomes (for example “less than or equal to 4” and “greater than or equal to 5”), the probabilities of all the outcomes will always sum to 1. (In this case, the probability of the first outcome is 2/3, and the probability of the second outcome is 1/3.)

Now suppose you are playing a gambling game based on a toss of a coin. If the coin comes up heads, you win $1. If it comes up tails, you win $0. When we look at a situation such as this, we are often interested in how much you would get, on average, if you played the game many times. In this example, it is easy to guess the answer. On average, you would expect to win half the time, so half the time you get $1, and half the time you get nothing. We say that the **expected value** of each flip of the coin is 50 cents.

Toolkit: Section 31.7

The expected value of a situation with financial risk is the measure of how much you would expect to win (or lose) on average, if the situation were to be replayed a large number of times. Expected value is calculated as follows:

1. For each outcome, multiply the probability of that outcome by the amount you will receive.

2. Add these amounts over all the possible outcomes.

Table 5.4.3 "Outcomes and Probabilities from Investment in Internet Venture" gives another example of expected value. Suppose a friend is planning on establishing a small Internet business and asks you to invest $1,000. He tells you (and you believe him) that there is a 50 percent chance that the business will fail, so you will lose your money. There is a 40 percent chance that the business will just break even, so you will get your $1,000 back but nothing more. And there is a 10 percent chance that the business will be very successful, so you will earn $16,000.

Table 5.4.3: Outcomes and Probabilities from Investment in Internet Venture

Outcome	Probability	Amount You Will Receive ($)
Failure	0.5	0
Break even	0.4	1,000
Success	0.1	16,000

In this case, the expected value of the investment is given by the following:

$$expected\ value = (0.5 \times \$0) + (0.4 \times \$1,000) + (0.1 \times \$16,000) = \$2,000.$$

Thus for your investment of $1,000, you could expect to get $2,000 back on average. This seems like a good investment. It is important to remember, though, what “on average” means. You will never actually *get* $2,000. You will receive either $16,000, $1,000, or nothing. Even though this is a good investment on average, you might still decide that you don’t want any part of it. Yes, you might get the big net gain of $15,000. But there’s also a 50 percent chance that you will be out $1,000. The gamble might seem too risky for you.

Coin tosses are special because it is relatively easy to determine the probability of a head or a tail. This is not the case for all the types of uncertainty you might face. In some cases, financial instruments—such as the mortgage-backed securities that played a big role in the financial crisis of 2007–2009—are so complex that investors find it difficult to assess the probabilities of various outcomes.

We often do a bad job of *estimating* probabilities. One reason for this is because we are unduly influenced by things that we can easily bring to mind. Psychologists call this the “availability heuristic.” For example, we tend to overestimate certain causes of death, such as car accidents, tornadoes, and homicides, and underestimate others, such as diabetes, stroke, and asthma.See Paul Slovic, Baruch Fischoff, and Sarah Lichtenstein, “Facts versus Fears: Understanding Perceived Risk,” in *Judgment under Uncertainty: Heuristics and Biases*, ed. Daniel Kahneman, Paul Slovic, and Amos Tversky (Cambridge, MA: Cambridge University Press, 1982), 463–89. We also often do a poor job at *using* probabilities; in particular, we often put too much emphasis on small probabilities. For example, consider two drugs that are equally effective in treating a disease, but suppose the older drug has a 1 in 10 million chance of having a certain side effect and the newer drug has a 1 in a 100 million chance of having the same side effect. Consumers might view the new drug as much more appealing, even though the side effect was already highly improbable with the older drug.

Diversification of Risk

In many cases, we would like to find some way of getting rid of—at least to some degree—the risks that we face. One way we eliminate risk is through insurance. Sometimes we purchase insurance on the market. Sometimes our employer provides us with insurance. Sometimes the government provides us with insurance. In the following subsections, we look at many different kinds of insurance, including property insurance, unemployment insurance, and deposit insurance.We do not discuss health insurance here. In Chapter 16, we discuss the provision of health care and the problems of health insurance in detail.

First, though, we need to understand how and why insurance works. Suppose you have a bicycle worth $1,000, and (for some reason) you cannot purchase insurance. You think that, in any given year, there is about a 1 percent chance that your bike will have to be replaced (because it is either stolen or written off in an accident).

Now, in expected value terms, this may not look too bad. Your expected loss from an accident is $0.01 × $1,000 = $10. So on average, you can expect to lose $10 a year. But the problem is that, if you are unlucky, you are stuck with a very big expense. Most of us dislike this kind of risk.

You are complaining about this to a friend, and she sympathizes, saying that she faces exactly (and we mean exactly) the same problem. She also has a bike worth $1,000 and thinks there is a 1 percent chance each year that she will need to replace it. And that's when you have the brilliant idea. *You can make an agreement that, if either one of you has to replace your bikes, you will share the costs*. So if you have to replace your bike, she will pay $500 of your costs, and if she has to replace her bike, then you will pay $500 of her costs. It is (almost exactly) twice as likely that you will have to pay *something*, but if you do, you will only have to pay half as much. With this scheme, your expected loss is unchanged. But you and your friend prefer this scheme because it is less risky; it is much less likely that you will have to make the big $1,000 payout.

We are implicitly assuming here that your chances of having to replace your bike are independent of the chance that she will have to replace her bike. (If you are likely to crash into her, or both of your bikes are stolen, then it is a different story.) There is also still a chance that you will both experience the unlucky 1-in-100 chance, in which case you would both still have to pay $1,000. But the likelihood of this happening is now tiny. (To be precise, the probability of both of you having an accident in the same year is 1 in 10,000 [that is, 0.0001]). This is because the probability that two independent events occur equals the probability of one multiplied by the probability of the other.)

But why stop here? If you can find two more friends with the same problem, then you can make it almost certain that you will have to pay out no more than $250. It is true that you would be even more likely to have to make a payment because you will have to pay if you or one of your friends has to replace his or her bike. But because the payment is now being shared four ways, you will have to pay only 25 percent of the expenses. This is an example of **diversification**, which is the insight that underlies insurance: people share their risks, so it is less likely that any single individual will face a large loss.

Diversification and insurance don't prevent bad stuff from happening. We live in a world where bicycles are stolen; where houses are destroyed by floods, fires, or storms; where people have accidents or become ill; and so on. There is not a lot we can do about the fact that bad things happen. But we can make the consequences of these bad things easier to deal with. Insurance is a means of sharing—diversifying—these risks.

Continuing with our bicycle insurance example, suppose you could find thousands of friends who would agree to be part of this arrangement. As more and more people join the scheme, it becomes increasingly likely that you have to make a payment each year, but the amount you would have to pay becomes smaller and smaller. With a very large number of people, you would end up very close to a situation where you pay out $10 with certainty each year. Of course, organizing thousands of your friends into such a scheme would present all sorts of practical problems. This is where insurance companies come in.

Insurance companies charge you a premium (an annual payment). In return, they promise to pay you an **indemnity** in the event you suffer a loss. The indemnity is usually not the full amount of the loss. The part of the loss that is not covered is called the **deductible**. In our example, there is no deductible, and the indemnity is $1,000. An insurance company would charge you a premium equal to the expected loss of $10 plus a little extra. The extra payment is how the insurance company makes money. You and everyone else are willing to pay this extra amount in return for the removal of risk.

The idea of diversification can also be applied to investment.We discuss this in Chapter 10. Think back to our example of your friend with the Internet venture. You might not want to invest $1,000 in his scheme because it seems too risky. But if you had 100 friends with 100 similar (but independent) schemes, you might be willing to invest $10 in each. Again, you would be diversifying your risk.

Risk Aversion

The preceding discussion of insurance and diversification is based on the presumption that people typically wish to avoid risk whenever possible. In our example, you have a 1 percent chance of suffering a $1,000 loss. Your expected loss is therefore $10. Now imagine we give you a choice between this gamble and a certain loss of $10. If you are just as happy in either case, then we say you are **risk-neutral**. But if you are like most people, then you would prefer a certain loss of $10 to the gamble whereby you have a 1 percent chance of losing $1,000. In that case, you are **risk-averse**.

Toolkit: Section 31.7

Suppose you are presented with the following gamble:

- Lose nothing (99 percent probability)
- Lose $1,000 (1 percent probability)

How much would you pay to avoid this gamble? If you are risk-neutral, you would be willing to pay only $10, which is the expected loss. If you are risk-averse, you would be willing to pay more than $10. The more risk-averse you are, the more you would be willing to pay.

It is risk aversion that allows insurance companies to make money. Risk-averse people prefer a sure thing to a gamble that has the same expected value. In fact, they will prefer the sure thing to a gamble with a slightly lower expected value. Because it can diversify risk, the insurance company cares only about the expected value. Thus an insurance company behaves as if it were risk-neutral.

Different Kinds of Insurance

There are many different kinds of insurance available to you. We briefly discuss some of the most important.

Property Insurance

Many forms of property are insured: houses, cars, boats, the contents of your apartment, and so on. Indeed, some insurance is often mandatory. People purchase insurance because there are risks associated with owning property. Houses burn down, cars are stolen, and boats are wrecked in storms. In an abstract sense, these risks are just like a coin flip: heads means nothing happens; tails means there is a fire, a robbery, or a storm.

Let us look at home insurance in more detail. Suppose you own a house that is worth $120,000. You might pay $1,000 per year as a premium for an insurance policy. If your house burns down, then the insurance company will pay you some money to recover *part* of the loss. If the deductible on the policy is $20,000, you would receive an indemnity of $100,000. You lose $20,000 when the house burns down because the insurance company does not fully cover your loss.

Thus, if your house burns down, the insurance company loses the indemnity minus the premium—a total of $99,000. You lose the deductible and the premium—a total of $21,000. Your joint loss is $120,000—the lost value of the house. This serves to remind us again that insurance is not some magic way of preventing bad things from happening. When the house does not burn down, the insurance company earns the $1,000 premium, and you pay the $1,000 premium. Your joint loss is zero in this case.

You may wonder why insurance companies typically insist on a deductible as part of an insurance contract. After all, you would probably prefer to be covered for the entire loss. Deductibles exist because insurance policies can have the effect of altering how people behave. We have assumed that the probability of a bad thing happening was completely random. But if you are fully insured, you might not be so careful about how you look after your house. You might worry less about turning off the stove, ensuring that you have put out the fire in the fireplace, falling asleep while smoking, and so on. Deductibles make sure that you still have a big incentive to take care of your property.

Unemployment Insurance

Not everyone who wants to work actually has a job. Some people are unemployed, meaning that they are actively looking for work but do not have jobs. The **unemployment rate** is the number of unemployed individuals divided by the sum of the number employed and the number unemployed.

Toolkit: Section 31.3

If you want to learn more details about the definition and measurement of unemployment, refer to the toolkit.

Since 1960, the unemployment rate in the United States has averaged slightly under 6 percent. This means that for every 100 people in the labor force (either working or looking for a job), 94 of them are working, and the other 6 are looking for jobs. The **labor market** is fluid so that, over time, unemployed workers find jobs, while some employed workers lose jobs and become unemployed. The unemployed find jobs, and others lose them and go through spells of unemployment.

If you are laid off from your job and become unemployed, you obviously still need to spend money for food and rent. During a spell of unemployment, you have several possible sources of income. If you have an existing stock of accumulated savings, then you can draw on these. If you are a member of a union, you may receive some support from the union. You may receive some severance pay when you lose your job. You might be able to rely on the support of your family and friends. And, most relevant for this chapter, you may be eligible to receive income from the government, called **unemployment insurance**.

Unemployment insurance is similar in some ways to health and property insurance. There is an unlucky event called unemployment, and the government provides insurance. Perhaps you think this is great news: after graduation, you can claim unemployment, collect from the government, and enjoy your leisure. Of course, life is not quite that good. First, to qualify for unemployment insurance, you have to hold a job for some period of time. The details of these regulations differ across countries and also across states within the United States. Second, unemployment benefits do not last forever, nor do they completely compensate for all of your lost income. Again, the details depend on the country or state in which you work.

Why is the government in the business of providing insurance? To answer this, look back at our example of home insurance. The typical insurance company will have many policies with many different households. Over the course of a year, some households will make a claim on their insurance, but most will not. As long as the insurance company has lots of policies in many locations, then, on average, the number of insurance claims will be nearly constant each year. Although individual households face risk, the insurance company is able to diversify almost all of this risk.

Unemployment is different. When the economy is doing well, unemployment is low, and few households need this form of insurance. When the economy is not doing well, then the unemployment rate can be very high. In such times, many people want to claim unemployment insurance *at the same time*. So unlike insurance policies for homeowners, there is no easy way to balance out the risks of unemployment. The risk of unemployment is not *independent* across all individuals.

If an individual insurance company tried to offer unemployment insurance, it might be unable to survive: during a period of low economic activity, the demands for insurance would be so severe that the insurance company might not be able to meet all the claims. The government has the ability to tax people and borrow as needed. This puts it in a much better position to offer unemployment insurance. So in many countries, the government raises revenue by taxing firms and workers and uses these funds to provide unemployment insurance.

Deposit Insurance

In the United States and in some other countries, deposits that you place in the bank are insured by the government. In the United States, the government provides insurance, up to $250,000 per deposit, to you in the event your bank closes.You can find details at FDIC, “Your Insured Deposits,” accessed March 14, 2011, www.fdic.gov/deposit/deposits/insured/basics.html. Deposit insurance in the United States dates from the time of the Great Depression in the 1930s. In this period many banks had insufficient funds on hand to meet the demands of their depositors and so went bankrupt. When this occurred, depositors lost the money they had put in the bank. After the Great Depression, the US federal government instituted deposit insurance. Similar programs exist in most other countries.

The argument for why the government should provide deposit insurance is similar to the argument for government provision of unemployment insurance. During periods of financial turbulence, many banks are prone to failure. If there were a private insurance company providing deposit insurance, it would probably be unable to meet all the claims. In addition, there is considerable social value to deposit insurance. It gives people greater confidence in the bank and in the banking system, which in turn makes bank failures less likely. Because bank failures put a great deal of stress on the financial system, government has an interest in insuring deposits.

In the summer of 2007, the British bank Northern Rock entered a financial crisis. Savers who had put their money in this institution started to worry that the bank would go bust, in which case they would lose their money. The British government, like the US government, provides deposit insurance. However, the amount of this insurance was limited to a maximum of about $70,000, so some people were still concerned about their savings. As lines started to form outside Northern Rock branches, the British government—concerned that the possible failure of Northern Rock would put other banks at risk—ended up guaranteeing all of its deposits.

Key Takeaways

- The probability of a particular outcome is the percentage chance that the outcome will occur.
- An expected value is calculated by multiplying the probability of each outcome by the value of that outcome and then adding these numbers for all outcomes.
- Risk aversion means a preference for a sure thing rather than a gamble with the same expected value.
- We face many types of risks in our lives, and we can often buy insurance as a way to deal with these risks.
- Insurance companies provide a way for individuals to diversify their individual risks.

check your understanding

1. Imagine you have a die that is fair: the probability of rolling each number is ⅙. Each time you roll the die, it is independent of the previous roll. Suppose you roll the die 10 times, and you roll a 5 each time. What is the probability of getting a 5 on the next roll of the die? Relate your answer to the common analysis of sportscasters who say that a baseball hitter who normally bats 0.300 “is due” when that batter has no hits in his last 20 at bats.
2. List some risks that you face that are not fully covered by insurance.

This page titled 5.4: Avoiding Risk is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

5.5: Embracing Risk 5.5: Embracing RiskEdit Update titlecancel

Learning Objectives

1. Why do people sometimes take on risks instead of avoiding them?
2. What are compensating wage differentials?

Insurance allows us to remove some risks from our lives, either partially or completely. There are other risks that we accept or even embrace. In some cases, we are willing to take on risks because we are compensated in some way for them. We discuss two examples here: job choice and investment in the stock market. In other cases, we actively seek risk, such as when we gamble, buy lottery tickets, or engage in extreme sports.

Job Choice

We have already discussed how your choice of career should be based on discounted present value, taking into account the fact that it can be costly to move from one job to another. Your choice of career also reflects how you feel about risk and uncertainty. There are numerous kinds of jobs that differ in many dimensions. When choosing a job, we pay particular attention to wages or salaries, vacation plans, health coverage, and other benefits. We also take into account the job's riskiness.

The most severe risk is, of course, the risk of death. Every year in the United States several thousand workers suffer fatal work injuries. For most of the last decade, this number has been between about 5,000 and 6,000, or the equivalent of roughly 4 deaths for every 100,000 workers. Figure 5.5.1 shows data on work-related deaths in the United States for 2006–9.The data come from Occupational Safety and Health Administration, "Census of Fatal Occupational Injuries," accessed February 24, 2011, http://stats.bls.gov/iif/oshwc/cfoi/cfch0008.pdf. The Occupational Safety and Health Administration (OSHA) is the federal agency that monitors workplace safety. The fatality rate has generally been declining over time. In the early 1990s, the fatality rate was in excess of 5 deaths for every 100,000 workers; in 2009, the fatality rate was 3.3.

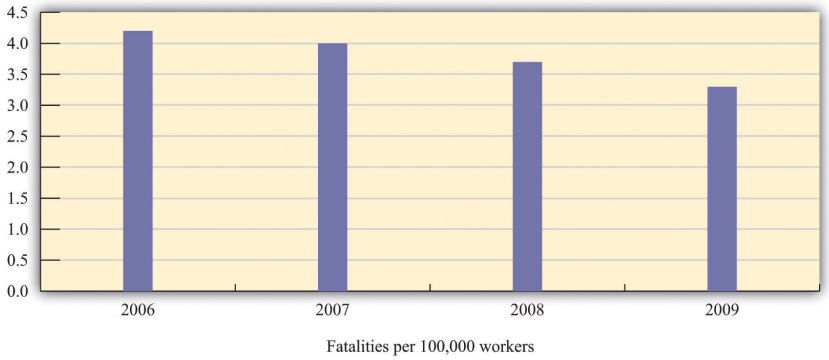


Figure 5.5.1: Work Fatalities in the United States

Source: US Department of Labor, Occupational Safety and Health Administration, http://stats.bls.gov/iif/oshwc/cfoi/cfch0008.pdf.

The fatality rate in the United States is high relative to many other developed countries. For example, the corresponding rate in England in 2006 was 0.8 per 100,000 workers.See Health and Safety Executive, "Fatal Injury Statistics," accessed March 14, 2011, http://www.hse.gov.uk/statistics/fatals.htm. Care must be taken when comparing international data because they are not always exactly comparable. Most importantly, European Union statistics exclude road traffic accidents, which account for about 43 percent of US deaths. Correcting for this, the US figures are similar to the European Union average. Figure 5.5.2 shows the work-related fatalities for countries in the European Union. Most importantly for the current discussion, the fatality rate varies significantly across jobs. In the United States, "Agriculture and mining recorded the highest fatal work injury rates among the major industry sectors in 2005—32.5 fatalities per 100,000 workers for agriculture and 25.6 fatalities per 100,000 workers for mining." In the European Union, construction, agriculture, and transportation are the most dangerous sectors.See US Department of Labor Bureau of Labor Statistics, "National Census of Fatal Occupational Injuries in 2005," accessed March 14, 2011, http://www.bls.gov/iif/oshwc/cfoi/cfnr0012.txt and Health and Safety Executive, "European Comparisons," accessed March 14, 2011, http://www.hse.gov.uk/statistics/european/fatal.htm.

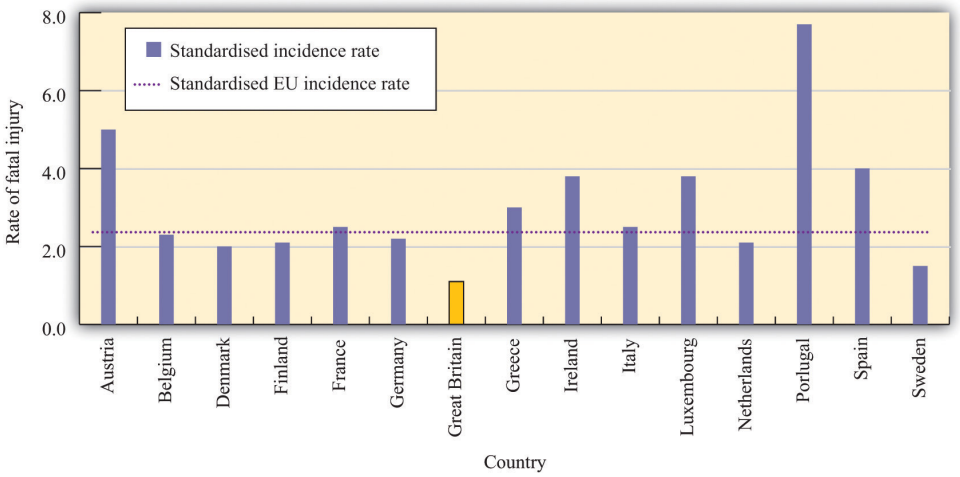


Figure 5.5.2: Work Fatalities in Europe

In the European Union, the average work-related fatality rate is lower than in the United States, although some countries—most notably Portugal—have higher rates. The lowest rate in the European Union is in Great Britain.

Source: Health and Safety Executive, Great Britain: http://www.hse.gov.uk/statistics/european/fatal.htm.

The risk of death is not the only job risk. Jobs also differ in terms of their risk of injury. In some cases, these injuries can have a severe impact on a person's quality of life; in other cases, they may prevent an individual from working in the future. Professional athletes and other performers face the risk of injuries that can end their careers. These individuals sometimes buy insurance to help mitigate some of these risks. In 2006, for example, pop star Mariah Carey purchased a $1 billion insurance policy on her legs after signing up for an advertising campaign ("Legs of a Goddess") with the Gillette shaving company. Bruce Springsteen's voice is insured for $5.6 million.See starpulse.com, "Mariah Carey Takes Out $1 Billion Insurance Policy For Her Legs," accessed March 14, 2011, www.starpulse.com/news/index.php/2006/05/30/mariah_carey_takes_out_1_billion_insuran for an account of Ms. Carey's insurance policy; Paul Bannister, "World's Biggest Insurer Takes on All Risks," accessed March 14, 2011, http://www.bankrate.com/brm/news/insurance/old-lloyds1.asp, details a number of unusual insurance policies.

Riskier jobs generally pay more. A firm that exposes workers to more risk must compensate them for that risk. A **compensating wage differential** is any difference in pay received by identical workers doing different jobs. (By "identical" we mean workers with comparable education, skills, experience, and so on.) Jobs that are unpleasant or dangerous will typically pay higher wages to compensate workers for the negative aspects of their jobs.

To the extent workers do not like to face risks, jobs that are viewed as riskier tend to pay more on average. For example, a recent study found that nurses who were more likely to be exposed to the AIDS virus (HIV) received higher wages than comparable nurses who were less likely to be exposed.Jeff DeSimone and Edward J. Schumacher, "Compensating Wage Differentials and AIDS Risk" (NBER Working Paper No. 10861, November 2004), accessed March 14, 2011, http://www.nber.org/papers/w10861.

Asset Portfolio Choice

Young people's portfolios of assets are usually very simple: a typical college student might have only a checking account and a savings account. As you grow older, though, you will typically acquire a broader portfolio. Even if you do not directly purchase assets such as stocks and bonds, you may own them indirectly when you sign up for a pension plan. Because the return on stocks (and other assets) is uncertain, owning these assets is another type of risk you *choose* to take.

Owning a stock is somewhat like buying a lottery ticket. You pay some money to buy a share of the stock of some company. In return, you may be paid some dividends; at some time in the future, you may sell the stock. But at the time you buy the stock, you don't know the payments you will receive in the future and you don't know the future price of the stock. So by purchasing a stock, you are gambling. Whether the gamble is favorable or not depends on the price of the stock, the chance it will pay dividends in the future, and the future price. Choosing how to allocate the assets in a portfolio is a type of gamble we all make.

We cannot completely avoid this kind of gamble. Perhaps you think that putting cash in your mattress would be a way to avoid this risk, but that is not the case. Ultimately you care about what that money can buy in terms of goods and services. The real value of

the money held in your mattress depends on the future prices of goods and services, which are not known to you today. The benefit of holding cash depends on the unknown inflation rate.

Buying a Lottery Ticket

Some people, at some times, are eager and willing to engage in risky activities. People engage in extreme sports, where the danger appears to be part of the attraction. People go to Las Vegas or to Monaco to gamble. In many countries, citizens can purchase a wide variety of lottery tickets sold by their governments. This is a form of gambling: you buy a ticket and if you have the lucky number, then you get a (sometimes large) prize. If you do not have the lucky number, your money is gone.

The existence of lotteries and other kinds of gambling seems like a puzzle. If people are risk-averse, then they are supposed to want to get rid of risk. The purchase of a lottery ticket is the exact opposite: you give up a sure thing (the price of the ticket) for an uncertain outcome. Unlike the purchase of insurance, which is a way to avoid risk, buying a lottery ticket is a demand for a gamble. Why do so many people buy lottery tickets? Why do governments sell them?

Do lottery tickets have an expected value that exceeds the cost of the ticket? If the difference is big enough, then even a risk-averse person might want to buy a ticket. Consider a very simply lottery. Suppose there is one fixed prize, and there is a probability that you win the prize. Then your expected gain is just the probability of winning times the prize:

$$expected\ gain = probability\ of\ winning \times value\ of\ prize.$$

Using this equation, you can determine whether the price for a lottery ticket is high or low. If the price of the ticket exceeds the expected gain from buying the ticket, then the ticket is not a good deal. But if the price is low relative to the expected gain, then you may want to accept the risk and buy the ticket.

Let us look at an example. One US lottery is called Powerball. On February 18, 2006, the prize was worth $365 million to the winner. The chance of getting the jackpot was 1 in 988,172,368.This is based on five balls selected ranging from 1 to 55 and a powerball ranging 1 to 42. See the Powerball site (http://www.usamega.com/powerball-howtoplay.htm) for details on this game. See also the odds calculator at CSGNetwork.com, “The Ultimate Lottery Games Odds Calculator,” accessed March 14, 2011, http://www.csgnetwork.com/oddscalc.html. The expected value of a ticket at that time was the value of the prize times the probability of winning: far less than the dollar price of the ticket. Despite the huge prize, the price of a ticket far exceeded its expected value.

Another perspective on the lottery is from the viewpoint of the government selling these tickets. Consider the Texas lottery.Texas Lottery home page, accessed March 14, 2011, http://www.txlottery.org/export/sites/default/index.html. The proceeds from the sale of tickets primarily support education. In 2005, about 60 percent of the income from the lottery went to payment for prizes and 28 percent went to a school fund. From the perspective of the Texas state government, selling lottery tickets is a way to fund programs. If the government is to make money on lottery tickets, then those buying the tickets must, on average, be losing money.

In fact, as this discussion suggests, the expected value of a lottery ticket is less—often substantially less—than the cost of the ticket. Why, then, do people buy lottery tickets? One possibility is that they simply enjoy gambling. This means that, at least with respect to these types of gambles, they are *risk-loving* rather than risk-averse. The pleasure of a lottery ticket is, among other things, the license to dream. Another possibility is that individuals overestimate their chances of winning.

Key Takeaways

- One reason that people take on risks is because they enjoy gambling.
- Some jobs are riskier than others and pay more to compensate people for the risks they face.

check your understanding

1. Suppose you take a job as an engineer at an oil company. There are two places where you can work. One is Houston, where you will live in the suburbs and commute to work. The other is on an oil rig in the North Sea (between Scotland and Norway), where you will spend most of your time on the rig. If the jobs are otherwise identical, which do you expect will pay more? Explain why.

This page titled 5.5: Embracing Risk is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

5.6: End-of-Chapter Material 5.6: End-of-Chapter MaterialEdit Update titlecancel

In Conclusion

Many decisions involve a trade-off between now and the future. Whenever we invest our time or money, we are giving up something today to obtain something in the future. So by saving some of our income, we give up consumption now for consumption in the future. When we go to a university, we give up income and leisure time today to get more consumption (through higher income) in the future.

Once you start thinking about trade-offs over time, it is difficult to avoid the reality that many of these decisions are made in the face of great uncertainty. When we save, we are not certain of the return on our saving. When we go to school, we are not guaranteed a job in the future nor are we guaranteed a specific salary. We have provided some insights into the nature of these uncertainties and how to deal with them. Discounted present value and expected value are techniques that are worthwhile to master, as they will help you make better decisions throughout your life.

Key Links

- Unemployment insurance
  - New York State: www.labor.state.ny.us/workerprotection/laborstandards/workprot/ui.shtm
  - Texas: http://www.twc.state.tx.us/ui/bnfts/claimant1.html
- Deposit insurance: http://www.fdic.gov/deposit/index.html
- Powerball lottery: http://www.powerball.com
- Fun site to calculate probabilities: http://stattrek.com/Tools/ProbabilityCalculator.aspx

exercises

1. Explain why an increase in the interest rate reduces the discounted present value of income.
2. What incentives exist for people to repay loans on education? On cars?
3. Suppose the nominal interest rate is 20 percent, the price level in the first year is 50, and the price level in the second year is 60. What is the real interest rate? How could you alter this example so that the real interest rate is 0?
4. Can the nominal interest rate ever be less than the real interest rate? If the real interest rate is negative, what happens to the slope of the budget line with two periods discussed in Chapter 5, Section 5.1?
5. Do households sometimes borrow and lend simultaneously? Why might that happen? Is the interest rate they borrow at usually higher or lower than the interest rate they receive, say in the form of bank deposits?
6. In describing how changes in income influence the supply of loans, we assumed that the increase in income occurs this year. Suppose instead that the increase in income will occur next year even though everyone in the economy knew it would happen today. How would the news of a future increase in income influence the current loan supply curve?
7. When the government changes taxes, do you know if it is permanent or temporary?
8. (Advanced) One way that real wages received by workers can change is through a change in income taxes. Considering the information in this chapter, would you except temporary tax changes to have a bigger or a smaller impact on labor supply than a permanent tax change? What if the tax change is not through a change in the tax rate but rather through a fixed payment to households? What would that policy do to labor supply?
9. Look back at Table 5.3.1 "Which Career Should You Choose?" in the section “Choosing a Career.” How would you edit the income entries in the table so that the insurance salesperson had a higher discounted present value than the lawyer even when the interest rate is 5 percent?
10. Look back at Table 5.3.1 "Which Career Should You Choose?" in the section “Choosing a Career.” Explain why an increase in the interest rate makes it less attractive to be a lawyer.
11. Besides discounted present values of income, what other factors are important in choosing a career? How do you balance these with differences in the discounted present value of income?
12. What are the risks associated with choosing a particular career? How do those risks depend on whether the skills you learn at your job can be used in other jobs?
13. Show the calculation of the discounted present value from work in Table 5.3.5 "Income from Going to College versus Taking a Job". Redo the comparison of college and work assuming an interest rate of 20 percent.

14. (Advanced) Look back at Table 5.3.1 "Which Career Should You Choose?" in the section “Choosing a Career.” Suppose the income of a lawyer increases by 20 percent each year after year 2 and the income of the insurance salesperson increases by $10,000 each year. Extend Table 5.3.1 "Which Career Should You Choose?" to 5 years. Which is a better profession?
15. Create your own version of Table 5.4.1 "Coin-Flipping Experiment" by flipping a coin 10 times. Imagine that each time the result is a head, you earn $1,000, and each time it is a tail, you lose $1,000. After you flip the coin 10 times, calculate how much you won (or lost). Now do this same experiment 20 times. Each time you flipped the coin 10 times, record how much you won (or lost), which will result in 20 numbers. What is the average of these 20 numbers? What is the expected value of how much money you will earn in each coin-flipping experiment??
16. Many products, such as computers, come with the option of an *extended warranty*. Suppose you are buying a computer with a one-year warranty. Thereafter, you can purchase an extended warranty for one more year, at a cost of $50. The warranty will repair or replace your computer in the event of breakdown. Suppose the average cost to the manufacturer of repairing or replacing the computer is $1,000. If the manufacturer is making no money from this warranty, what is the implied probability that the computer will need repair?
17. We wrote “As long as the insurance company has lots of policies in many locations, then, on average, the number of insurance claims will be nearly constant each year.” Why did we include the statement about many locations?
18. Why is it difficult to diversify job risk? Is it possible to do some diversification within a family?
19. In the United States, the provision of unemployment insurance is partly at the state level and partly at the federal level. For your state, find out what the benefits are and what federally funded unemployment insurance might be available to you.

Economics Detective

1. Study the insurance policy you can buy when you purchase a new cell phone. Exactly what does this insurance protect you against? Given the price of the insurance and the coverage, what is the implied probability that you will make a claim for a new phone under the insurance? Is there a deductible? Why is it part of the policy?
2. Look at the insurance policy (if you have one) for the place where you are living. What is the deductible? List the ways in which you take actions to reduce the risk of fire where you live.
3. Our example of homeowners’ insurance did not use real numbers. Find a homeowners’ policy and determine the coverage, the premium, and the deductible.
4. One form of insurance occurs when you rent a car. Using the Internet or phoning local insurance agents, find out the kinds of insurance that are available when you rent a car. What is the cost per day? Exactly what risks do these policies protect you against? Given the price of the insurance and the coverage, what is the implied probability that you will have an accident and make a claim under the insurance? Does this probability seem reasonable to you?
5. What is the average price of a house in the United States? In your hometown?
6. What does it cost to insure a $100,000 house in your city? What does it cost to insure a $1,000,000 house in your city? Explain the differences in the insurance costs.
7. Pick a state in the Unites States. Suppose you work there and earn $2,000 each week as a manager. One day, the firm tells you that you are no longer needed. What unemployment insurance could you collect? Would you qualify for unemployment insurance? How much would the benefits be? How long would the benefits last?
8. Go to your local bank and see if there are any signs that indicate deposit insurance is provided. Ask about details of the program.
9. Use the Internet to find out about deposit insurance programs in the United States and in another country. How do these programs compare?
10. For the state in which you live, does the government sponsor a lottery? If so, how are the funds used?
11. In the financial crisis of 2008–9, was deposit insurance provided in the United States? In other countries?
12. In the state you live in, find out about the unemployment insurance program. How long do you receive benefits and how generous are the benefits?

Spreadsheet Exercises

1. Write a spreadsheet program to create a version of Table 5.4.1 "Coin-Flipping Experiment" for any combination of income flows and interest rates.
2. (Advanced) Create a spreadsheet program to simulate the flipping of a coin. Do T experiments with 5 flips per experiment. For each experiment, calculate the mean of the outcome. When you are finished, you will have T means. What does the distribution of the T means look like? What is the mean of that distribution? What happens as T gets very large?

This page titled 5.6: End-of-Chapter Material is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

6: eBay and craigslist

6.1: Buying on the Internet
6.2: craigslist and the Gains from Trade
6.3: eBay
6.4: Supply and Demand
6.5: Production Possibilities
6.6: End-of-Chapter Material

This page titled 6: eBay and craigslist is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

6.1: Buying on the Internet 6.1: Buying on the InternetEdit Update titlecancel

eBay is one of the most famous sites on the Internet (www.ebay.com; Figure 6.1.1). It was founded in 1995 and is now a very large company, with $60 billion in sales in 2009 and over 90 million active users worldwide.See eBay Inc., “Who We Are,” accessed February 24, 2011, http://www.ebayinc.com/who. One of the many ways in which the Internet is changing the world is that there are now ways of buying and selling—such as eBay—that were completely unavailable to people 20 years ago.



Figure 6.1.1: The eBay Home Page

In or around the second and third centuries BCE, the island of Delos, in Greece, was a major center of trade—both of goods and slaves. At its height of activity, Delos, an island of five square kilometers, had a population of about 25,000 people.MyKonos Web, “History of Delos,” accessed January 25, 2011, http://www.mykonos-web.com/mykonos/delos_history.htm. This means Delos was about as densely populated as modern-day cities such as Istanbul, London, Chicago, Rio de Janeiro, or Vancouver. Visitors today see different things when they visit the ruins of Delos. Some see dusty pieces of shattered rock; others see the remains of a great culture. An economist sees the ruins of a trading center: a place where people such as the trader shown in Figure 6.1.2 were, in a sense, the eBay of their times.



Figure 6.1.2: A Trader in Delos

Delos and eBay are separated by almost two and a half millennia of history, yet both are founded on a basic human activity: the trading of goods and services. How basic? Consider the following.

- Children learn to trade at an early age. First graders may be trading Pokémon cards before they have even learned the meaning of money.

- In prisoner-of-war camps during World War II, prisoners would receive occasional parcels from the Red Cross, containing items such as chocolate bars, cigarettes, jam, razor blades, writing paper, and so on. There was extensive trading of these items. In some cases, cigarettes even started to play the role of money. For the classic discussion of trade in a prisoner-of-war camp, see R. A. Radford, "The Economic Organisation of a P.O.W. Camp," *Economica* (*New Series)* 12, no. 48 (1945): 189–201.
- In Nazi concentration camps, where people lived close to starvation in conditions of extreme danger and deprivation, the prisoners traded with each other. They traded scraps of bread, undergarments, spoons, basic medicines, and even tailoring services.Jill Klein, "Calories for Dignity: Fashion in the Nazi Concentration Camp," *Advances in Consumer Research* 15 (2003): 34–37.

Trade has played a central role in determining where many of us live today.

- In England, you will find the town of Market Harborough; in Germany, you can find Markt Isen; in Sweden, Lidköping. *Markt* and *köping* both mean market. These towns, and many others like them, date from the medieval period and, as their names suggest, owe their existence to the markets that were established there.
- Many of the great cities of the world developed in large part as ports, where goods were imported and exported. London, New Orleans, Hong Kong, Cape Town, Singapore, Amsterdam, and Montreal are a few examples.

Much of economics is about how we interact with each other. We are not alone in the economic world. We buy goods and services from firms, retailers, and each other. We likewise sell goods and services, most notably our labor time. In this chapter, we investigate different kinds of economic interactions and answer two of the most fundamental questions of economics:

1. How do we trade?
2. Why do we trade?

Road Map

The chapter falls naturally into two parts corresponding to these two questions. We begin by thinking about the ways in which individuals exchange goods and services.

In modern economies, most trade is highly *disintermediated*. You usually don't buy a good from its producer. Perhaps the producer sells the good to a retail store that then sells to you. Or perhaps the good is first sold to a wholesaler who then sells to a retailer who then sells to you. Goods are often bought and sold many times before you get the opportunity to buy them.Such transactions are the focus in Chapter 7 and Chapter 8. For the moment, however, we have a different emphasis. We do not yet get into the details of retailing in the economy but instead focus on trade among individuals—the kind of transaction that you can carry out on eBay and craigslist.

Specifically, we want to understand how potential buyers and sellers are matched up. We also want to know what determines the prices at which people exchange goods and services. Broadly speaking, prices can be established in the following ways.

- Some prices are the result of bargaining and negotiation. If you buy a car or a house, you will engage in a one-on-one negotiation with the seller in which there will typically be several rounds of offers and counteroffers before a final price is agreed on. Similarly, if you go to street markets in many countries in the world, you will not find posted prices but will have to bargain and haggle.
- Some prices are determined by auction, such as trinkets sold on eBay and antiques sold by Christie's. Auctions are a type of bargaining that must follow some preset rules.
- Some prices are chosen by the seller: she simply displays a price at which she is willing to sell a unit of the good or service. Even this is a very simple type of bargaining called the "take-it-or-leave-it offer." The seller posts a price (an "offer"), and prospective buyers then have a simple choice: either they buy at that price (they "take it") or they do not buy (they "leave it").

Take-it-or-leave-it offers are the most common form of price-setting in retail markets. The prices displayed in your local supermarket can be thought of as thousands of take-it-or-leave-it offers that the supermarket makes to you and other shoppers. Whenever you go to the supermarket, you reject most of these offers (meaning you don't buy most of the goods on display), but you accept some of them. Take-it-or-leave-it offers also occur when individuals trade. Classified advertisements in newspapers or on Internet sites like craigslist typically involve take-it-or-leave-it offers.

Once we understand how individuals trade with one another, we turn to an even more basic question: why do we trade? Whether we are talking about first graders swapping Pokémon cards, the purchase of a camera on eBay, the auction of a Renoir

painting at Sotheby's, or traders in the Mediterranean islands over two millennia ago, there is one reason for trade: I have something you want, and you have something I want. (In many cases, one of these "somethings" is money. Keep in mind, though, that people don't want money for its own sake; they want money to buy goods and services.)

We therefore explain how differences in what we have and what we want provide a motive for trade and how such trade creates value in the economy. Then we go deeper. In a modern economy, trade is an essential part of life. We consume a large number of goods and services, but we play a role in the production of very few. Put differently, modern economies exhibit a great deal of specialization. We carefully investigate how specialization lies right at the heart of the gains from trade.

This page titled 6.1: Buying on the Internet is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

6.2: craigslist and the Gains from Trade 6.2: craigslist and the Gains from TradeEdit Update titlecancel

Learning Objectives

1. What are the gains from trade?
2. How does a market with take-it-or-leave-it offers work?
3. How are the gains from trade split between buyers and sellers?
4. What is economic efficiency?

To begin our investigation of why and how we trade, let us examine craigslist (http://www.craigslist.org), an Internet site devoted to exchange. The craigslist site is very similar to the classified advertisements in a newspaper except that the advertisements are online. It is local, in the sense that there is a different site for different places. You can find craigslist sites for cities and states throughout the United States, and—at the time of this writing—for 14 cities and 54 countries around the world. If you visit the craigslist website, you will see there are many types of goods and services listed. For now, we focus on the purchase of a good. Later, we will consider the purchase and even the exchange of services.

Pricing on craigslist commonly takes a take-it-or-leave-it form. The seller posts a price and then buyers and sellers communicate through (anonymous) e-mails. Of course, the buyer always has the option of trying to turn this take-it-or-leave-it scenario into back-and-forth bargaining by making a counteroffer. Once they have agreed to trade, the buyer and seller must find a way to consummate the transaction—delivering the good and making payment.

The Gains from Trade from a Single Transaction

Suppose you are interested in buying a car. You go to craigslist in your area and search through offers to sell cars. These offers typically provide lots of information about the product, usually including photos and a price. If you want to inquire about a particular car, you can contact the seller. If you want to buy the car, you can accept the seller's offer. If you want to negotiate, you can do so as well. To get at the heart of this kind of exchange, let us first take a simple case where there is a single seller and a single buyer.

Economists generally think that individuals make decisions in their own self-interests. If a seller is willing to sell a good at a given price, and a buyer is willing to buy at that price, our presumption is that this exchange makes them both better off. This deceptively simple idea is the very heart of economics: *voluntary trade makes both participants better off*. The word *voluntary* matters here. We are supposing that both people freely enter into this trade. If two people make a deal of their own free will and if they are rational, in the sense that they can make decisions in their own best interests, then the deal must make them both better off.

The demand for a car is an example of a **unit demand curve** because you are deciding whether or not to buy at all rather than how much you should buy.See Chapter 4 for more discussion. The buyer has a **valuation** for the good, which represents the most he would be willing to pay for it. For example, suppose you see a used car on craigslist, and your valuation of this car is $3,000. This means that you would be equally happy either having the car and forfeiting $3,000 worth of other goods and services or not having the car. Figure 6.2.1 shows what your demand curve looks like in this case. You are choosing to buy either zero units or one unit, so if the price is above your valuation, you do not buy the good, whereas if the price is below your valuation, you buy the car.

Toolkit: Section 31.1

You can review unit demand and valuation in the toolkit.

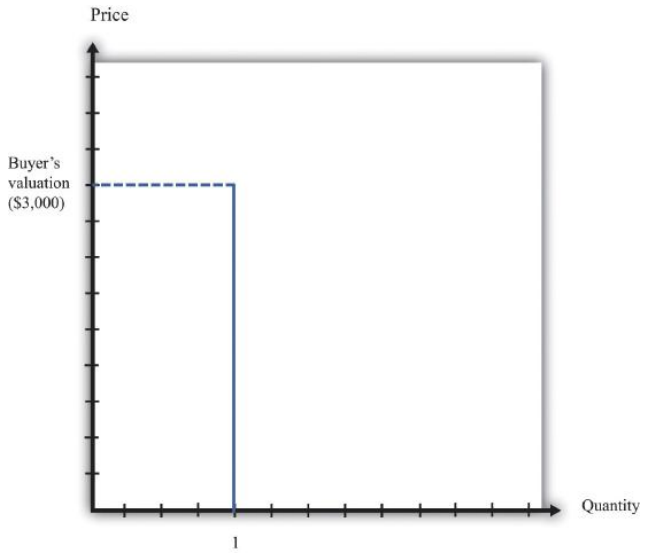


Figure 6.2.1: The Buyer's Valuation. The buyer follows the decision rule: "Buy if the price is less than the valuation."

If your valuation were $3,000, then you would, of course, prefer to pay much less. If the car is for sale for $2,990, then it is true that you would be better off buying the car than not, but you won't get much out of the deal. You would be happier only to the tune of $10 (more precisely, $10 worth of goods and services). If the car is for sale for $2,400, then you will be happier by an amount equivalent to $600 worth of goods and services. On the other hand, if the car were for sale for $3,001, you definitely would not want to buy at that price. Buying the car would actually make you slightly less happy.

The seller also has a valuation of the car. The seller is not willing to sell it at any price. For example, if her valuation is $2,000, she is equally happy keeping the car or not having the car and having an extra $2,000 worth of goods and services. If she can sell the car for more than $2,000, she will be better off. She won't sell the car for less than $2,000 because then she would be less happy than before. We can show the seller's willingness to sell in a way analogous to the buyer's willingness to buy. Figure 6.2.2 shows that she will not sell the car at a price less than $2,000, but she will sell once the price is greater than $2,000.

By analogy to unit demand, we call this the **unit supply curve**. It tells us the price at which she is willing to sell. Below her valuation, she is unwilling to supply to the market. Above her valuation, she is willing to sell the good. Whereas the buyer's valuation is the *absolute maximum* that the buyer is willing to pay, the seller's valuation is the *absolute minimum* that the seller is willing to accept.

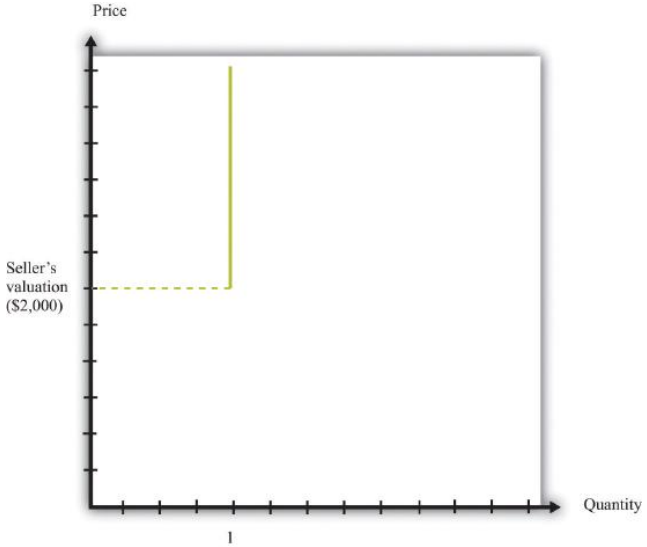


Figure 6.2.2: The Seller's Valuation. The seller follows the decision rule: "Sell if the price is greater than the valuation."

The buyer's valuation in our example is larger than the seller's valuation. This means it is possible to make both the buyer and the seller better off. The mere fact of transferring a good from someone who values it less to someone who values it more is an act that

creates value in the economy. We say that there are gains from trade available here.

Toolkit: Section 31.10

Total surplus is a measure of the gains from trade. In a single transaction,

$$total\ surplus\ =\ buyer's\ valuation\ -\ seller's\ valuation.$$

In this example, therefore, the **total surplus** is $1,000. This is the value created in the economy by the simple fact of transferring the car from a seller who values it less to a buyer who values it more. Figure 6.2.3 shows this graphically by combining the unit demand curve and the unit supply curve.

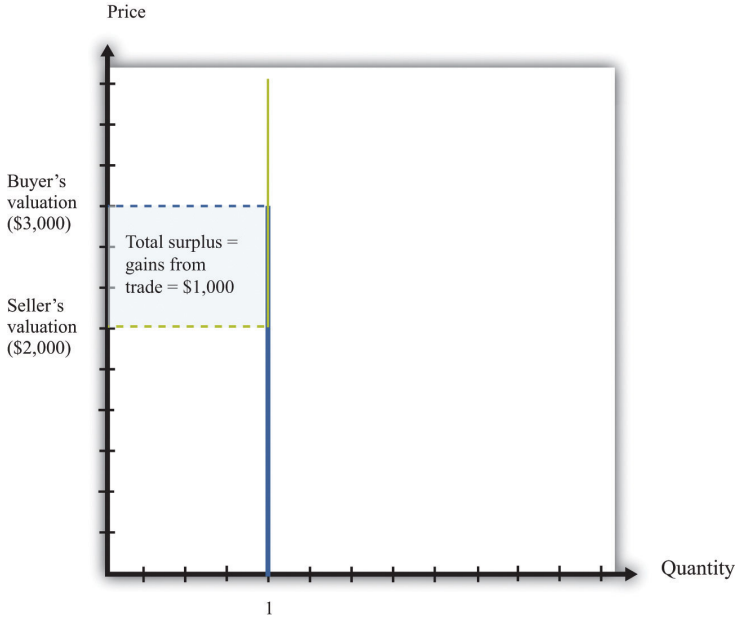


Figure 6.2.3: Buyer and Seller Valuations. The total surplus from a transaction is equal to the buyer's valuation minus the seller's valuation. Graphically, total surplus can be represented as a rectangle. The height of the rectangle is the difference in the valuations. The base of the rectangle is 1 because only one unit is being traded.

The buyer wants the price to be as low as possible, whereas the seller wants the price to be as high as possible. If both agree on a price of $2,100, for example, the buyer gets most of the surplus, and the seller does not get very much. If they agree on a price of $2,900, the situation is reversed: most of the benefit goes to the seller. The *distribution* of the value created depends on the price. Either way, though, they are both made better off by the trade, and in both cases the total surplus is the same (Figure 6.2.4).

Toolkit: Section 31.10

The buyer surplus is a measure of how much the buyer gains from a transaction, and the seller surplus is a measure of how much the seller gains from a transaction:

$$buyer\ surplus\ =\ buyer's\ valuation\ -\ price$$

and

$$seller\ surplus\ =\ price\ -\ seller's\ valuation.$$

The total surplus is the sum of the buyer surplus and the seller surplus.



Figure 6.2.4: The Distribution of Total Surplus. The distribution of surplus between the buyer and the seller depends on the price. A low price means that the buyer will get most of the surplus, while a high price means that the seller will get most of the surplus. The total surplus, however, is the same no matter what the price.

Economic Efficiency

When the buyer purchases the car from the seller, there is a reallocation of society's resources. Dollars have gone from the buyer to the seller, and the car has gone from the seller to the buyer. Economists have developed a specific criterion, called **efficiency**, for evaluating the way in which resources are allocated in a society.

It is actually easier to understand efficiency by looking at its opposite. Economists say that an allocation of resources is *inefficient* if there is some way to reallocate those resources that will make some people better off (that is, happier) without making anyone else worse off. For example, think about the situation where the buyer and seller have not traded the car. This allocation is inefficient. The buyer places a greater value on the car than does the seller, so it is inefficient for the car to remain with the seller. Any rearrangement of resources that makes some people better off without making anyone else worse off is *welfare improving*.

Toolkit: Section 31.11

Efficiency is the basis that economists use for judging the allocation of resources in an economy. Resources are allocated efficiently if there is no way to reallocate them to make someone better off without making anybody else worse off.

Before the buyer and the seller trade, the allocation of resources is inefficient. However, there are many different trades that make both the buyer and the seller better off. In fact, any trade between a price of $2,000 and a price of $3,000 is welfare improving. The only thing that matters for economic efficiency is that a trade takes place, so the **gains from trade** can be realized. No matter how the surplus is distributed between the buyer and the seller, the outcome is efficient as long as the trade occurs.

Determining a Price on craigslist

We now know that as long as the buyer's valuation for a good exceeds the seller's valuation, there are potential gains from trade. We have not yet explored the mechanisms that allow trade to occur, nor have we explained what determines the price at which trades occur. To begin with, we ignore the possibility of bargaining. Then there are only two steps for selling an item on craigslist:

1. The seller of an item posts a price (makes a take-it-or-leave-it offer).
2. The buyer either accepts or rejects the offer.

If the buyer accepts the seller's offer, then an exchange is made. But what offer will the seller make? The answer depends on how much the seller knows about the *buyer's* valuation of the good. There are two cases to consider:

1. **The seller knows the buyer's valuation.** The seller would like her surplus to be as large as possible. If she knows the buyer's valuation, what should she do? To answer this question, she must put herself in the buyer's shoes. If she sets a price that is greater than the buyer's valuation, then the buyer will reject the offer. But as long as the price is less than the buyer's valuation, the buyer will accept the offer. With this in mind, the seller will set a price slightly less than the buyer's valuation to capture almost the entire surplus. In our example, the seller should put the car on sale for $2,999. The seller gets $999 worth of surplus,

and the buyer gets $1. (Does it matter if the buyer also knows the seller's valuation? As a matter of pure economic theory, the answer is no. The buyer should be willing to buy provided he is getting some surplus—even if it is very little. After all, something is better than nothing. However, if the buyer knows that the seller is getting a lot of surplus, he may perceive this as unfair and might even choose not to buy out of spite. In reality, sellers often set a lower price, "giving away" some of the surplus to the buyer, to avoid this possibility.Technically, a take-it-or-leave-it offer is an example of an ultimatum game, which is discussed in Chapter 13.)

2. **The seller does not know the buyer's valuation.** This case is more likely and also much harder. The seller must trade off two different concerns. If she picks too high a price, then there is a risk of not making a sale at all. But the lower the price, the less surplus she gets in the event of a sale. There is no simple rule to know what price she should set. An economist looking from the outside finds this case more worrying than the first because it is possible that the gains from trade will be missed. If the seller offers a price that is greater than the buyer's valuation, then—under a take-it-or-leave-it offer—no trade takes place.

The knowledge of the buyer also matters. Suppose that the buyer knows the seller's valuation. Then he knows that there are possible gains from trade. In this case, it is natural to think that the buyer will try to negotiate with the seller, rather than just accept or reject the seller's offer. Indeed, if the buyer knows the seller's valuation, then we have the reverse of the first case. If the buyer offers a price slightly above the seller's valuation, then the buyer should be able to capture the entire surplus. We summarize this in Figure 6.2.5.

In practice, the buyer is also likely to try to negotiate if the seller's price leaves the buyer with very little surplus. Thus even though craigslist is apparently based on take-it-or-leave-it offers, a great deal of bargaining does in fact take place.

The seller makes a take-it-or-leave-it offer		Does the buyer know the seller's valuation?	
		Yes	No
Does the seller know the buyer's valuation?	Yes	The seller may need to let the buyer have some surplus. Otherwise the buyer may say no out of spite!	The seller sets a price just below the buyer's valuation and gets almost all the surplus.
	No	There is a risk of the price being too high. However, if the buyer can make a counteroffer, he will offer a price just above the seller's valuation and get almost all the surplus.	There is a risk of the price being too high, in which case the gains from trade are missed,

Figure 6.2.5: The Outcomes from a Take-It-or-Leave-It Offer

Key Takeaways

- If the seller's valuation of an object is less than the buyer's valuation of the same object, then there are gains to trade.
- One mechanism to reap the gains from trade when valuations are known is for the seller to post a price and the buyer to decide to purchase the good or not—that is, the seller makes a take-it-or-leave it offer.
- The way the gains to trade are split between the buyer and the seller depends on the way the bargaining occurs and the information the parties have about each other.
- An allocation is efficient if there is no way to make someone better off without also making somebody else worse off.

check your understanding

1. In this discussion, we assumed that the seller's valuation was less than the buyer's valuation. What would happen if that were not true?
2. Suppose the seller's valuation is less than the buyer's but that the buyer, not the seller, sets the price. What price would the buyer set? Would there still be gains from trade?

This page titled 6.2: craigslist and the Gains from Trade is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

6.3: eBay 6.3: eBayEdit Update titlecancel

Learning Objectives

1. What are the economics of an eBay auction?
2. How should I bid on eBay?
3. How are gains to trade determined and shared when there are multiple buyers?
4. What is the winner's curse?

So far we have supposed that there is only a single buyer and a single seller. If you are thinking about selling a good on craigslist, however, there are many potential buyers of your good. In addition, you probably don't know very much about the valuations of the different buyers. You might then like to find some way to make your buyers compete with each other. In other words, you might consider auctioning off the good instead.

You have probably at least visited the eBay site, and you may even have bought or sold an item on eBay. If so, you know it can be a convenient and efficient way to buy and sell goods. But what exactly is eBay? We answer this question by looking at the site from the perspective of participants. First we review how eBay works and look at it from the point of view of both a buyer and a seller. Then we bring some economic analysis to bear to better understand what is taking place on eBay and in other auctions.

Buying on eBay

Suppose you want to purchase something, such as a leather jacket, some cycling gloves, or a cell phone; the list of things that might interest you is endless. On the eBay page, you can search for the exact item you want to buy. Your search must be specific: if you search for "cell phone," you will find thousands of products. You need to know the exact model of phone you want, and even then you may find multiple items for sale.

Auctions on eBay have several characteristics, including the identity of the seller, the time limit on the auction, the acceptable means of payment, the means of delivery, and the reserve price.

- **Seller identity.** Unlike when you purchase from a store, you do not get to see the seller on eBay, and you cannot simply walk away with the good. This may concern you. After all, how do you know the seller will actually ship you the good after you have paid? How do you know if the product will work? This is a worry for you, but it is also a worry for any reputable seller. Sellers want you to trust them, so there are mechanisms to allow you to find out about sellers. On the eBay page, you can find detailed information about the seller, including a number—called a feedback score—that indicates the number of positively rated sales by that seller. If you dig a bit deeper, you can even find reviews of the performance of the seller.
- **Time remaining in the auction.** Online auctions have a fixed time limit. When you go to the auction site for a particular good, you will see the amount of time left in the auction. Much of the action in an auction often comes very near the end of the bidding.
- **Means of payment.** When you buy a good on eBay, you are not able to simply give the seller some cash. The means of payment accepted by the seller are indicated in the auction information. In many cases, sellers use an electronic payment system, such as PayPal.
- **Means of delivery.** The seller must have a way of shipping the good to you—perhaps via FedEx or another package delivery service. The auction specifies who pays the cost of shipping.
- **Reserve price.** The seller will frequently specify a minimum price, called a reserve price. As a potential buyer, however, you will not see the actual reserve price; the only information you will see is whether or not the reserve price has been reached. A natural reserve price for the seller is her valuation of the good. (More precisely, the reserve price would also include the fee that the seller must pay to eBay in the event of a sale. The only reason for a seller to set a higher valuation is if she thinks she might do better trying to auction the good again at some point in the future rather than settling for a low price in the current auction.)

You participate as a buyer in an eBay auction by placing a bid. For some products, you also have an option of clicking "Buy It Now," where you can purchase the good immediately. In other words, sellers sometimes make a take-it-or-leave it offer as well as offer an auction. To understand the details of the bidding process, look first at the description of how to bid on eBay:

Once you find an item you're interested in, it's easy to place a bid. Here's how:

1. Once you're a registered eBay member, carefully look over the item listing. Be sure you really want to buy this item before you place a bid.

2. Enter your maximum bid in the box at the bottom of the page and then click the Place Bid button.
3. Enter your User ID and password and then click the Confirm Bid button. That's it! eBay will now bid on your behalf up to the maximum amount you're willing to pay for that item.

You'll get an email confirming your bid. At the end of the listing, you'll receive another email indicating whether you've won the item with an explanation of next steps.eBay Inc., "Help: How to Bid," accessed March 14, 2011, pages.ebay.com/help/new/contextual/bid.html.

Because participants in an eBay auction are not all present to bid at the same time, eBay bids for you. All you have to do is to tell it how much you are willing to pay, and eBay takes over. This is known as "proxy bidding" or "automatic bidding."

The exact way in which eBay bids for you is not transparent from this description. It works as follows. Once you input your maximum bid, eBay compares this to the highest existing bid. If your maximum bid is higher than the existing highest bid, then eBay raises the bid by an increment on your behalf. Unless someone bids more, you will win the auction. If someone does bid more (the maximum bid exceeds the highest bid), then you, by proxy, will respond. In this way, the highest bid increases. This process ends with the item going to the bidder with the highest maximum bid. However, the buyer does not pay the amount of the maximum bid. The buyer pays the amount of the next highest bid, plus the increment.

Let us see how this works through an example. Suppose there are two buyers who put in maximum bids of $100 and $120 for a cell phone. Suppose that the increment is $1 and the bidding starts at $50. Because the maximum bids exceed $50, the highest bid will increase by increments of $1 until reaching $100. At this point, the higher of the two maximums, $120, will cause the highest bid to increase by another increment to $101. After this, there is no further action: the other bidder effectively drops out of the auction. The item goes to the buyer who bid $120, and he pays a price of $101 (provided this exceeds the seller's reserve price).

A Decision Rule for Bidding on eBay

Now that you understand how the auction works, you must decide how to bid. Suppose there is only one auction for the good you want (rather than multiple sellers of similar goods). In this case, there is a remarkably simple decision rule to guide your bidding.

1. Decide on your valuation of the good—that is, the most you would be willing to pay for it.
2. Bid that amount.

This seems surprising. Your first reaction might well be that it is better to bid less than your valuation. But here is the key insight: the amount you actually pay if you win the auction doesn't depend on your bid. Your bid merely determines whether or not you win the auction.

If you pursue this strategy and win the auction, you will gain some surplus: the amount you pay will be less than the valuation you place on the good. If you don't win the auction, you get nothing. So winning the auction is better than not winning. If you bid more than your valuation, then there is a chance that you will have to pay more than your valuation. In particular, if the second-highest bidder puts in a bid that exceeds your valuation, then you will lose surplus. So this does not seem like a good strategy. Finally, if you bid less than your valuation, there is a chance that you won't win the auction even though you value the good more highly than anyone else. Therefore, you will lose the chance of getting some surplus. This is also not a good strategy.

Figure 6.3.1 illustrates one way of thinking about this. There are two possibilities: either your valuation is bigger than the highest competing bid or your valuation is smaller than the highest competing bid. We don't have to know anything about how other bidders are making their decisions. Part (a) of Figure 6.3.1 shows the first case. Here, there is a risk that you will lose out on surplus that you could have received. If you bid below the competing bid, you will lose the auction and hence lose out on the surplus. The surplus is the difference between your valuation and the competing bid, minus the increment.

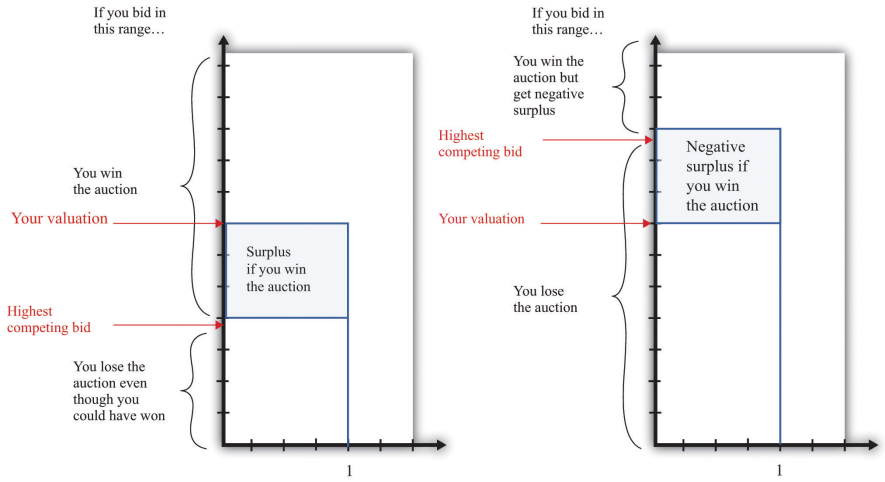


Figure 6.3.1: Why You Should Bid Your Valuation in an eBay Auction. These illustrations show why you should always bid your valuation on eBay.

Part (b) of Figure 6.3.1 shows the case where your valuation is less than the competing bid. Here the risk is that you will win the auction and then regret it. If you bid an amount greater than the competing bid, then you will win the auction, but the amount you will have to pay exceeds your valuation. Your loss is the difference between the competing bid and your valuation, plus the increment.

Although automatic bidding by proxy sounds very fancy, the eBay auction is really the same as "English auctions" that are familiar from television and movies. In an English auction, an auctioneer stands at the front of the room and invites bids. Bidding increases in increments until all but one bidder drops out. The winning bidder pays the amount of his bid. The winning bidder therefore pays an amount equal to the highest competing bid, plus the increment, just as in the eBay auction. The amount that she wins is her valuation minus the price she pays, just as in the eBay auction.

The bidding in an English auction can be exciting to watch; you can also have the excitement of seeing how bids evolve on eBay (at least if you are willing to keep logging on and hitting the "refresh" button). But we could also imagine that eBay could do something simpler. It could carry out a "Vickrey auction," which is named after its inventor, the Nobel-prize-winning economist William Vickrey. In a Vickrey auction, the auctioneer (1) collects all the bids, (2) awards victory in the auction to the highest bidder, and (3) makes this person pay an amount equal to the *second-highest* bid.

The Vickrey auction sometimes goes by the more technical name of a **second-price sealed-bid auction**. Most people think this kind of auction sounds very odd when they first hear of it. Why make the winner pay the second-highest bid? Yet it is almost exactly the same as the eBay auction or the English auction. In those auctions, as in the Vickrey auction, the winner is the person with the highest bid, and the winner pays the amount of the second-highest bid (plus the increment). The only difference is that there is no increment in the Vickrey auction; because the increment is typically a very small sum of money, this is a minor detail.

Selling on eBay

Sellers on eBay typically provide information on the product being sold. This is often done by creating a small web page that describes the object and includes a photograph. Sellers also pay a listing fee to eBay for the right to sell products. They also specify the costs for shipping and handling. After the sale is completed, the buyer and the seller communicate about the shipping, and the buyer makes a payment. Then the seller ships the product, pays eBay for the right to sell, and pockets the remainder.

As we mentioned previously, the buyer can provide feedback on the transaction with the seller. This feedback is important to the seller because transactions require some trust. A seller who has built a reputation for honesty will be able to sell more items, potentially at a higher price.

An Economic Perspective on Auctions

As an individual participating in an auction, you have two concerns: (1) whether or not you win and (2) how much you have to pay. As economists observing from outside, there are other perspectives. Auctions play a very valuable role in the economy. They represent a leading way in which goods are bought and sold—that is, they represent a mechanism for trade.

As we have already stated, voluntary trade is a good thing because it creates value in the economy. Every transaction allows a good or service to be transferred from someone who values it less to someone who values it more. The English auction, such as on eBay, is attractive to economists because it does something more. It ensures that the good or the service goes to the person who values it the most—that is, it ensures that the outcome is efficient. It also has the fascinating feature that it induces people to reveal their valuations through their bids.

eBay is just one example of the many auctions you could participate in. There are auctions for all types of goods: treasury securities, art, houses, the right to broadcast in certain ranges of the electromagnetic spectrum, and countless others. These auctions differ not only in terms of the goods traded but also in their rules. For example, firms competing for a contract to improve a local road may submit sealed bids, with the contract going to the lowest bidder to minimize the costs of the project. Of course, other elements of the bid, including the reputation of the bidder, may also be taken into account.

Complications

The eBay auction sounds almost too good to be true. It is easy to understand, brings forth honest bids, and allocates the good to the person who values it the most. Are there any problems with this rosy picture?

Multiple Sellers

Suppose you have a video-game system for sale. You can put it up for auction on eBay, but you must be aware that many other people could be listing the identical item. What will happen?

First, your potential buyers will most likely look (and bid) at multiple auctions, not only your auction. Second, potential buyers will not be eager to bid in your auction. After all, if they don't win your auction, they can always hope for better luck in another auction. It follows that buyers may decide it is no longer such a good strategy to bid their valuations. Buyers who bid their valuations might end up paying a high price if they win an auction where someone else placed a relatively high bid. Such buyers might be more successful taking the chance of losing one auction and winning another in which the bidding is lower. As buyers monitor other auctions, they will also start to get a sense of how much other people are willing to pay and will adjust their bidding accordingly.

Unfortunately, we can't give you such simple advice about what to do as a buyer in these circumstances. It is not easy to develop the best bidding strategy. In fact, problems like this can be so hard that even expert auction theorists have not fully worked them out.

Tacit Collusion

Another concern is that bidders might want to find some way to collude. As a simple example, suppose there are three bidders for a good with an increment of $1. One bidder has a valuation of $50, one has a valuation of $99, and one has a valuation of $100. In an eBay auction, the winning bid would be $100, but the winner would end up with no surplus (because he would pay $99 plus the $1 increment). Now suppose that the two high-value bidders make an agreement. As soon as the third bidder drops out, they toss a coin. If it comes up heads, Mr. $99 drops out. If it comes up tails, Ms. $100 drops out.

This means that with 50 percent probability, Ms. $100 wins, pays $51, and gets a surplus of $49. With 50 percent probability Mr. $99 wins, pays $51, and gets a surplus of $48. Both buyers prefer this. It's certainly better for Mr. $99, who had no chance of winning before. It is also better for Ms. $100 because even though she may no longer win the auction, she stands to get some surplus if she does win. Of course, the seller wouldn't like this arrangement at all. And the dispassionate economist observing from afar doesn't like it either because sometimes the good may not go to the person who values it the most.

Explicit collusion of this type may very well be illegal, and it is also very hard to carry out. Yet it may be possible for buyers to collude indirectly, and there is speculation that such collusion is sometimes observed on eBay.

The Winner's Curse

We have been supposing throughout that potential buyers know their own valuations of the good being auctioned. In most circumstances, this seems reasonable. Valuations are typically a personal matter that depend on the tastes of the individual buyer.

Occasionally, however, a good with an objective monetary value that is *unknown* to potential buyers may be auctioned. A classic example is the drilling rights to an oilfield. There is a certain amount of oil in the ground, and it will earn a certain price on the market. However, bidders do not know these values in advance and must make their best guess.

It is easiest to see what can happen here with a numerical example. Suppose the true (but unknown) value of an oilfield is $100 million. Suppose there are five bidders, whose guesses as to the value of the oilfield are summarized in Table 6.3.1 "Valuations of Different Bidders in a Winner's Curse Auction". Notice that these bidders are right on average, but two overestimate the value of the field, and two underestimate it. Imagine that the bidders decide to follow the strategy that we recommended earlier and bid up to their best guess. Bidder E will win. He will have to pay the second-highest bid of $105 million, which is more than the oilfield is worth. He will lose $5 million.

Table 6.3.1: Valuations of Different Bidders in a Winner's Curse Auction

	Bidder A	Bidder B	Bidder C	Bidder D	Bidder E
Valuation ($ million)	90	95	100	105	110

The problem here is that the person who will win the auction is the person who makes the *worst overestimate* of the value of the field. Evidently it is not a good strategy in this auction to bid your best guess. You should recognize that your best guess may be inaccurate, and if you overestimate badly, you may win the auction but lose money. This phenomenon is known as the **winner's curse**. Your best strategy is therefore to bid less than you actually think the oilfield is worth. But how much less should you bid? That, unfortunately, is a very hard question for which there is no simple answer. It depends on how accurate you think your guess is likely to be *and* how accurate you think other bidders' guesses will be.

Key Takeaways

- On eBay, the best strategy is to bid your true valuation of the object.
- Auctions, like eBay, serve to allocate goods from sellers to buyers.
- If the winner's curse if present, then you will want to bid less than your estimate of the value of the object.

check your understanding

1. Suppose you bid less than your valuation on eBay. Explain how you could do better by bidding a little more.
2. Why didn't the winner's curse have an effect on your bidding in eBay?

This page titled 6.3: eBay is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

6.4: Supply and Demand 6.4: Supply and DemandEdit Update titlecancel

Learning Objectives

1. What is the outcome of an auction with a large number of buyers and sellers?
2. What is the market demand curve?
3. What is the market supply curve?
4. What is the equilibrium of perfectly competitive markets?

An auction mechanism such as eBay is a natural thing for a seller to use if there are a large number of potential buyers for a good. But what happens if there is also a large number of potential sellers? In this section, we consider what might happen when we have a large number of potential buyers *and* a large number of potential sellers of a good.

We have already explained that it is very difficult to analyze what would happen on eBay when there are multiple buyers and sellers, but we can make a better guess about what will happen on a site like craigslist. As a buyer, you will look for the lowest price out there, bargain with sellers who post high prices, or both. As a seller, you would look at the prices posted by others and realize that you probably should set your price fairly close to those prices. In addition, we have some evidence that can help us understand the likely outcome in a world of many buyers and many sellers. It comes from looking at "double oral auctions." "Double" refers to the fact that there are a large number of buyers and sellers. "Oral" refers to the way in which the auction is conducted.

Double Oral Auctions

In a **double oral auction**, there is a large number of buyers, each of whom potentially has a different valuation of the good. There is also a large number of sellers, each of whom potentially has a different valuation of the good. Buyers and sellers negotiate with each other, one on one. If they cannot agree to a deal, either party can move on at any time and try to find someone else to bargain with.

Until quite recently, auctions such as this were common in many financial markets and commodity markets. These markets sometimes go by the name *pit markets* because buyers and sellers meet in a frenzy of activity in a trading area called the pit. Traders can hear and see the negotiations of others and often have access to the prices at which deals have been done. This means that both buyers and sellers have lots of information about what price is prevailing in the market. Chapter 10 has much more to say about these markets.

Economists have also conducted experiments in which they have put people in simulated pit markets to find out how they behave. The result is quite remarkable, but before we explain what happens, we need a framework to help us think about such markets.

Demand: Many Buyers

Suppose we are considering the purchase of a gaming console by a group of buyers. Each potential buyer has his own valuation. Some might be willing to pay as much as $700. Others might be willing to spend much less. After all, how much you are willing to pay for a gaming console depends on your income, how much you like playing, what equipment you currently own, and so on.

Each potential buyer has a unit demand curve like the one we saw in Figure 6.4.1. We can add these unit demand curves together to get a picture of demand in the entire market: the **market demand curve**. For example, suppose only one person is willing to buy if the price is $700. However, suppose there is another buyer with a valuation of $660. If consoles were on sale for $660, then both individuals would want to purchase. At $660, in other words, the quantity demanded is 2. Perhaps the buyer with the next-highest valuation is willing to pay $640. If the price is $640, the quantity demanded is 3. Figure 6.4.1 shows what happens when we add together all these unit demand curves. The result is a downward sloping relationship that shows us how many units would be demanded at any given price.

Toolkit: Section 31.9

The market demand curve tells us how many units of a good or a service will be demanded at any given price. The market demand curve is obtained by adding together the individual demand curves in the economy and obeys the law of demand: as the price decreases, the quantity demanded increases.

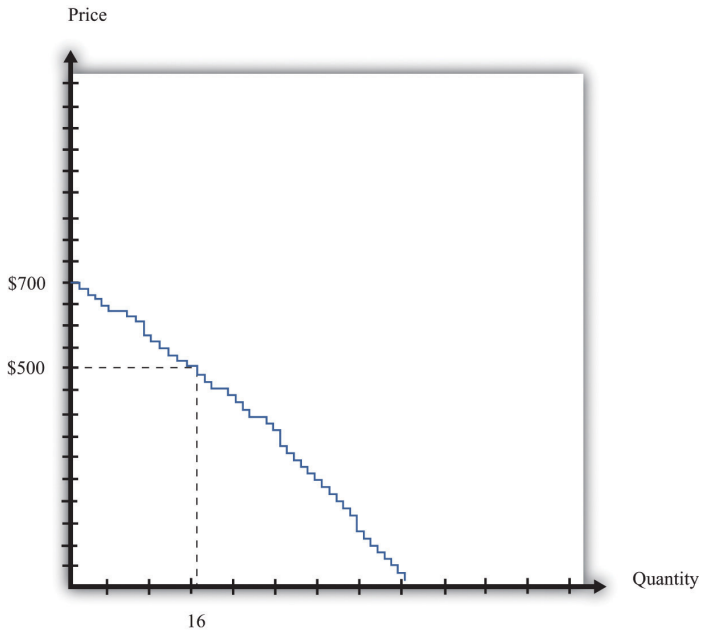


Figure 6.4.1: Obtaining the Market Demand Curve. We can add together the unit demand curves of different individuals in the economy to get the ***market demand curve***.

Supply: Many Sellers

We saw earlier that each potential seller has a unit supply curve. If the price is less than a seller's valuation, she will not sell the good, but when the price becomes greater than her valuation, she will be willing to sell. Just as we added together the unit demand curves to get the market demand curve, so too can we add together the unit supply curves to get the **market supply curve**.

Toolkit: Section 31.9

The market supply curve tells us how many units of a good or a service will be supplied at any given price. The market supply curve is obtained by adding together the individual supply curves in the economy and typically slopes upward: as the price increases, the quantity supplied to the market increases.

In Figure 6.4.2, we see that the lowest valuation in the market is $150. There is one seller willing to sell a console at that price. As the price increases, more and more sellers will find the price attractive and will want to sell. For example, there are 11 potential sellers with a valuation less than $350. Thus, at this price, 11 consoles will be supplied to the market.

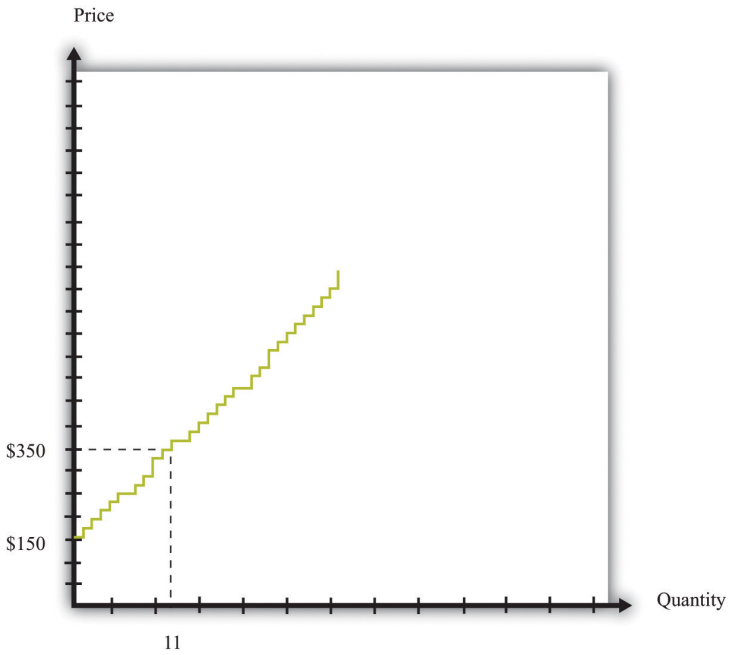


Figure 6.4.2: Obtaining the Market Supply Curve. We can add together the unit supply curves of different individuals in an economy to get the ***market supply curve***.

Equilibrium

Figure 6.4.1 and Figure 6.4.2 tell us the number of buyers willing to buy and the number of sellers willing to sell at each price.

Figure 6.4.3 shows what happens if we combine the demand curve and the supply curve on the same diagram. One point jumps out at us: the place where the demand and supply curves meet. In our example, this is at $480 and a quantity of 21 units. At this point, *the number of buyers with a valuation greater than the price is the number of sellers with a valuation less than the price*. If buyers and sellers were presented with this price, none would find themselves unable to transact. At this price, there is an exact match between the number of buyers and sellers.

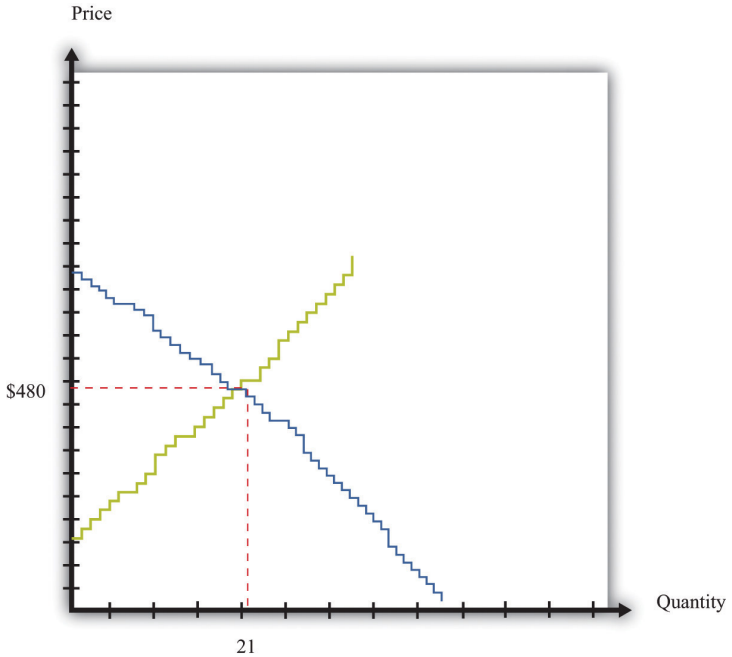


Figure 6.4.3: Market Equilibrium. In this figure, we combine the demand and supply curves to find the ***equilibrium*** *price and quantity in the market.*

Toolkit: Section 31.9

Equilibrium in a market refers to an equilibrium price and an equilibrium quantity and has the following features:

- Given the equilibrium price, sellers supply the equilibrium quantity.
- Given the equilibrium price, buyers demand the equilibrium quantity.

Equilibrium is not only a point on a graph. It is a prediction about a possible outcome in a situation where a large number of buyers and sellers meet with the possibility of trading. It seems plausible that in a situation where a large number of buyers and sellers can meet and trade with each other, most will end up trading at or near the **equilibrium price**.

The equilibrium outcome is plausible because, at any other price, there will be a mismatch of buyers and sellers. Imagine, by contrast, that the buyers and sellers of our example are currently trading at $600, well above the equilibrium price of $480. At this high price, many more people want to sell than want to buy. Buyers would rapidly realize that they are in a strong bargaining position: if many sellers want your business, you can make them compete with each other and force price decreases. In fact, whenever the price is above equilibrium, the mismatch of buyers and sellers will tend to decrease prices.

By similar reasoning, a price of, say, $400 would also result in a mismatch between buyers and sellers. In this case, though, there are more people who want to buy than sell. Sellers can make buyers compete with each other, leading to price increases. At any price below the equilibrium price, prices will tend to increase.

Perfectly Competitive Markets

Economists formalize the intuition we have just developed with the most famous framework in all of economics: supply and demand.The definition is repeated and discussed in more detail in Chapter 8; we make extensive use of it in other chapters.

Toolkit: Section 31.9

Supply and demand is a framework we use to explain and predict the equilibrium price and quantity of a good. This framework illustrates the willingness to sell (market supply) and buy (market demand) on a graph with price on the vertical axis and units of the good or the service on the horizontal axis. A point on the market supply curve shows the quantity that suppliers are willing to sell for a given price. A point on the market demand curve shows the quantity that demanders are willing to buy for a given price. The intersection of supply and demand determines the equilibrium price and quantity that will prevail in the market. A basic supply-and-demand framework is shown in Figure 6.4.4.

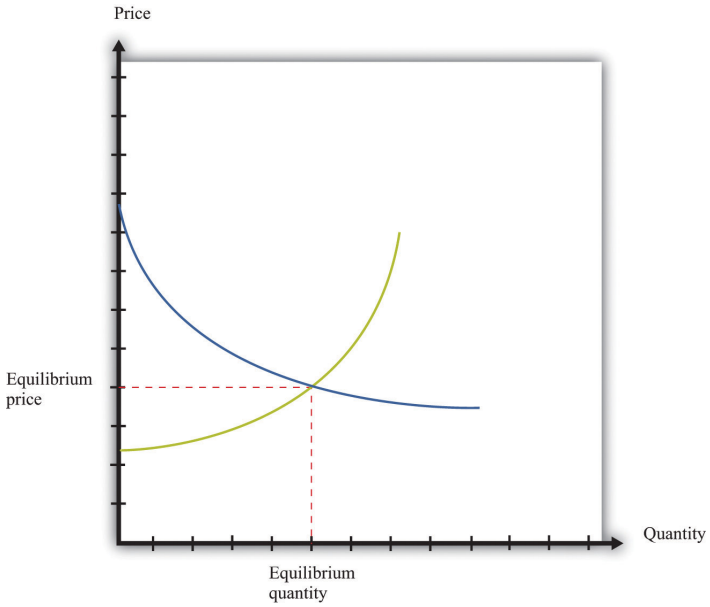


Figure 6.4.4: Supply and Demand. When we have a large number of buyers and sellers of an identical good or service, the equilibrium price and quantity are determined by the intersection of the supply and demand curves.

The *position* of the demand curve depends on many things, such as income and the prices of other goods. A change in any of these will cause the entire demand curve to shift. Likewise, the *position* of the supply curve depends on factors such as a supplier's costs. A change in these will cause the entire supply curve to shift. When one (or both) of the curves shifts, the equilibrium price and quantity change.

Experience with double oral auctions, both in the laboratory and in actual pit markets, tells us that trading will typically settle down close to the equilibrium price within a relatively short period of time. In a situation where there is a large number of people buying and selling an identical good, we say that we have a **competitive market**. We expect that most trades will take place at or close to the equilibrium price, and the quantity traded will be approximately equal to the **equilibrium quantity**. In fact, even when the number of participants in the auction is relatively small, we often find that a double oral auction still gets close to this equilibrium price and quantity. This is the remarkable finding that we mentioned earlier: in a double oral auction, the number of transactions and the prices of these transactions are usually very close to the equilibrium predicted by **supply and demand**.

Toolkit: Section 31.9

Suppose a market has the following two characteristics:

1. There are many buyers and many sellers, all of whom are small relative to the market.
2. The goods being traded are perfect substitutes.

In this case we say that we have a competitive market (sometimes called a perfectly competitive market). Buyers and sellers both take the price as given. This means that they think their actions have no effect on the price in the market, which in turn means we can employ the supply-and-demand framework.

The Gains from Trade in Equilibrium

Suppose all the transactions in Figure 6.4.3 take place at the equilibrium price of \$480. What can we say about the surplus received by buyers and sellers? Each individual transaction looks like those we examined in Chapter 6, Section 6.2. The total surplus from any given transaction is equal to the difference between the buyer's valuation and the seller's valuation. The **buyer surplus** is the difference between his valuation and \$480. The **seller surplus** is the difference between the price and her valuation. For example, Figure 6.4.5 shows the gains from trade if a buyer with a valuation of \$630 matches up with a seller whose valuation is \$230:

$$buyer\ surplus\ =\ \$630-\$480=\$150,\ \ seller\ surplus\ =\ \$480-\$230=\$250,$$

and

$$total\ surplus\ =\ \$150+\$250=\$400.$$

The transaction generates \$400 worth of surplus: \$150 goes to the buyer, and \$250 goes to the seller.

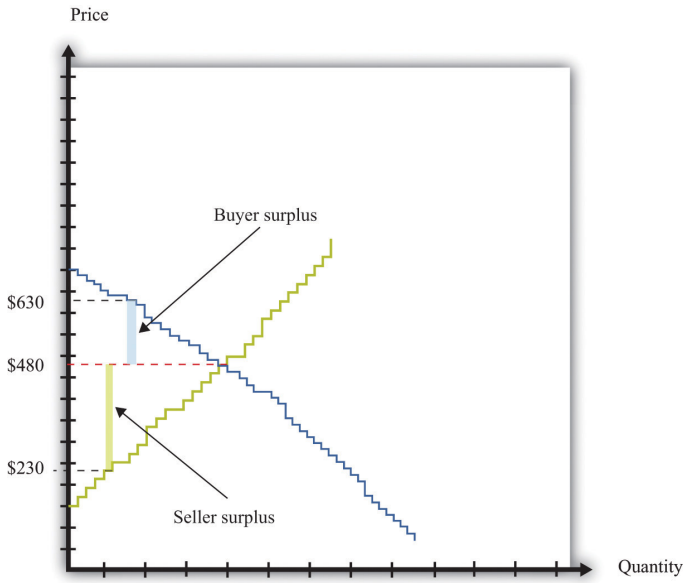


Figure 6.4.5: The Gains from Trade in a Single Transaction in Market Equilibrium. Each transaction in the market generates surplus.

We could draw exactly the same diagram for all 21 transactions in the market. If we combine them, we would end up with Figure 6.4.6. The total surplus accruing to the buyers is equal to the area below the demand curve and above the price. The total surplus accruing to the sellers is equal to the area above the supply curve and below the price. The total surplus—that is, the total gains from trade in this market—is the sum of the buyer surplus in the market and the seller surplus in the market. The total surplus is therefore the area between the supply curve and the demand curve.

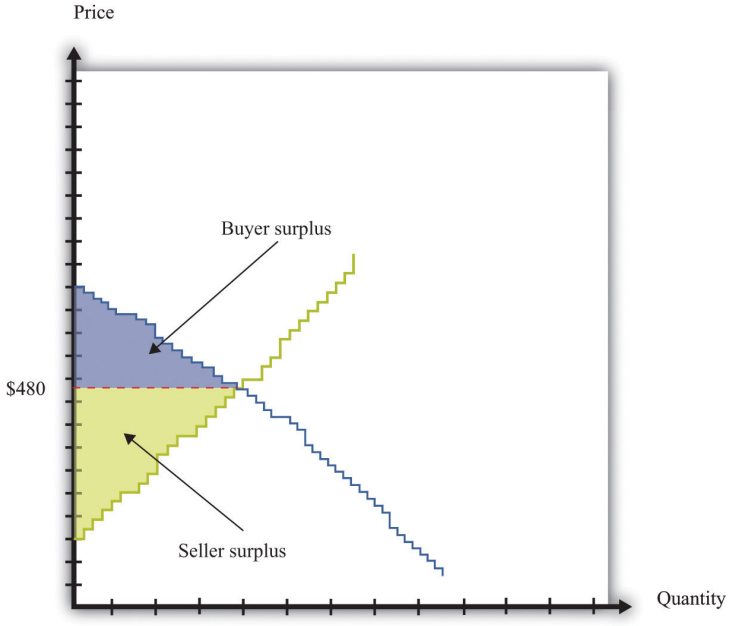


Figure 6.4.6: *Surplus in Equilibrium*. If we add the surplus from all trades in the market, supposing that they all take place at $480, we obtain the total surplus in the market.

If you look at Figure 6.4.6, something else may become apparent to you. *All the gains from trade have been exhausted in the market*. If buyers and sellers trade at the market price, then they manage to achieve all the gains from trade that are possible in this market because

- every transaction that has been carried out has created surplus;
- any further transaction would generate negative surplus.

The first statement is true because all trades are voluntary. We can see that the second statement is true by imagining trying to match up another buyer and seller. All the buyers with valuations greater than $480 have now made a purchase. So every remaining potential buyer has a valuation less than $480. All the sellers with valuations less than $480 have now made a sale. So every potential seller has a valuation greater than $480. It follows that there is no mutually beneficial transaction to be carried out.

This is a truly remarkable result. A market where all potential buyers and sellers take as given the equilibrium price allows all the possible gains from trade to be realized. Thus a market is a very effective mechanism for generating an efficient allocation of resources. This is why economists place so much emphasis on markets and "market solutions" to economic problems. Markets allow buyers and sellers to come together to make mutually beneficial trades. Economists believe that, as far as possible, we should create circumstances in which people can meet and carry out voluntary transactions.

Although this argument for markets is very powerful, we must be careful. Buyers and sellers may benefit from trading, but sometimes other people not involved in the transaction may also be affected. For example, suppose you fill up your car with gas at your local gas station. Presumably, you benefit from this transaction—otherwise you wouldn't have bought the gas. Likewise, the gas station owner benefits from the transaction—otherwise the owner wouldn't have sold it to you. But your purchase will contribute to smog and air pollution when you drive the car, affecting other people in the vicinity. To the extent that you make a contribution to global climate change, your little transaction has the potential to have an effect—a very tiny effect but an effect nonetheless—on everyone else on the planet. As a more positive example, going to college is presumably a mutually beneficial transaction between you and your school. But many others may eventually benefit from your education as well. In Chapter 14, we consider such uncompensated costs and benefits in detail.

Key Takeaways

- In a perfectly competitive market, buyers and sellers take the prices as given.
- In the equilibrium of a perfectly competitive market, there are no further gains to trade.
- The outcome of a double oral auction and the supply-and-demand framework are the same.

check your understanding

1. Look at Figure 6.4.3. How could the equilibrium price be greater than $480?
2. Suppose there are two buyers. The first has a demand curve given by $quantity = 5 - 0.5 \times price$. The second one has a demand curve of $quantity = 15 - 1.5 \times price$. What is the market demand at $1? Suppose there is a total supply of 10 units in this market. What is the equilibrium price? How is the surplus allocated?

This page titled 6.4: Supply and Demand is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

6.5: Production Possibilities 6.5: Production PossibilitiesEdit Update titlecancel

Learning Objectives

1. Where do the gains from trade come from?
2. What determines who produces which good?

So far we have discussed several different ways in which individuals trade with one another, including individual bargaining and Internet sites such as eBay and craigslist. We have considered situations with one seller and one buyer, one seller and many buyers, and many sellers and buyers. But why do we trade so much? Why is trade so central to our lives and indeed to the history of the human race?

On a typical craigslist website, many services are offered for sale. They are listed under categories such as financial, legal, computer, beauty, and so on. If you click on one of these headings and follow one of the offers, you typically find that someone is willing to provide a service, such as legal advice, in exchange for money. Sometimes there are offers to barter: to exchange a service for some other service or for some specific good. For example, we found the following offers listed on craigslist.These are actual offers that we found on craigslist, edited slightly for clarity.

Hello, I am looking for a dentist/oral surgeon who is willing to remove my two wisdom teeth in exchange for furniture repair and refinishing. If preferred, I'll come to your office and show you my teeth beforehand. Take a look at some of the work I have posted on my web page…quality professional furniture restoration. Bring new life to your antiques!

We are new to the area and are looking for a babysitter for casual or part-time help with our three little girls. My husband is a chiropractor and offers adjustments, and I am a vegan and raw foods chef offering either culinary classes or prepared food in exchange for a few hours of babysitting each week.

I have a web design company,…I figure I'd offer to barter in this slow economy. If you got something you'd be willing to trade for a website, let me know and maybe we can work something out!

These offers provide a glimpse into why people trade. Some people are relatively more productive than others in the production of certain goods or services. Hence it makes sense that people should perform those tasks they are relatively good at and then in some way exchange goods and services. These offers reveal both a reason for trade and a mechanism for trade.

As individuals, we are involved in the production of a very small number of goods and services. The person who cuts your hair is probably not a financial advisor. It is unlikely that your economics professor also moonlights as a bouncer at a local nightclub. By contrast, we buy thousands of goods and services—many more goods and services than we produce. *We specialize in production* and *generalize in consumption*. One motivation for trade is this simple fact: we typically don't consume the goods we produce, and we certainly want to consume many more goods than we produce. Yet that prompts the question of why society is organized this way. Why do we live in such a specialized world?

To address this question, we leave our modern, complicated world—the world of eBay, craigslist, and the Internet—behind and study some very simple economies instead. In fact, we begin with an economy that has only one individual. This allows us to see what a world would look like without any trade at all. Then we can easily see the difference that trade makes.

Production Possibilities Frontier for a Single Individual

Inspired by the craigslist posts that we saw earlier, imagine an economy where people care about only two things: web pages and vegan meals. Our first economy has a single individual—we call him Julio—who has 8 hours a day to spend working. Julio can spend his time in two activities: web design and preparing vegan meals. To be concrete, suppose he can produce 1 web page per hour or 2 vegan meals per hour. Julio faces a time allocation problem: how should he divide his time between these activities?We study the time allocation problem in Chapter 4.

The answer depends on both Julio's productivity and his tastes. We start by looking at his ability to produce web pages and vegan meals in a number of different ways. Table 6.5.1 "Julio's Production Ability" shows the quantity of each good produced per hour of Julio's time. Julio can produce either 2 vegan meals or 1 web page in an hour. Put differently, it takes Julio half an hour to prepare a meal and 1 hour to produce a web page. These are the technologies—the ways of producing output from inputs—that are available to Julio.

Toolkit: Section 31.17

A technology is a means of producing output from inputs.

Table 6.5.1: Julio's Production Ability

Vegan Meals per Hour	Web Pages per Hour
2	1

We could write these two technologies as equations:

$$quantity\ of\ vegan\ meals\ =\ 2\ \times\ hours\ spent\ cooking$$

and

$$quantity\ of\ web\ pages\ =\ hours\ spent\ on\ web\ design.$$

Or we can draw these two technologies (Figure 6.5.1). The equations, the figure, and the table are three ways of showing exactly the same information.

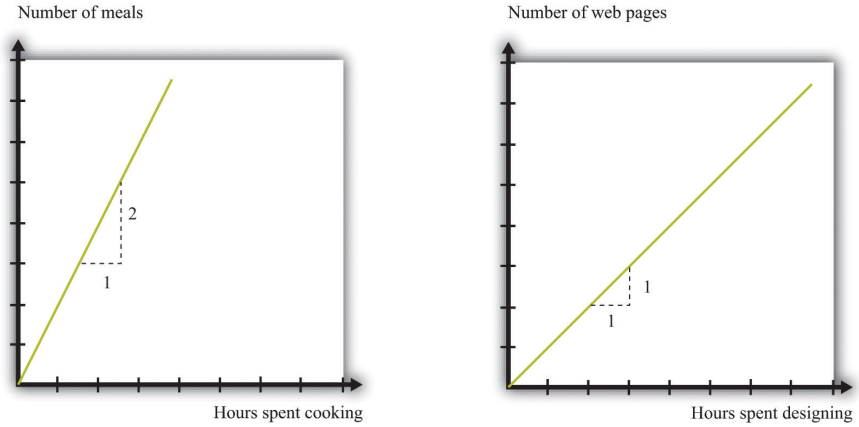


Figure 6.5.1: Julio's Production Ability. These figures show Julio's technologies for producing vegan meals and producing web pages.

When we add in the further condition that Julio has 8 hours available each day, we can construct his different possible production choices—the combinations of web pages and vegan cuisine that he can produce given his abilities and the time available to him. The first two columns of Table 6.5.2 "Julio's Production Possibilities" describe five ways Julio might allocate his 8 hours of work time. In the first row, Julio allocates all 8 hours to preparing vegan meals. In the last row, he spends all of his time in web design. The other rows show what happens if he spends some time producing each service. Note that the total hours spent in the two activities is always 8 hours.

The third and fourth columns provide information on the number of vegan meals and web pages that Julio produces. Looking at the first row, if he works only on vegan meals, then he produces 16 meals and 0 web pages. If Julio spends all of his time designing web pages, then he produces 0 vegan meals and 8 web pages.

Table 6.5.2: Julio's Production Possibilities

Time Spent Producing		Goods Produced	
Vegan Meals	**Web Pages**	**Vegan Meals**	**Web Pages**
8	0	16	0
2	6	12	2
4	4		

Time Spent Producing		Goods Produced	
6	2		
0	8	0	8

We can also illustrate this table in a single graph (Figure 6.5.2) that summarizes Julio's production possibilities. The quantity of vegan meals is on the horizontal axis, and the quantity of web pages is on the vertical axis. To understand Figure 6.5.2, first consider the vertical and horizontal intercepts. If Julio spends the entire 8 hours of his working day on web design, then he will produce 8 web pages and no vegan meals (point A). If Julio instead spends all his time cooking vegan meals and none on web design, then he can produce 16 vegan meals and 0 web pages (point B).

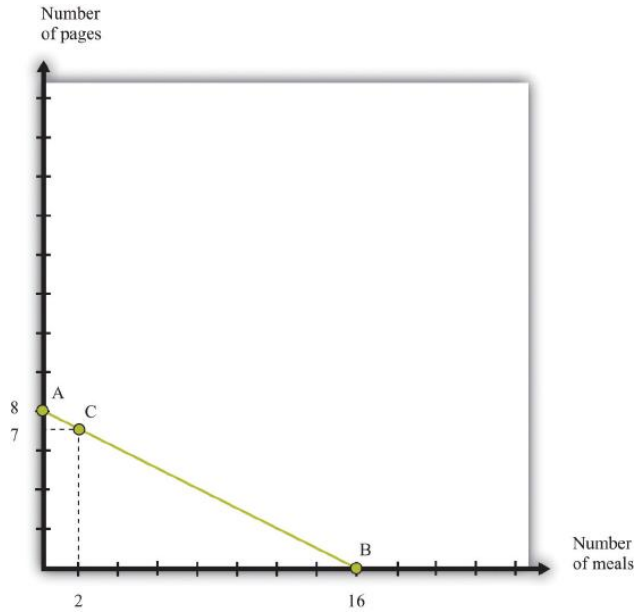


Figure 6.5.2: Julio's Production Possibilities. Julio's production possibilities frontier shows the combinations of meals and web pages that he can produce in an 8-hour day.

The slope of the graph is −1/2. To see why, start at the vertical intercept where Julio is producing only web pages. Suppose that he reduces web-page production by 1 page. This means he will produce only 7 web pages, which requires 7 hours of his time. The hour released from the production of web design can now be used to prepare vegan meals. This yields 2 vegan meals. The resulting combination of web pages and vegan meals is indicated as point C. Comparing points A and C, we can see why the slope is −1/2. A reduction of web-page production by 1 unit (the rise) yields an increase in vegan meals production of 2 (the run). The slope—rise divided by run—is −1/2.

Given his **technology** and 8 hours of working time, all the combinations of vegan meals and web pages that Julio can produce lie on the line connecting A and B. We call this the **production possibilities frontier**. Assuming that Julio equally likes both web design and vegan meals and is willing to work 8 hours, he will choose a point on this frontier.All of this may seem quite familiar. The production possibility frontier for a single individual is the same as the time budget line for an individual. See Chapter 4 for more information.

Toolkit: Section 31.12

The production possibilities frontier shows the combinations of goods that can be produced with available resources. It is generally illustrated for two goods.

What is the cost to Julio of cooking one more meal? To cook one more meal, Julio must take 30 minutes away from web design. Because it takes 30 minutes to produce the meal, and Julio produces 1 web page per hour, *the cost of producing an additional vegan meal is half of a web page*. This is his **opportunity cost**: to do one thing (produce more vegan meals), Julio must give up the opportunity to do something else (produce web pages). Turning this around, we can determine the opportunity cost of producing an extra web page in terms of vegan meals. Because Julio can produce 1 web page per hour or cook 2 meals per hour, *the opportunity*

cost of 1 web page is 2 vegan meals. The fact that Julio must give up one good (for example, web pages) to get more of another (for example, vegan meals) is a direct consequence of the fact that Julio's time is scarce.

Could Julio somehow produce more web pages and more vegan meals? There are only two ways in which this could happen. First, his technology could improve. If Julio were able to become better at either web design or vegan meals, his production possibilities frontier would shift outward. For example, if he becomes more skilled at web design, he might be able to produce 3 (rather than 2) web pages in 2 hours. Then the new production possibilities frontier would be as shown in part (a) of Figure 6.5.3.

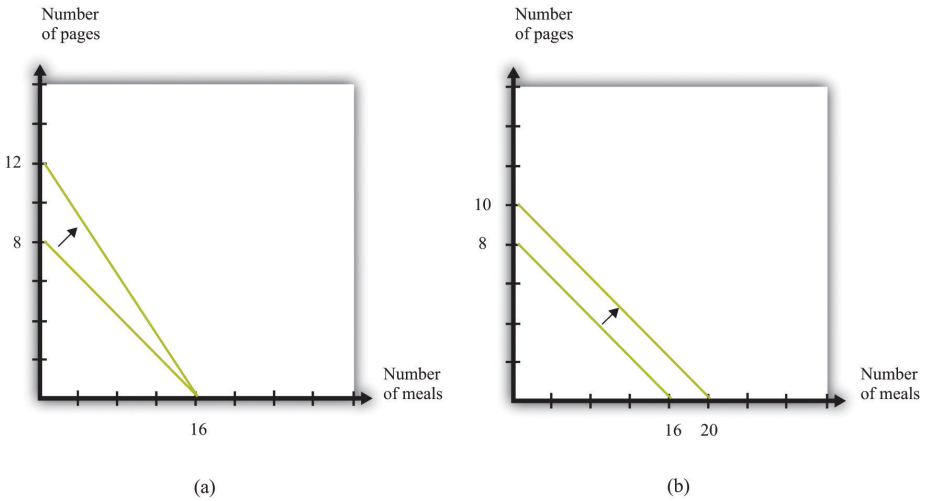


Figure 6.5.3: Two Ways of Shifting the Production Possibilities Frontier Outward. There are two ways in which Julio could produce more in a day: he could become more skilled, or he could work harder.

Alternatively, Julio could decide to work more. We have assumed that the amount of time that Julio works is fixed at 8 hours. Part (b) of Figure 6.5.3 shows what the production possibilities frontier would look like if Julio worked 10 hours per day instead of 8 hours. This also has an opportunity cost. If Julio works longer, he has less time for his leisure activities.

We have not yet talked about where on the frontier Julio will choose to allocate his time. This depends on his tastes. For example, he might like to have 2 vegan meals for each web page. Then he would consume 4 web pages and 8 meals, as in Figure 6.5.4.

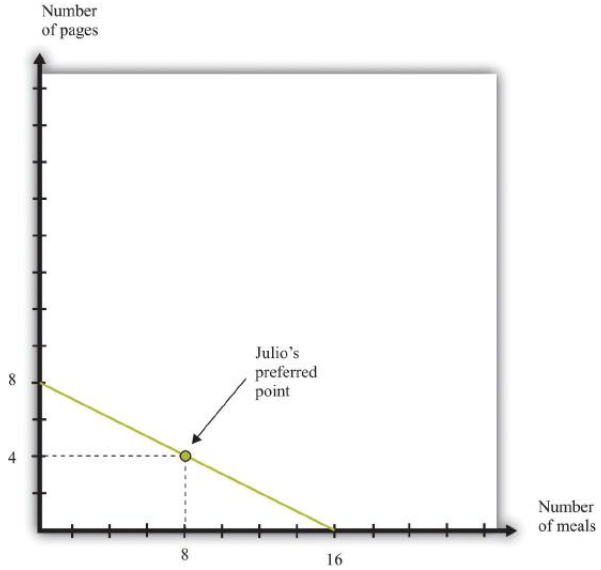


Figure 6.5.4: Julio's Allocation of Time to Cooking Meals and Producing Web Pages. If Julio likes to consume pages and meals in fixed proportions (2 vegan dishes for every web page), he will allocate his time to achieve the point shown in the figure.

The Production Possibilities Frontier with Two People

If this were the end of the story, we would not have seen the advertisements on craigslist to trade vegan meals or web pages. Things become more interesting and somewhat more realistic when we add another person to our economy.

Hannah has production possibilities that are summarized in Table 6.5.3 "Production Possibilities for Julio and Hannah". She can produce 1 vegan meal in an hour or produce 1.5 web pages in an hour. Hannah, like Julio, has 8 hours per day to allocate to production activities. In Table 6.5.3 "Production Possibilities for Julio and Hannah" we have also included their respective opportunity costs of producing web pages and vegan meals.

Table 6.5.3: Production Possibilities for Julio and Hannah

	Vegan Meals per Hour	Web Pages per Hour	Opportunity Cost of Vegan Meals (in Web Pages)	Opportunity Cost of Web Pages (in Vegan Meals)
Julio	2	1	1/2	2
Hannah	1	1.5	3/2	2/3

Table 6.5.3 "Production Possibilities for Julio and Hannah" reveals that Hannah is more productive than Julio in the production of web pages. By contrast, Julio is more productive in vegan meals. Hannah's production possibilities frontier is illustrated in Figure 6.5.5. It is steeper than Julio's because the opportunity cost of vegan meals is higher for Hannah than it is for Julio.

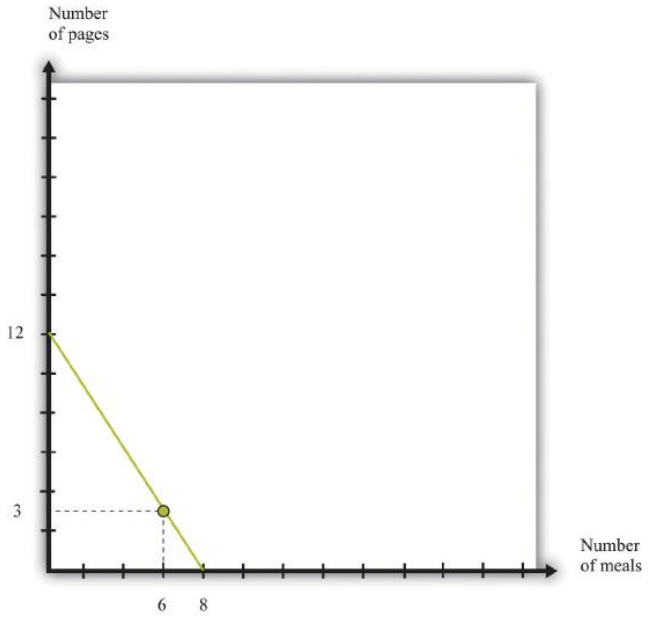


Figure 6.5.5: Hannah's Production Possibilities Frontier. Hannah's production possibilities frontier is steeper than Julio's.

Economists use the ideas of **absolute advantage** and **comparative advantage** to compare the productive abilities of Hannah and Julio.

Toolkit: Section 31.13

Comparative advantage and absolute advantage are used to compare the productivity of people (or firms or countries) in the production of a good or a service. A person has an absolute advantage in the production of a good if that person can produce more of that good in a unit of time than another person. A person has a comparative advantage in the production of one good if the opportunity cost, measured by the lost output of the other good, is lower for that person than for another person.

Both absolute and comparative advantage are relative concepts because they compare two people. In the case of absolute advantage, we compare the productivity of two people for a given good. In the case of comparative advantage, we compare two people and two goods because opportunity cost is defined across two goods (web pages and vegan meals in our example). Comparing Hannah and Julio, we see that Hannah has an absolute advantage in the production of web pages. She is better at producing web pages than Julio. By contrast, Julio has an absolute advantage in the production of vegan meals. Therefore, it is not surprising that Hannah also has a comparative advantage in the production of web pages—the opportunity cost of web pages is lower for her than it is for Julio—whereas Julio has a comparative advantage in vegan meals.

It is entirely possible for one person to have an absolute advantage in the production of both goods. For example, were Hannah's productivity in both activities to double, she would be better both at both web design and vegan meals. However, her opportunity cost of web pages would be unchanged. Julio would still have a comparative advantage in vegan meals. In general, one person always has a comparative advantage in one activity, while the other person has a comparative advantage in the other activity (the only exception is the case where the two individuals have exactly the same opportunity costs).

Suppose that Hannah's tastes for vegan meals and web pages are the same as Julio's: like Julio, Hannah wants to consume 2 meals for every web page. Acting alone, she would work for 2 hours in web design and spend 6 hours cooking, ending up with 3 web pages and 6 meals. But there is something very odd going on here. Julio, who is good at preparing meals, spends half his time on web design. Hannah, who is good at web design, spends three-quarters of her time cooking. Each has to spend a lot of time doing an activity at which he or she is unproductive.

This is where we see the possibility of gains from trade. Imagine that Julio and Hannah join together and become a team. What is their *joint* production possibilities frontier? If both Julio and Hannah devote their 8 hours of time to the production of web pages, then the economy can produce 20 web pages (8 from Julio and 12 from Hannah). At the other extreme, if both Julio and Hannah devote their 8 hours to cooking vegan meals, then the economy can produce 24 meals (16 from Julio and 8 from Hannah). These two points, which represent specialization of their two-person economy in one good, are indicated by point A and point D in Figure 6.5.6.

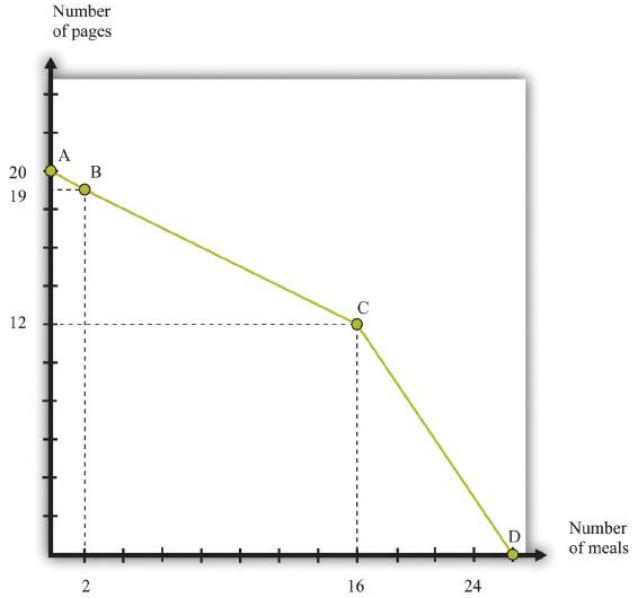


Figure 6.5.6: Julio and Hannah's Joint Production Possibilities Frontier. Julio and Hannah's joint production possibilities frontier.

To fill in the rest of their joint production possibilities frontier, start from the vertical intercept, where 20 web pages are being produced. Suppose Julio and Hannah jointly decide that they would prefer to give up 1 web page to have some vegan meals. If Hannah were to produce only 11 web pages, she could free up 2/3 of an hour for vegan meals. She would produce 2/3 of a meal. Conversely, if Julio produced 1 fewer web page, that would free an hour of his time (because he is less efficient than Hannah at web design), and he could create 2 vegan meals. Evidently, it makes much more sense for Julio to shift from web design to cooking (see point B in Figure 6.5.6).

The most efficient way to substitute web design production for vegan meals production, starting at point A, is to have Julio switch from producing web pages to producing vegan meals. Julio should switch because he has a comparative advantage in cooking. As we move along the production possibilities frontier from point A to point B to point C, Julio continues to substitute from web pages to meals. For this segment, the slope of the production possibilities frontier is −1/2, which is Julio's opportunity cost of web pages.

At point C, both individuals are completely specialized. Julio spends all 8 hours on vegan meals and produces 16 meals. Hannah spends all 8 hours on web design and produces 12 web pages. If they would like to have still more vegan meals, it is necessary for Hannah to start producing that service. Because she is less efficient at cooking and more efficient at web design than Julio, the cost of extra vegan meals increases. Between point A and point C, the cost of vegan meals was 1/2 a webpage. Between point C and point D, the cost of a unit of vegan meals is 1.5 web pages. The production possibilities frontier becomes much steeper. If you look

carefully at Figure 6.5.4, Figure 6.5.5, and Figure 6.5.6, the joint production possibilities frontier is composed of the two individual frontiers joined together at point C.

Gains from Trade Once Again

Figure 6.5.7 again shows the production possibilities frontier for the Julio-Hannah team: all the combinations of web pages and vegan meals that they can produce in one day, using the technologies available to them.

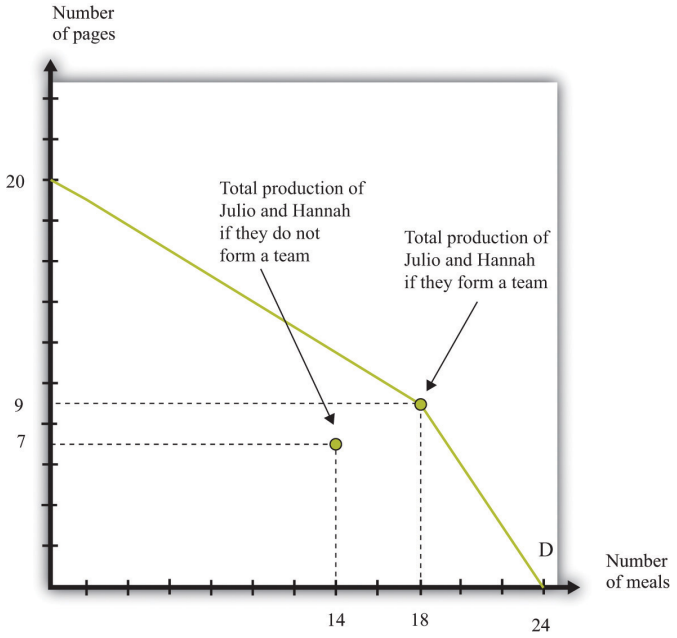


Figure 6.5.7: Julio and Hannah's Preferred Point. Julio and Hannah can benefit from joining forces.

Julio and Hannah have a great deal to gain by joining forces. We know that Julio, acting alone, produces 4 web pages and 8 vegan meals. Hannah, acting alone, produces 3 web pages and 6 vegan meals. The joint total is 7 web pages and 14 vegan meals. In Figure 6.5.7, we have labeled this as "Total production of Julio and Hannah if they do not form a team." But look at what they can achieve if they work together: they can produce 9 web pages and 18 vegan meals.

Evidently they both can be better off when they work together. For example, each could get an additional web page and 2 vegan meals. Julio could have 5 web pages and 10 vegan meals (instead of 4 and 8, respectively), and Hannah could have 4 web pages and 8 vegan meals (instead of 3 and 6, respectively).

How do they do this? Julio specializes completely in vegan meals. He spends all 8 hours of his day cooking, producing 16 vegan meals. Hannah, meanwhile, gets to spend most of her time doing what she does best: designing web pages. She spends 6 hours on web design, producing 9 web pages, and 2 hours cooking, producing 2 vegan meals.

The key to this improvement is that we are no longer requiring that Julio and Hannah consume only what they can individually produce. Instead, they can produce according to their comparative advantage. They each specialize in the production of the good that they produce best and then trade to get a consumption bundle that they are happy with. The gains from trade come from the ability to specialize.

It is exactly such gains from trade that people are looking for when they place advertisements on craigslist. For example, the first ad we quoted was from someone with a comparative advantage in fixing furniture looking to trade with someone who had a comparative advantage in dental work. Comparative advantage is one of the most fundamental reasons why people trade, and sites like craigslist allow people to benefit from trade. Of course, in modern economies, most trade does not occur through individual barter; stores, wholesalers, and other intermediaries mediate trade. Although there are many *mechanisms* for trade, comparative advantage is a key *motivation* for trade.

Specialization and the History of the World

Economics famously teaches us that there is no such thing as a free lunch: everything has an opportunity cost. Paradoxically, economics also teaches us the secret of how we can make everyone better off than before simply by allowing them to trade—and if

that isn't a free lunch, then what is?

This idea is also the story of why the world is so much richer today than it was 100 years, 1,000 years, or 10,000 years ago. The ability to specialize and trade is a key to prosperity. In the modern world, almost everybody is highly specialized in their production, carrying out a very small number of very narrow tasks. Specialization permits people to become skilled and efficient workers. (This is true, by the way, even if people have similar innate abilities. People with identical abilities will still usually be more efficient at producing one good rather than two.) Trade means that even though people specialize in production, they can still generalize in consumption. At least in the developed world, we enjoy lives of luxury that were unimaginable even a couple of centuries ago. This luxury would be impossible without the ability to specialize and trade.

The story of Julio and Hannah is therefore much more than a textbook exercise. One of the first steps on the ladder of human progress was the shift from a world where people looked after themselves to a world where people started producing in hunter-gatherer teams. Humans figured out that they could be more productive if some people hunted and others gathered. They also started to learn the benefits of team production—hunting was more efficient if a group of hunters worked together, encircling the prey so it could not escape. Such hunting teams are perhaps the first example of something that looks like a firm: a group of individuals engaged jointly in production.

Production Possibilities Frontier for a Country

Before we finish with this story, let us try to get a sense of how we can expand it to an entire economy. We begin by adding a third individual to our story: Sergio. Sergio is less efficient than both Julio and Hannah. He has no absolute advantage in anything; he is no better at web design than Hannah, and he is no better preparing vegan meals than Julio. Remarkably, Julio and Hannah will still want to trade with him.

Table 6.5.4: Production Possibilities for Julio, Hannah, and Sergio

	Vegan Meals per Hour	Web Pages per Hour	Opportunity Cost of Web Pages (in Vegan Meals)	Opportunity Cost of Vegan Meals (in Web Pages)
Julio	2	1	1/2	2
Hannah	1	1.5	3/2	2/3
Sergio	1	1	1	1

We begin by constructing the production possibilities frontier for these three individuals. The logic is the same as before. Start from the position where the economy produces nothing but web pages (see Figure 6.5.8, point A). Together, Julio, Hannah, and Sergio can produce a total of 28 web pages in a day. Then we first shift Julio to cooking because he has the lowest opportunity cost of that activity. As before, the slope of this first part of the frontier is −1/2, up to point B. At point B, Julio is producing nothing but vegan meals, and Hannah and Sergio are devoting all of their time to producing web pages. At point B, the economy produces 20 web pages and 16 vegan meals.

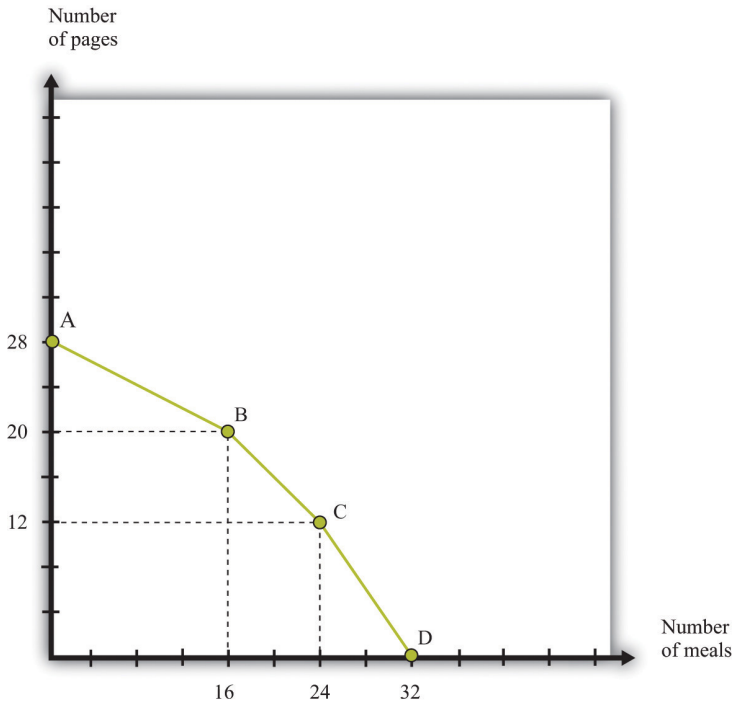


Figure 6.5.8: The Production Possibilities Frontier with Three People. Here we show a production possibilities frontier with three individuals. Notice that it is smoother than the production possibilities frontier for two people.

If they want more vegan meals, who should switch next? The answer is Sergio because his opportunity cost of vegan meals is lower than Hannah's. As Sergio starts shifting from web design to vegan meals, the frontier has slope −1, and we move from point B to point C. At point C, Sergio and Julio cook meals, and Hannah produces web pages. The number of web pages produced is 12, and there are 24 vegan meals. Finally, the last segment of the frontier has slope −3/2, as Hannah also shifts from web design to vegan meals. At point D, they all cook, with total production equal to 32 meals.

Let us suppose that Sergio, like the others, consumes in the ratio of 2 vegan meals for every web page. If the economy is at point C, the economy can produce 12 web pages and 24 vegan meals. Earlier, we saw that Hannah and Julio could together produce 9 web pages and 18 vegan meals, so bringing in Sergio allows for an extra 3 web pages and 6 vegan meals, which is more than Sergio could produce on his own. Even though Sergio is less efficient than both Julio and Hannah, there are still some gains from trade. The easiest way to see this is to note that it would take Sergio 9 hours to produce 3 web pages and 6 vegan meals. In 8 hours he could produce almost 3 web pages and over 5 vegan meals.

Where do the gains from trade come from? They come from the fact that, relative to Hannah, Sergio has a comparative advantage in vegan meals. Previously, Hannah was devoting some of her time to vegan meals, which meant she had to divert time from web design. This was costly because she is good at web design. By letting Sergio do the vegan meals, Hannah can specialize in what she does best. The end result is that there are extra web pages and vegan meals for them all to share.

Comparing Figure 6.5.7 and Figure 6.5.8, you can see that the frontier becomes "smoother" when we add Sergio to the picture. Now imagine that we add more and more people to the economy, each with different technologies, and then construct the frontier in the same way. We would get a smoother and smoother production possibilities frontier. In the end, we might end up with something like Figure 6.5.9.

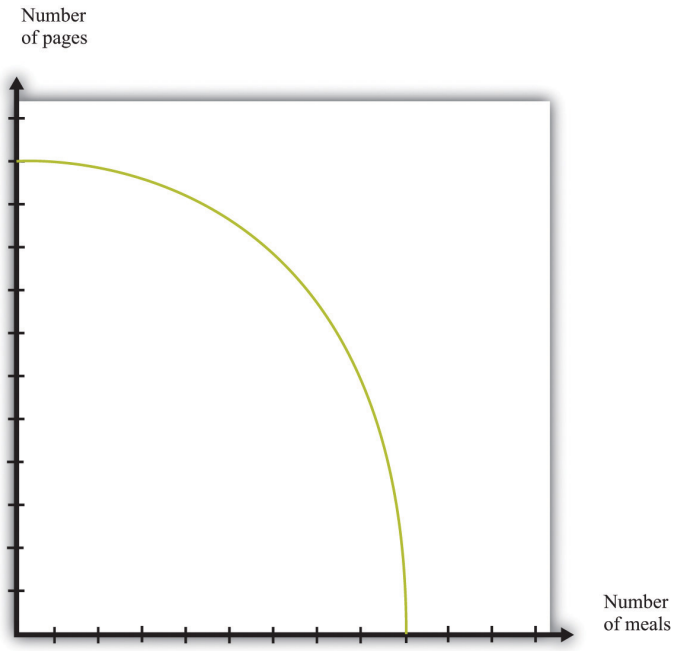


Figure 6.5.9: The Production Possibilities Frontier with a Large Number of People. As we add more and more people to the economy, the production possibilities frontier will become smoother.

It is easy enough to imagine that Julio, Hannah, and Sergio could all get together, agree to produce according to the principle of comparative advantage, and then share the goods that they have produced in a way that makes them all better off than they would be individually. Exactly how the goods would be shared would involve some kind of negotiation and bargaining among them. Once we imagine an economy with a large number of people in it, however, it is less clear how they would divide up the goods after they were produced. And that brings us full circle in the chapter. It is not enough that potential gains from trade exist. There must also be mechanisms, such as auctions and markets, that allow people to come together and realize these gains from trade.

Key Takeaways

- Gains in trade partly come from the fact that individuals specialize in production and generalize in consumption.
- The efficient way to organize production is by looking at comparative advantage.
- Gains to trade emerge when individuals produce according to comparative advantage and then trade goods and services with one another.

check your understanding

1. Fill in the missing values in Table 6.5.2 "Julio’s Production Possibilities".

This page titled 6.5: Production Possibilities is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

6.6: End-of-Chapter Material 6.6: End-of-Chapter MaterialEdit Update titlecancel

In Conclusion

This chapter is our first look at how individuals exchange. We have emphasized two points:

1. *How* individuals trade. You have seen that some very familiar things, such as eBay and craigslist, provide mechanisms to facilitate trade.
2. *Why* individuals trade. These gains may simply arise from differences in how people value items, as in Chapter 6, Section 6.3. Or, as in Chapter 6, Section 6.6, these gains may reflect the fact that people differ in their abilities to produce different goods and services.

In reality, individuals differ across these two dimensions and more. Chapter 10 explores two further reasons for trade: differences in information and differences in attitudes toward risk.

Auctions such as eBay, newspaper classified advertisements, and sites such as craigslist are all means by which *individuals* in an economy can trade with one another. Of course, these are not the only forms of trade. Our discussion, by design, has ignored other common forms of trade in the economy, such as individuals buying goods and services from a firm (perhaps through a retailer) and individuals selling their labor services to firms.Such exchanges are discussed in Chapter 7 and Chapter 8.

The biggest insight you should take away from this chapter is the fact that exchange is a means of creating value. When a seller sells a good or a service to a buyer, there is a presumption that both become better off. We have such a presumption because people enter into trades voluntarily: nobody forces a buyer to buy; nobody forces a seller to sell. The fact that voluntary exchange creates value is one of the most powerful ideas in economics.

Key Links

- eBay: www.ebay.com
- craigslist: http://www.craigslist.org

exercises

1. Would you expect to get an item for less by buying it through craigslist rather than a regular store? If so, why?
2. (Advanced) Suppose the government imposed a tax on trading through craigslist, so the seller had to pay 5 percent of the price to the government. What might be the impact on trade in this market? What might happen to prices?
3. If you are a seller on craigslist, what would be the cost of setting a very high price?
4. If the price of an item traded increases, can both the surplus to the buyer and the surplus of the seller increase simultaneously?
5. If the owner of a car values it at $5,000 and there is a prospective buyer who is willing to pay $7,000 for that car, what does efficiency dictate about the price the buyer should pay for the car?
6. What are the differences between buying an item on eBay and buying that same item on craigslist?
7. In what settings do you have to be aware of the winner's curse?
8. In what way does a double oral auction differ from craigslist? From eBay?
9. If more people come into an auction, how should that affect your bidding in a winner's curse situation? Should you bid more or less? Why?
10. Suppose that instead of producing 2 vegan meals each hour, Julio can produce 3 vegan meals each hour. Draw his production possibilities frontier. What is his opportunity cost of web pages in terms of vegan meals with this alternative technology?
11. Suppose that Julio can produce 3 vegan meals each hour but requires 2 hours to design a web page. Draw his production possibilities frontier.
12. If Julio had a choice between the technology in Table 6.5.1 "Julio's Production Ability" and the one described in question 10, which would he prefer? Explain why.
13. What would the production possibilities frontier look like if, starting from point A in Figure 6.6.1, we first shifted Hannah rather than Julio to vegan meal production?
14. Show how the production possibilities frontier shifts if Hannah becomes more productive in producing web pages.

15. (Advanced) Suppose that both Julio and Hannah like each of the goods in a ratio of 6 vegan meals to 1 web page. Show that there are still gains to trade using the technologies described in Table 6.3.1 "Valuations of Different Bidders in a Winner's Curse Auction".
16. Explain what it means to "specialize in production and generalize in consumption." How many jobs do people usually have at a point in time? How many items does a food shopper usually have in his or her basket at the store?
17. Explain the connection between opportunity cost and comparative advantage.

Economics Detective

1. Find an auction to buy or sell the following items: a house, a car, a government bond, and licenses for the electromagnetic spectrum. What do you have to do to become a bidder at one of these auctions? How is the auction conducted?
2. Suppose you want to purchase a painting at Sotheby's, a famous English auction house. How would you do so? How would the auction operate? In what ways would it differ from buying art on eBay or craigslist?

Spreadsheet Exercise

1. (Advanced) Create a spreadsheet to input data like that in the first two columns of Table 6.5.3 "Production Possibilities for Julio and Hannah". Suppose there are two people who can produce two goods. Enter into the spreadsheet how much of each good they can produce in an hour.
    1. Calculate the opportunity costs, as in the last two columns of Table 6.5.3 "Production Possibilities for Julio and Hannah".
    2. Assuming each has 8 hours a day to work, use the spreadsheet to calculate the total amount of each good each individual could produce if they produced only that good.
    3. Use this information to graph the production possibilities frontier for each person.
    4. Use this information to graph the production possibilities frontier for the two people combined.
    5. As you input different levels of output per hour per person, watch how these graphs change.
    6. Where do you see comparative advantage coming into play?

This page titled 6.6: End-of-Chapter Material is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

7: Where Do Prices Come From?

7.1: The Price of a Pill
7.2: The Goal of a Firm
7.3: The Revenues of a Firm
7.4: The Costs of a Firm
7.5: Markup Pricing- Combining Marginal Revenue and Marginal Cost
7.6: The Supply Curve of a Competitive Firm
7.7: End-of-Chapter Material

This page titled 7: Where Do Prices Come From? is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

7.1: The Price of a Pill 7.1: The Price of a PillEdit Update titlecancel

If you walk down the aisles of a supermarket, you will see thousands of different goods for sale. Each one will have a price displayed, telling you how much money you must give up if you want the good in question. On the Internet, you can find out how much it would cost you to stay in a hotel in Lima, Peru, or how much you would have to pay to rent a four-wheel drive vehicle in Nairobi, Kenya. On your television every evening, you can see the price that you would have to pay to buy a share of Microsoft Corporation or other companies.

Prices don't appear by magic. Every price posted in the supermarket or on the Internet is the result of a decision made by one or more individuals. In the future, you may find yourself trying to make exactly such a decision. Many students of economics have jobs in the marketing departments of firms or work for consulting companies that provide advice on what prices firms should charge. To learn about how managers make such decisions, we look at a real-life pricing decision.

In 2003, a major pharmaceutical company was evaluating the performance of one of its most important drugs—a medication for treating high blood pressure—in a Southeast Asian country. (For reasons of confidentiality, we do not reveal the name of the company or the country; other than simplifying the numbers slightly, the story is true.) Its product was known as one of the best in the market and was being sold for $0.50 per pill. The company had good market share and income in the country. There was one major competing drug in the market that was selling at a higher price and a few less important drugs.

In pharmaceutical companies, one individual often leads the team for each major drug that the company sells. In this company, the head of the product team—we will call her Ellie—was happy with the performance of the drug. Nonetheless, she wondered whether her company could make higher profits by setting a higher or lower price. In many countries, the prices of pharmaceutical products are heavily regulated. In this particular country, however, pharmaceutical companies were largely free to set whatever price they chose. Together with her team, therefore, Ellie decided to review the pricing strategy for her product. In this chapter, we therefore tackle the following question:

How should a firm set its price?

Road Map

Price-setting in retail markets typically takes the form of a take-it-or-leave-it offer. The seller posts a price, and prospective customers either buy or don't buy at that price. The prices you encounter every day in a supermarket, a coffee shop, or a fast-food restaurant, for example, are all take-it-or-leave-it offers that the retailer makes to you and other customers. Chapter 6 has more discussion.

In this chapter, we put you in the place of a marketing manager who has been given the job of determining the price that a firm should charge for its product. We first discuss the goals of this manager: what is she trying to achieve? We then show what information she needs to make a good decision. Finally, we derive some principles that allow her to set the right price. The chapter is built around two ideas:This chapter and Chapter 6 are linked because they are both about mechanisms that allocate goods and services. In Chapter 6 we explain how eBay, craigslist, and newspapers are ways in which individuals exchange goods and services. In this chapter, we study how goods and services are allocated from firms to households. At the end of this chapter, we show that the supply-and-demand framework introduced in Chapter 6 is also a useful framework *when the same product is produced by a large number of firms*. In particular, we show that our ideas about pricing also allow us to understand the foundations of supply.

1. **The law of demand.** Each firm faces a demand curve for its product. This demand curve obeys the law of demand: if a firm sets a higher price, it must be willing to sell a smaller quantity; if a firm wishes to sell a larger quantity, it must set a lower price.
2. **Profit accounting.** Firms earn income from selling their goods and services, but they also incur costs from producing those goods and services. These costs include the costs of raw materials, the wages paid to the firm's workers, and so on. The difference between a firm's revenues and its costs is the firm's profits.

The choice of price, via the demand curve, determines the amount of output a firm sells. The amount of output determines a firm's revenues and costs. Together, revenues and costs determine the profits of a firm.

This page titled 7.1: The Price of a Pill is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

7.2: The Goal of a Firm 7.2: The Goal of a FirmEdit Update titlecancel

Learning Objective

1. What is the goal of a firm?

Firms devote substantial resources to their decisions about pricing. Large firms often have individuals or even entire departments whose main job is to make pricing decisions. Consulting firms specialize in providing advice to firms about the prices that they should charge. Some companies, such as airlines, have dedicated software to help them make these decisions. It isn't hard to understand why firms pay so much attention to the prices they charge. More than anything else, price determines the profits that a firm earns.

Economists are prone to talk about the decisions and objectives of a firm, and we often use the same shorthand. A firm, though, is just a legal creation—a collection of individuals who use some kind of technology. A firm takes labor, raw materials, and other inputs and turns them into products that people want to buy. Some of the people in a firm—the managers—decide how many workers it should hire, what prices it should set, and so on.

To understand pricing, we begin with the goal of a firm (that is, its managers). If a firm's managers are doing their jobs well, they should be making decisions to serve the interests of the owners of that firm. The owners of a firm are its shareholders. If you buy a share in a firm, then you own a fraction (your *share*) of the firm, which gives you the right to a fraction of the firm's earnings. Shareholders, for the most part, have one reason for buying and owning shares: to earn income. So the managers, if they are doing their jobs well, want a firm to make as much money as possible. We need to be careful, though. What matters is not the total amount of money received by a firm, but how much is available to be distributed to its owners. The owners of a firm hope to earn as high a return as possible on their shares.

Toolkit: Section 31.15

The money that is available for distribution to the shareholders of a firm is called a firm's profits. A firm pays money for raw materials, energy, and other supplies, and it pays wages to its workers. These expenses are a firm's costs of production. When it sells the product(s) it has produced, a firm earns revenues. Accountants analyze these revenues and costs in more detail, but in the end all the monies that flow in and out of a firm can be classified as either revenues or costs. Thus

$$profits = revenues - costs.$$

Consider, then, a marketing manager who wants to set the best price for a product—such as Ellie choosing the price for her company's blood pressure medication. She wants to find the price that will yield the most profits to her company. In an ideal world, a marketing manager might have access to a spreadsheet table, such as Figure 7.2.1, which displays a firm's monthly profits for different possible prices that it might set. Then Ellie's job would be easy: she would just have to look at the table, find the cell in column B with the highest number, and set the corresponding price. In this case, she would set a price of $15.

Monthly Profits.xls

	A	B	C	D	E	F
1	Price ($)	Monthly Profit ('000 $)				
2						
3						
4	1	-6840				
5	2	-5760				
6	3	-4760				
7	4	-3840				
8	5	-3000				
9	6	-2240				
10	7	-1560				
11	8	-960				
12	9	-440				
13	10	0				
14	11	360				
15	12	640				
16	13	840				
17	14	960				
18	15	1000				
19	16	960				
20	17	840				
21	18	640				
22	19	360				
23	20	0				
24	21	-440				
25	22	-960				
26	23	-1560				
27	24	-2240				
28	25	-3000				

Figure 7.2.1: A Spreadsheet That Would Make Pricing Decisions Easy

But the reality of business is different. It is very difficult and expensive—perhaps even impossible—to gather information such as that in Figure 7.2.1. You might imagine that a firm could experiment, trying different prices and seeing what **profits** it earned. Unfortunately, this would be very costly because most of the time a firm would earn much lower profits than it could. Experimenting might even generate losses. For example, suppose that, one September, Ellie chose to try a price of $2 per pill. The firm would lose nearly $6 million—the equivalent of about six months' profits even at the very best price. Ellie would rapidly find herself looking for another job.

It is clear that trial and error—choosing different prices at random and seeing how much profit you get—could lead to costly mistakes, and there is no guarantee that you would ever find the best price. By adding some structure to a trial-and-error process, though, there is a simple strategy for finding the best price: begin by slightly raising the firm's price. If profits increase, then you are on the right track. Keep raising the price, little by little, until profits stop increasing. On the other hand, if profits decrease when you raise the price, then you should try lowering the price instead. If profits increase, then keep lowering the price little by little.

Figure 7.2.2 shows how a change in price translates into a change in profits. A change in a firm's price leads to a change in the quantity demanded. As a result, the **revenues** and **costs** of a firm change, as do its profits. Figure 7.2.3 shows the profits a firm will earn at different prices. Our pricing strategy simply says the following. You are trying to get to the highest point of the profit hill in Figure 7.2.3, and you will get there eventually if you always walk uphill. At the very top of the hill, the change in profits is zero.

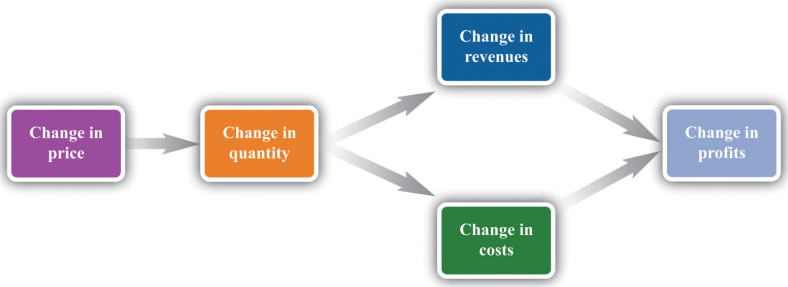


Figure 7.2.2: A Change in Price Leads to a Change in Profits. If a firm changes its price, then there will be a change in demand. This then leads to changes in revenues and costs, which changes in the profits of a firm.

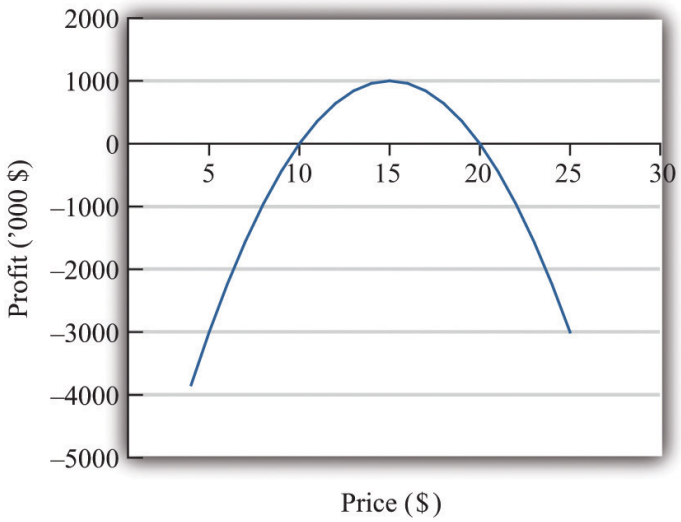


Figure 7.2.3: The Profits of a Firm

We could end the chapter right here. But we want to dig deeper and uncover some principles that tell us more about how pricing works. Then we can learn what information Ellie and other managers like her need to make better pricing decisions—and how they can make these decisions effectively. Our starting point is our earlier observation that

$$profits = revenues - costs.$$

key takeaway

- The objective of a firm is to maximize its profits, defined as revenues minus costs.

check your understanding

1. If the manager of a firm chose a price to maximize sales, what would that price be? What would profits be at that price?
2. Explain in words why the profit function has the shape shown in Figure 7.2.3.

This page titled 7.2: The Goal of a Firm is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

- **8.6: Beyond Perfect Competition** by Anonymous is licensed CC BY-NC-SA 3.0. Original source: https://2012books.lardbucket.org/books/theory-and-applications-of-economics.
- **Current page** by Anonymous is licensed CC BY-NC-SA 3.0. Original source: https://2012books.lardbucket.org/books/theory-and-applications-of-economics.

7.3: The Revenues of a Firm 7.3: The Revenues of a FirmEdit Update titlecancel

Learning Objectives

1. What is the demand curve faced by a firm?
2. What is the elasticity of demand? How is it calculated?
3. What is marginal revenue?

A firm's revenues are the money that it earns from selling its product. Revenues equal the number of units that a firm sells times the price at which it sells each unit:

$$revenues \ = \ price \ \times \ quantity.$$

For example, think about a music store selling CDs. Suppose that the firm sells 25,000 CDs in a month at $15 each. Then its total monthly revenues are as follows:

$$revenues \ = \ 15 \times 25,000 = \$375,000.$$

There are two ways in which firms can obtain higher revenues: sell more products or sell at a higher price. So if a firm wants to make a lot of revenue, it should sell a lot of its product at a high price. Then again, you probably do not need to study economics to figure that out. The problem for a manager is that her ability to sell a product is limited by what the market will bear. Typically, we expect that if she sets a higher price, she will not be able to sell as much of the product:

Equivalently, if she wants to sell a larger quantity of product, she will need to drop the price:

This is the law of demand in operation (Figure 7.3.1).

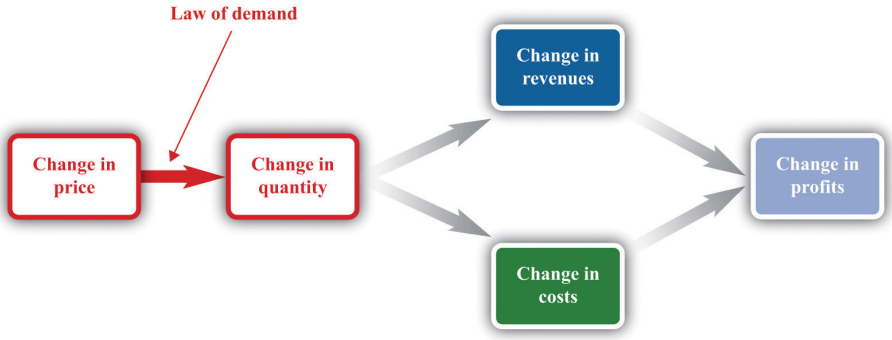


Figure 7.3.1: A Change in the Price Leads to a Change in Demand. An increase in price leads to a decrease in demand. A decrease in price leads to an increase in demand.

The Demand Curve Facing a Firm

There will typically be more than one firm that serves a market. This means that the overall demand for a product is divided among the different firms in the market. We have said nothing yet about the kind of "market structure" in which a firm is operating—for example, does it have a lot of competitors or only a few competitors? Without delving into details, we cannot know exactly how the market demand curve will be divided among the firms in the market. Fortunately, we can put this problem aside—at least for this chapter.We look at market structure in Chapter 15. For the moment, it is enough just to know that each firm faces a demand curve for its own product.

When the price of a product increases, individual customers are less likely to think it is good value and are more likely to spend their income on other things instead. As a result—for almost all products—a higher price leads to lower sales.

Toolkit: Section 31.15

The demand curve facing a firm tells us the price that a firm can expect to receive for any given amount of output that it brings to market or the amount it can expect to sell for any price that it chooses to set. It represents the market opportunities of the firm.

An example of such a demand curve is

$$quantity\ demanded\ =\ 100 - (5 \times price).$$

Table 7.3.1 "Example of the Demand Curve Faced by a Firm" calculates the quantity associated with different prices. For example, with this demand curve, if a manager sets the price at $10, the firm will sell 50 units because 100 − (5 × 10) = 50. If a manager sets the price at $16, the firm will sell only 20 units: 100 − (5 × 16) = 20. For every $1 increase in the price, output decreases by 5 units. (We have chosen a demand curve with numbers that are easy to work with. If you think that this makes the numbers unrealistically small, think of the quantity as being measured in, say, thousands of units, so a quantity of 3 in this equation means that the firm is selling 3,000 units. Our analysis would be unchanged.)

Table 7.3.1: Example of the Demand Curve Faced by a Firm

Price ($)	Quantity
0	100
2	90
4	80
6	70
8	60
10	50
12	40
14	30
16	20
18	10
20	0

Equivalently, we could think about a manager choosing the quantity that the firm should produce, in which case she would have to accept the price implied by the demand curve. To write the demand curve this way, first divide both sides of the equation by 5 to obtain

Now add “price” to each side and subtract “ ” from each side:

For example, if the manager wants to sell 70 units, she will need a price of $6 (because 20 − 70/5 = 6). For every unit increase in quantity, the price decreases by 20 cents.

Either way of looking at the demand curve is perfectly correct. Figure 7.3.2 shows the demand curve in these two ways. Look carefully at the two parts of this figure and convince yourself that they are really the same—all we have done is switch the axes.

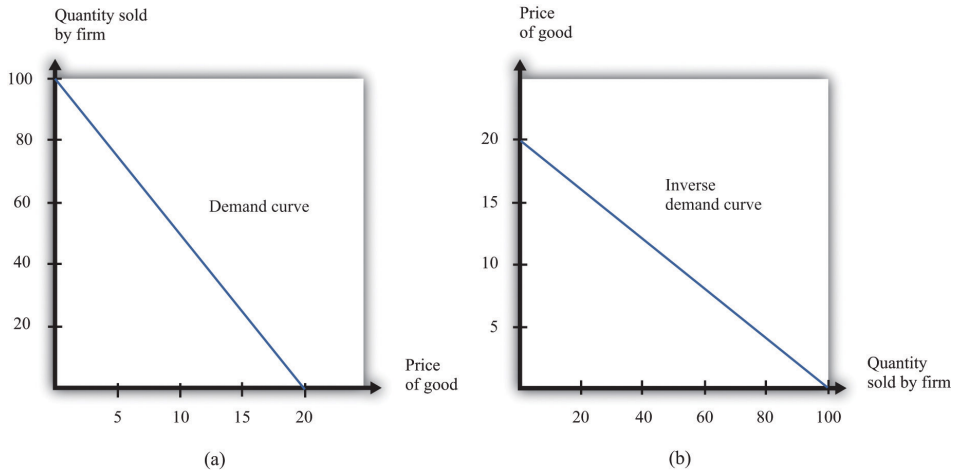


Figure 7.3.2: Two Views of the Demand Curve. There are two ways that we can draw a ***demand curve***, *both of which are perfectly correct. (a) The demand curve has price on the horizontal axis and quantity demanded on the vertical axis (b). The demand curve has price on the vertical axis, which is how we normally draw the demand curve in economics.*

The firm faces a trade-off: it can set a high price, such as $18, but it will be able to sell only a relatively small quantity (10). Alternatively, the firm can sell a large quantity (for example, 80), but only if it is willing to accept a low price ($4). The hard choice embodied in the demand curve is perhaps the most fundamental trade-off in the world of business. Of course, if the firm sets its price too high, it won't sell anything at all. The **choke price** is the price above which no units of the good will be sold. In our example, the choke price is $20; look at the vertical axis in part (b) of Figure 7.3.2.A small mathematical technicality: the equation for the demand curve applies only if both the price and the quantity are nonnegative. At any price greater than the choke price, the quantity demanded is zero, so the demand curve runs along the vertical axis. A negative price would mean a firm was paying consumers to take the product away.

Every firm in the economy faces some kind of demand curve. Knowing the demand for your product is one of the most fundamental necessities of successful business. We therefore turn next to the problem Ellie learned about the demand curve for her company's drug.

The Elasticity of Demand: How Price Sensitive Are Consumers?

Marketing managers understand the law of demand. They know that if they set a higher price, they can expect to sell less output. But this is not enough information for good decision making. Managers need to know whether their customers' demand is very sensitive or relatively insensitive to changes in the price. Put differently, they need to know if the demand curve is steep (a change in price will lead to a small change in output) or flat (a change in price will lead to a big change in output). We measure this sensitivity by the **own-price elasticity of demand**.

Toolkit: Section 31.2

The own-price elasticity of demand (often simply called the *elasticity of demand*) measures the response of quantity demanded of a good to a change in the price of that good. Formally, it is the percentage change in the quantity demanded divided by the percentage change in the price:

When price increases (the change in the price is positive), quantity decreases (the change in the quantity is negative). The price elasticity of demand is a negative number. It is easy to get confused with negative numbers, so we instead use

which is always a positive number.

- If −(elasticity of demand) is a large number, then quantity demanded is sensitive to price: increases in price lead to big decreases in demand.
- If −(elasticity of demand) is a small number, then quantity demanded is insensitive to price: increases in price lead to small decreases in demand.

Throughout the remainder of this chapter, you will often see −(elasticity of demand). Just remember that this expression always refers to a positive number.

Calculating the Elasticity of Demand: An Example

Go back to our earlier example:

$$quantity\ demanded\ =\ 100 - 5 \times price.$$

Suppose a firm sets a price of $15 and sells 25 units. What is the elasticity of demand if we think of a change in price from $15 to $14.80? In this case, the change in the price is −0.2, and the change in the quantity is 1. Thus we calculate the elasticity of demand as follows:

1. The percentage change in the quantity is (4 percent).
2. The percentage change in the price is (approximately −1.3 percent).
3. −(elasticity of demand) is

The interpretation of this elasticity is as follows: when price decreases by 1 percent, quantity demanded increases by 3 percent. This is illustrated in Figure 7.3.3.

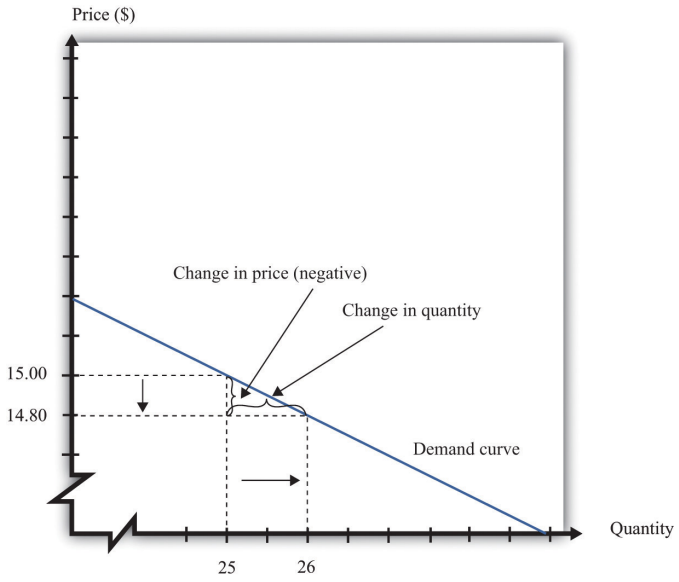


Figure 7.3.3: The Elasticity of Demand. When the price is decreased from $15.00 to $14.80, sales increase from 25 to 26. The percentage change in price is −1.3 percent. The percentage change in the quantity sold is 4. So −(elasticity of demand) is 3.

One very useful feature of the elasticity of demand is that it does not change when the number of units changes. Suppose that instead of measuring prices in dollars, we measure them in cents. In that case our demand curve becomes

$$quantity\ demanded\ = 100 - 500 \times price.$$

Make sure you understand that this is exactly the same demand curve as before. Here the slope of the demand curve is −500 instead of −5. Looking back at the formula for elasticity, you see that the change in the price is 100 times greater, but the price itself is 100 times greater as well. The percentage change is unaffected, as is elasticity.

Market Power

The elasticity of demand is very useful because it is a measure of the **market power** that a firm possesses. In some cases, some firms produce a good that consumers want very much—a good in which few substitutes are available. For example, De Beers controls much of the world's market for diamonds, and other firms are not easily able to provide substitutes. Thus the demand for De Beers' diamonds tends to be insensitive to price. We say that De Beers has a lot of market power. By contrast, a fast-food restaurant in a mall food court possesses very little market power: if the fast-food Chinese restaurant were to try to charge significantly higher prices, most of its potential customers would choose to go to the other Chinese restaurant down the aisle or even to eat sushi, pizza, or burritos instead.

Ellie's company had significant market power. There were a relatively small number of drugs available in the country to treat high blood pressure, and not all drugs were identical in terms of their efficacy and side effects. Some doctors were loyal to her product

and would almost always prescribe it. Some doctors were not very well informed about the price because doctors don't pay for the medication. For all these reasons, Ellie had reason to suspect that the demand for her drug was not very sensitive to price.

The Elasticity of Demand for a Linear Demand Curve

The elasticity of demand is generally different at different points on the demand curve. In other words, the market power of a firm is not constant: it depends on the price that a firm has chosen to set. To illustrate, remember that we found −(elasticity of demand) = 3 for our demand curve when the price is $15. Suppose we calculate the elasticity for this same demand curve at $4. Thus imagine that that we are originally at the point where the price is $4 and sales are 80 units and then suppose we again decrease the price by 20 cents. Sales will increase by 1 unit:

1. The percentage change in the quantity is (1.25 percent).
2. The percentage change in the price is (5 percent).
3. −(elasticity of demand) is

The elasticity of demand is different because we are at a different point on the demand curve.

When −(elasticity of demand) increases, we say that demand is becoming more elastic. When −(elasticity of demand) decreases, we say that demand is becoming less elastic. As we move down a linear demand curve, −(elasticity of demand) becomes smaller, as shown in Figure 7.3.4.

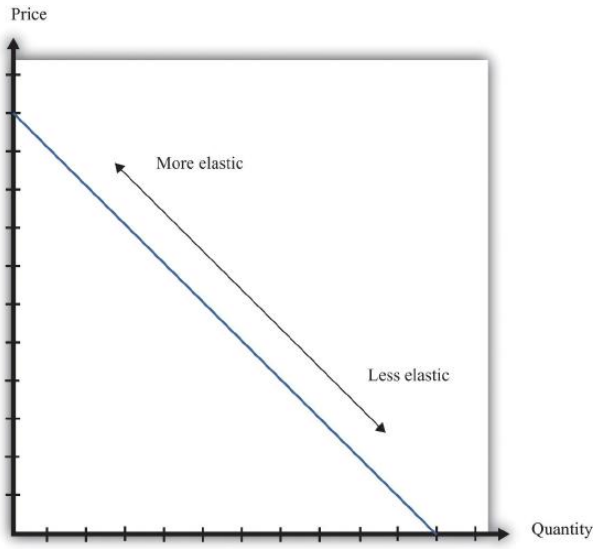


Figure 7.3.4: The Elasticity of Demand When the Demand Curve Is Linear. The elasticity of demand is generally different at different points on a demand curve. In the case of a linear demand curve, −(elasticity of demand) becomes smaller as we move down the demand curve.

Measuring the Elasticity of Demand

To evaluate the effects of her decisions on revenues, Ellie needs to know about the demand curve facing her firm. In particular, she needs to know whether the quantity demanded by buyers is very sensitive to the price that she sets. We now know that the elasticity of demand is a useful measure of this sensitivity. How can managers such as Ellie gather information on the elasticity of demand?

At an informal level, people working in marketing and sales are likely to have some idea of whether their customers are very price sensitive. Marketing and sales personnel—if they are any good at their jobs—spend time talking to actual and potential customers and should have some idea of how much these customers care about prices. Similarly, these employees should have a good sense of the overall market and the other factors that might affect customers' choices. For example, they will usually know whether there are other firms in the market offering similar products, and, if so, what prices these firms are charging. Such knowledge is much better than nothing, but it does not provide very concrete evidence on the demand curve or the elasticity of demand.

A firm may be able to make use of existing sales data to develop a more concrete measure of the elasticity of demand. For example, a firm might have past sales data that show how much they managed to sell at different prices, or a firm might have sales data from different cities where different prices were charged. Suppose a pricing manager discovers data for prices and quantities like those in

part (a) of Figure 7.3.5. Here, each dot marks an observation—for example, we can see that in one case, when the price was $100, the quantity demanded was 28.

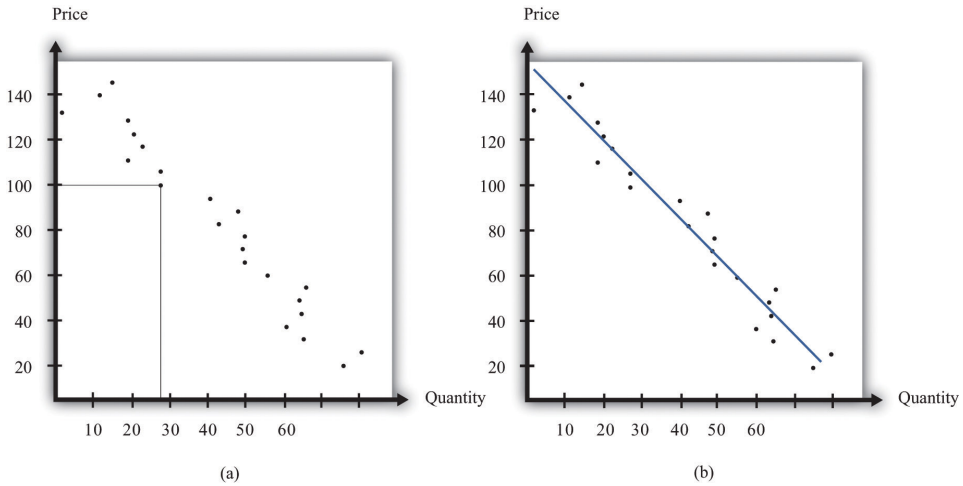


Figure 7.3.5: Finding the Demand Curve. (a) This is an example of data that a manager might have obtained for prices and quantities. (b) A line is fit to the data that represents a best guess at the underlying demand curve facing a firm.

The straight-line demand curves that appear in this and other books are a convenient fiction of economists and textbook writers, but no one has actually seen one in captivity. In the real world of business, demand curves—if they are available at all—are only a best guess from a collection of data. Economists and statisticians have developed statistical techniques for these guesses. The underlying idea of these techniques is that they *fit a line* to the data. (The exact details do not concern us here; you can learn about them in more advanced courses in economics and statistics.) Part (b) of Figure 7.3.5 shows an example. It represents our best prediction, based on available data, of how much people will buy at different prices.

If a firm does not have access to reliable existing data, a third option is for it to generate its own data. For example, suppose a retailer wanted to know how sensitive customer demand for milk is to changes in the price of milk. It could try setting a different price every week and observe its sales. It could then plot them in a diagram like Figure 7.3.5 and use techniques like those we just discussed to fit a line. In effect, the store could conduct its own experiment to find out what its demand curve looks like. For a firm that sells over the Internet, this kind of experiment is particularly attractive because it can randomly offer different prices to people coming to its website.

Finally, firms can conduct market research either on their own or by hiring a professional market research firm. Market researchers use questionnaires and surveys to try to discover the likely purchasing behavior of consumers. The simplest questionnaire might ask, “How much would you be willing to pay for product x?” Market researchers have found such questions are not very useful because consumers do not answer them very honestly. As a result, research firms use more subtle questions and other more complicated techniques to uncover consumers’ willingness to pay for goods and services.

Ellie decided that she should conduct market research to help with the pricing decision. She hired a market research firm to ask doctors about how they currently prescribed different high blood pressure medications. Specifically, the doctors were asked what percentage of their prescriptions went to each of the drugs on the market. Then they were asked the effect of different prices on those percentages. Based on this research, the market research firm found that a good description of the demand curve was as follows:

$$quantity\ demanded\ =\ 252 - 300 \times price.$$

Remember that the drug was currently being sold for $0.50 a pill, so

$$quantity\ demanded\ =\ 252 - 300 \times 0.5 = 102.$$

The demand curve also told Ellie that if she increased the price by 10 percent to $0.55, the quantity demanded would decrease to 87 (252 − 300 × 0.55 = 87). Therefore, the percentage change in quantity is From this, the market research firm discovered that the elasticity of demand at the current price was

How Do Revenues Depend on Price?

The next step is to understand how to use the demand curve when setting prices. The elasticity of demand describes how quantity demanded depends on price, but what a manager really wants to know is how *revenues* are affected by price. Revenues equal price times quantity, so we know immediately that a firm earns $0 if the price is $0. (It doesn't matter how much you give away, you still get no money.) We also know that, at the choke price, the quantity demanded is 0 units, so its revenues are likewise $0. (If you sell 0 units, it doesn't matter how high a price you sell them for.) At prices between $0 and the choke price, however, the firm sells a positive amount at a positive price, thus earning positive revenues. Figure 7.3.6 is a graphical representation of the revenues of a firm. Revenues equal price times quantity, which is the area of the rectangle under the demand curve. For example, at $14 and 30 units, revenues are $420.

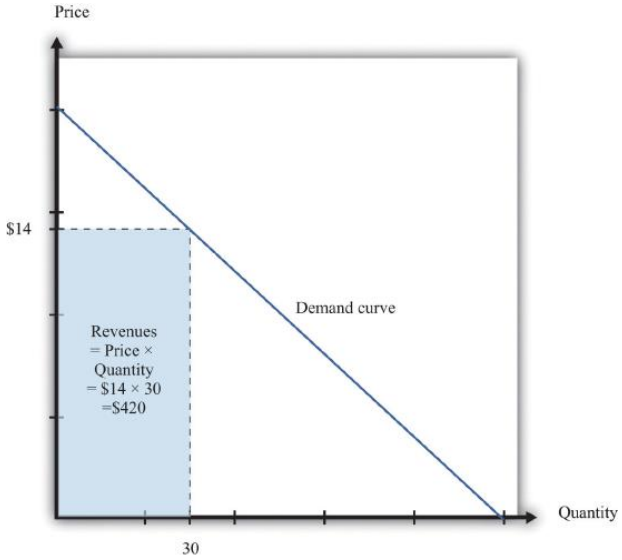


Figure 7.3.6: Revenues. The ***revenues of a firm*** *are equal to the area of the rectangle under the demand curve.*

We can use the information in Table 7.3.1 "Example of the Demand Curve Faced by a Firm" to calculate the revenues of a firm at different quantities and prices (this is easy to do with a spreadsheet). Table 7.3.2 "Calculating Revenues" shows that if we start at a price of zero and increase the price, the firm's revenues also increase. Above a certain point, however (in this example, $10), revenues start to decrease again.

Table 7.3.2: Calculating Revenues

Price($)	Quantity	Revenues ($)
0	100	0
2	90	180
4	80	320
6	70	420
8	60	480
10	50	500
12	40	480
14	30	420
16	20	320
18	10	180
20	0	0

Marginal Revenue

Earlier we suggested that a good strategy for pricing is to experiment with small changes in price. So how do small changes in price affect the revenue of a firm? Suppose, for example, that a firm has set the price at $15 and sells 25 units, but the manager contemplates decreasing the price to $14.80. We can see the effect that this has on the firm's revenues in Figure 7.3.7.

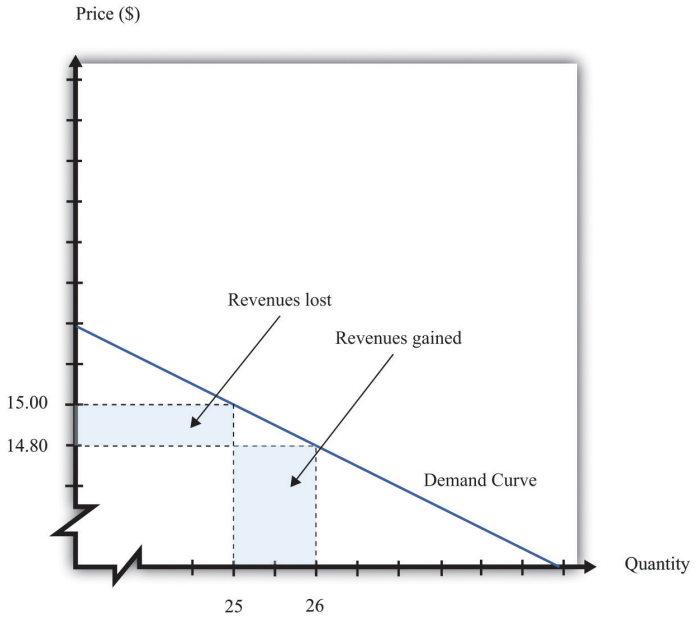


Figure 7.3.7: Revenues Gained and Lost. If a firm cuts its price, it sells more of its product, which increases revenues, but sells each unit at a lower price, which decreases revenues.

The firm will lose 20 cents on each unit it sells, so it will lose $5 in revenue. This is shown in the figure as the rectangle labeled "revenues lost." But the firm will sell more units: from the demand curve, we know that when the firm decreases its price by $0.20, it sells another unit. That means that the firm gains $14.80, as shown in the shaded area labeled "revenues gained." The overall change in the firm's revenues is equal to $14.80 − $5.00 = $9.80. Decreasing the price from $15.00 to $14.80 will increase its revenues by $9.80.

Look carefully at Figure 7.3.7 and make sure you understand the experiment. *We presume throughout this chapter that a firm must sell every unit at the same price*. When we talk about moving from $15.00 to $14.80, we are *not* supposing that a firm sells 25 units for $15 and then drops its price to $14.80 to sell the additional unit. We are saying that the manager is choosing between selling 25 units for $15.00 or 26 units for $14.80

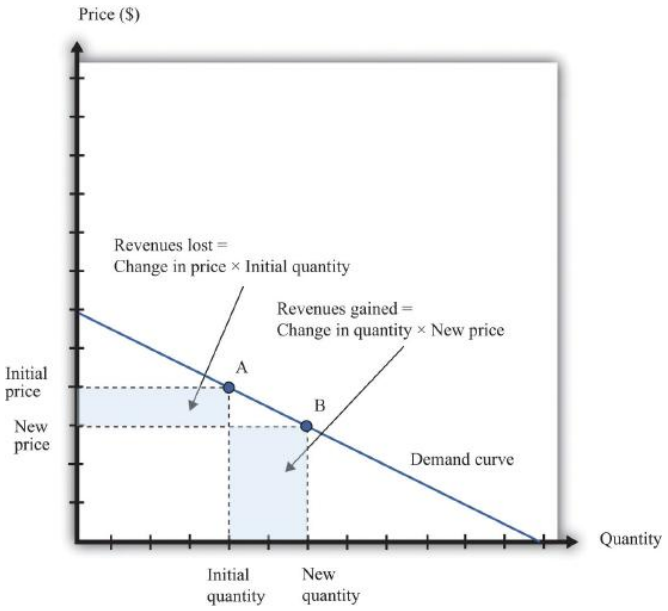


Figure 7.3.8: Calculating the Change in Revenues. If a manager has an idea about how much quantity demanded will decrease for a given increase in price, she can calculate the likely effect on revenues.

Figure 7.3.8 explains this idea more generally. Suppose a firm is originally at point A on the demand curve. Now imagine that a manager decreases the price. At the new, lower price, the firm sells a new, higher quantity (point B). The change in the quantity is the new quantity minus the initial quantity. The change in the price is the new price minus the initial price (remember that this is a negative number). The change in the firm's revenues is given by

$$change\ in\ revenues\ =\ (change\ in\ quantity\ \times\ new\ price)\ +\ (change\ in\ price\ \times\ initial\ quantity).$$

The first term is positive: it is the extra revenue from selling the extra output. The second term is negative: it is the revenue lost because the price has been decreased. Together these give the effect of a change in price on revenues, which we call a firm's **marginal revenue**.

Toolkit: Section 31.15

Marginal revenue is the *change* in revenue associated with a *change* in quantity of output sold:

We can write this asFor the derivation of this expression, see the toolkit.

Marginal Revenue and the Elasticity of Demand

Given the definitions of marginal revenue and the elasticity of demand, we can write

It may look odd to write this expression with two minus signs. We do this because it is easier to deal with the positive number: −(elasticity of demand). We see three things:

1. Marginal revenue is always less than the price. Mathematically, Suppose a firm sells an extra unit. If the price stays the same, then the extra revenue would just equal the price. But the price does not stay the same: it decreases, meaning the firm gets less for every unit that it sells.
2. Marginal revenue can be negative. If –(elasticity of demand) < 1, then and When marginal revenue is negative, increased production results in *lower* revenues for a firm. The firm sells more output but loses more from the lower price than it gains from the higher sales.
3. The gap between marginal revenue and price depends on the elasticity of demand. When demand is more elastic, meaning −(elasticity of demand) is a bigger number, the gap between marginal revenue and price becomes smaller.

These three ideas are illustrated in Figure 7.3.9. The demand curve shows us the price at any given quantity. The marginal revenue curve lies below the demand curve because of our first observation: at any quantity, marginal revenue is less than price.When a demand curve is a straight line, the marginal revenue curve is also a straight line with the same intercept, but it is twice as steep. The marginal revenue curve intersects the horizontal axis at 50 units: when output is less than 50 units, marginal revenue is

positive; when output exceeds 50, marginal revenue is negative. We explained earlier that a linear demand curve becomes more inelastic as you move down it. When the demand curve goes from being relatively elastic to relatively inelastic, marginal revenue goes from being positive to being negative.

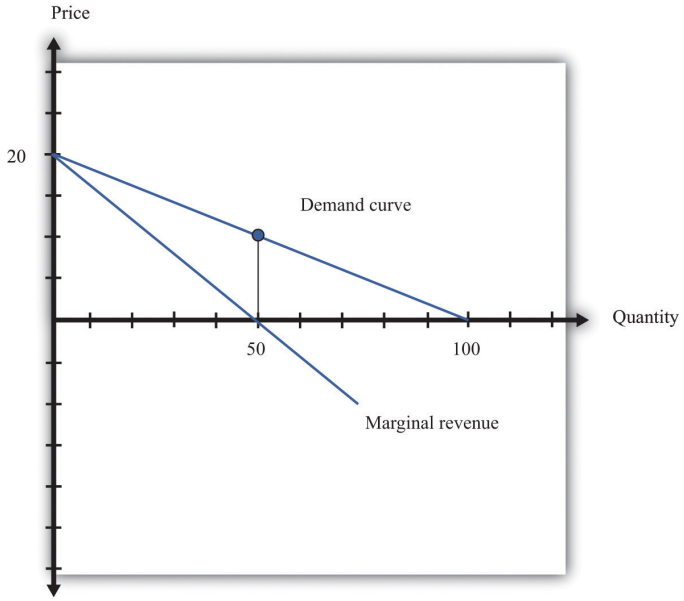


Figure 7.3.9: Marginal Revenue and Demand. The marginal revenue curve lies below the demand curve because at any quantity, marginal revenue is less than price.

Earlier, we showed that when a firm sets the price at \$15, −(elasticity of demand) = 3. Thus we can calculate marginal revenue at this price:

What does this mean? Starting at \$15, it means that if a firm decreases its price—and hence increase its output—by a small amount, there would be an increase in the firm's revenues.

When revenues are at their maximum, marginal revenue is zero. We can confirm this by calculating the elasticity of demand at \$10. Consider a 10 percent increase in price, so the price increases to \$11. At \$10, sales equal 50 units. At \$11, sales equal 45 units. In other words, sales decrease by 5 units, so the decrease in sales is 10 percent. It follows that

Plugging this into our expression for marginal revenue, we confirm that

At \$10, a small change in price leads to no change in revenue. The benefit from selling extra output is exactly offset by the loss from charging a lower price.

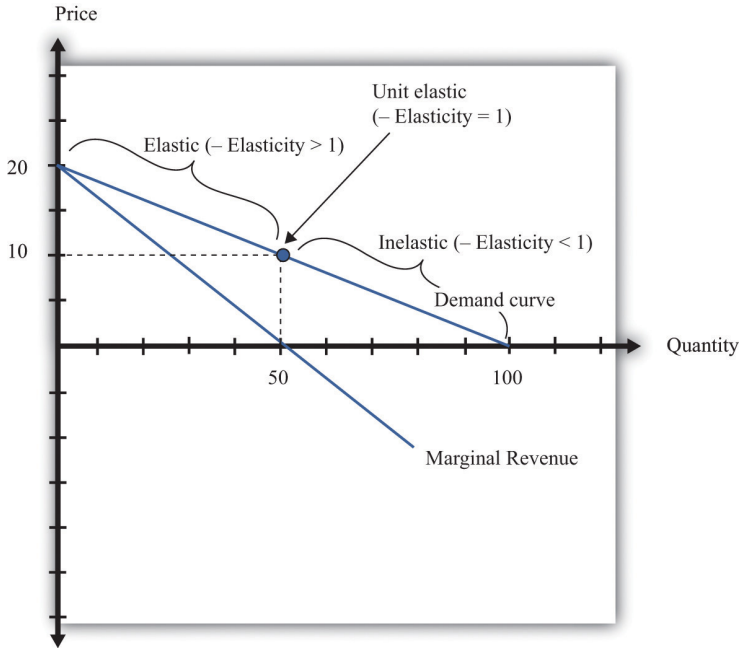


Figure 7.3.10: Marginal Revenue and the Elasticity of Demand. The demand curve can be divided into two parts: at low quantities and high prices, marginal revenue is positive and the demand curve is elastic; at high quantities and low prices, marginal revenue is negative and the demand curve is inelastic.

We can thus divide the demand curve into two parts, as in Figure 7.3.10. At low quantities and high prices, a firm can increase its revenues by moving down the demand curve—to lower prices and higher output. Marginal revenue is positive. In this region, −(elasticity of demand) is a relatively large number (specifically, it is between 1 and infinity) and we say that the demand curve is relatively elastic. Conversely, at high quantities and low prices, a decrease in price will decrease a firm's revenues. Marginal revenue is negative. In this region, −(elasticity of demand) is between 0 and 1, and we say that the demand curve is inelastic. Table 7.3.3 represents this schematically.

Table 7.3.3

−(Elasticity of Demand)	Demand	Marginal Revenue	Effect of a Small Price Decrease
• > –(elasticity of demand) > 1	Relatively elastic	Positive	Increase revenues
−(elasticity of demand) = 1	Unit elastic	Zero	Have no effect on revenues
1 > –(elasticity of demand) > 0	Relatively inelastic	Negative	Decrease revenues

Maximizing Revenues

The market research company advising Ellie made a presentation to her team. The company told them that if they increased their price, they could expect to see a decrease in revenue. At their current price, in other words, marginal revenue was positive. If Ellie's team wanted to maximize revenue, they would need to recommend a reduction in price: down to the point where marginal revenue is $0—equivalently, where −(elasticity of demand) = 1.

Some members of Ellie's team therefore argued that they should try to decrease the price of the product so that they could increase their market share and earn more revenues from the sale of the drug. Ellie reminded them, though, that their goal wasn't to have as much revenue as possible. It was to have as large a *profit* as possible. Before they could decide what to do about price, they needed to learn more about the costs of producing the drug

key takeaways

- The demand curve tells a firm how much output it can sell at different prices.
- The elasticity of demand is the percentage change in quantity divided by the percentage change in the price.
- Marginal revenue is the change in total revenue from a change in the quantity sold.

check your understanding

1. Earlier, we saw that the demand curve was

$$quantity\ demanded\ =\ 252 - 300 \times price.$$

   - Suppose Ellie sets the price at $0.42. What is the quantity demanded?
   - Suppose Ellie sets a price that is 10 percent higher ($0.462). What is the quantity demanded?
   - Confirm that −(elasticity of demand) = 1 when the price is $0.42.

1. If a firm's manager wants to choose a price to maximize revenue, is this the same price that would maximize profits?
2. If a demand curve has the same elasticity at every point, does it also have a constant slope?

This page titled 7.3: The Revenues of a Firm is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

7.4: The Costs of a Firm 7.4: The Costs of a FirmEdit Update titlecancel

Learning Objectives

1. What is marginal cost?
2. What costs matter for a firm's pricing decision?

The goods and services that firms put up for sale don't appear from nowhere. Firms produce these goods and incur costs as a result. When a marketing manager is thinking about the price that she sets, she must take into account that different prices lead to different levels of production and hence to different costs for the firm.

Your typical image of a firm probably involves a manufacturing process. This could be very simple indeed. For example, there are firms in Malaysia that produce palm oil. A firm in this context is little more than a big piece of machinery in the middle of a jungle of palm trees. The production process is hot, noisy, and very straightforward: (1) laborers harvest palm nuts from the trees surrounding the factory; (2) these palm nuts are crushed, heated, and pressed to extract the oil; and (3) the oil is placed into barrels and then sold. If you wanted to run a palm oil production business in Malaysia, you would need to purchase the following:

- Machine for extracting the oil
- Truck to transport the oil to market
- Generator to power the machinery
- Fuel to power the generator
- Gasoline for the truck
- Labor time from workers—to harvest the nuts, run the machinery, and transport the oil to market

That's it. It is not difficult to become a palm oil entrepreneur! In this case, it is quite easy to list the main costs of production for the firm.

In other businesses, however, it is much more difficult. Imagine trying to make a similar list for Apple Computer, with all its different products, production plants in different countries, canteens for their workers, pension plans, and so on. Of course, Apple's accountants still need to develop a list of Apple's expenses, but they keep their jobs manageable by grouping Apple's expenditures into various categories. If this were an accounting textbook, we would discuss these categories in detail. Our task here is simpler: we only need to determine how these costs matter for pricing decisions.

Marginal Cost

Earlier, we showed how a firm's revenues change when there is a change in quantity that the firm produces. If we also know how a firm's costs change when there is a change in output, we have all the information we need for good pricing decisions. As with revenues, we scale this change by the size of the change in quantity. Figure 7.4.1 shows how **marginal cost** fits into our road map for the chapter.

Toolkit: Section 31.14

Marginal cost is the *change* in cost associated with a *change* in quantity of output produced:

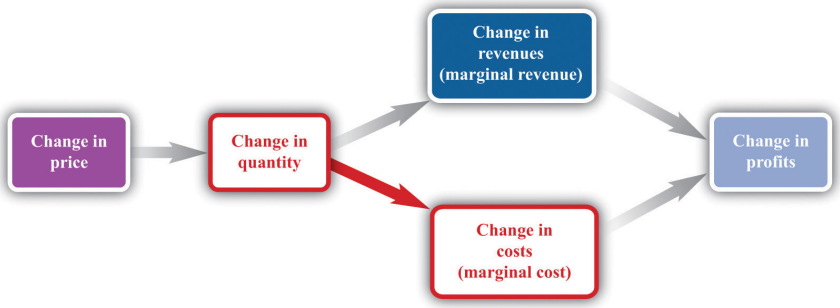


Figure 7.4.1: Marginal Cost. When a firm sets a higher price, it sells a smaller quantity and its costs of production decrease.

Table 7.4.1: Marginal Cost

Output	Total Costs	Marginal Cost ($)
0	50	
		10
1	60	
		10
2	70	
		10
3	80	
		10
4	90	
		10
5	100	
…		

Table 7.4.1 "Marginal Cost" shows an example of a firm's costs. It calculates marginal cost in the last column. We have presented this table with marginal cost on separate rows to emphasize that marginal cost is the cost of going from one level of output to the next. In our example, marginal cost—the cost of producing one more unit—is $10. If you want to produce one unit, it will cost you $60. If you want to produce two units, your must pay an *additional* $10 in costs, for a total of $70. If you want to produce three units, you must pay the $70 to produce the first two units, plus the additional marginal cost of $10, for a total cost of $80, and so on. Graphically, marginal cost is the *slope* of the cost line, as shown in Figure 7.4.2.

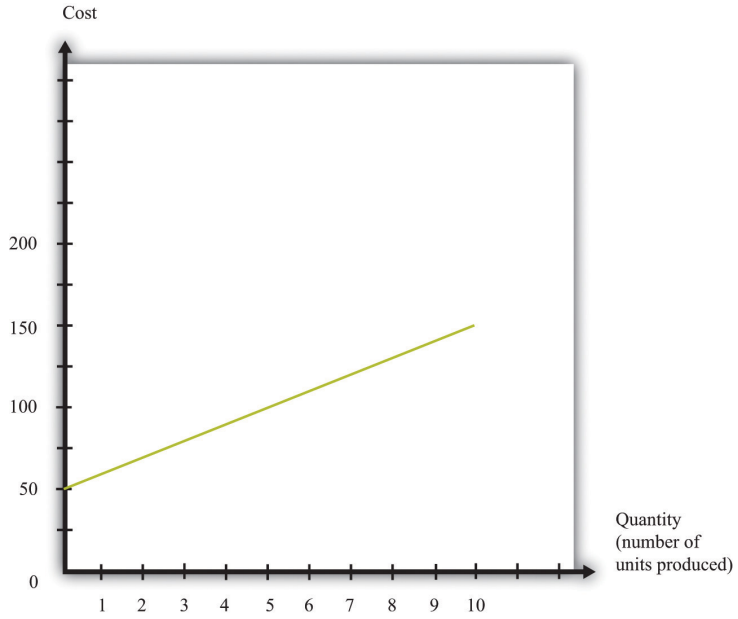


Figure 7.4.2: An Example of a Cost Function. This graph illustrates the cost function for a firm. In this case, the cost function has the equation cost = 50 + 10 × quantity. The cost of producing each additional unit (the marginal cost) is $10.

To emphasize again, *only those costs that change matter for a firm's pricing decision*. When a firm considers producing extra output, many of its costs do not change. We can completely ignore these costs when thinking about optimal pricing. This is not to say that other costs don't matter; quite the contrary. They are critical for a different decision—whether the firm should be in business at all.See Chapter 9. But as long as we are interested in pricing, we can ignore them.

The costs for developing pharmaceutical products are typically quite high. The drug that Ellie was responsible for was first developed in research laboratories, then tested on animals, and then run through a number of clinical studies on human patients. These studies were needed before the Food and Drug Administration in the United States, and equivalent drug safety organizations in other countries, would approve the drug for sale. But these development costs have no effect on marginal cost because they were all incurred before a single pill could be sold.

Surprisingly, this means that even though the drug was very expensive to develop, Ellie's team—quite correctly—paid no attention to that fact. In determining what price to set, they looked at the price sensitivity of their customers and at marginal cost. They understood that the development costs of the drug were not relevant to the pricing decision. In fact, Ellie's team had only a very vague idea how much the drug had cost to develop: after all, that development had been carried out by a completely different arm of the company in other parts of the world.

You will often hear the opposite argument. It is common for people to say that pharmaceutical companies charge high prices because it costs so much to develop their drugs. This argument is superficially appealing, but it is completely backward. Pharmaceutical companies don't charge high prices because they incur large development costs. They are willing to incur large development costs because they can charge high prices.

Estimating Marginal Cost

Estimating marginal cost is generally much easier than estimating demand because the cost side of the business is largely under the control of the firm. The firm's costs depend on its technology and the decisions made about using that technology. In most medium- or large-sized firms, there is an "operations department" that takes care of the production process. The marketing manager ought to be able to consult with her colleagues in operations and learn about the costs of the firm. Most importantly, even if it is unreasonable to expect an operations manager to know the firm's entire cost function, the operations manager should have a good idea about marginal cost (that is, should know how much it would cost per unit to scale up operations by a small amount). And that is the information the marketing manager needs for her pricing decisions.

key takeaways

- Marginal cost measures the additional costs from producing an extra unit of output.
- It is only the change in costs—marginal cost—that matter for a firm's pricing decision.

checking your understanding

If the marginal cost in Table 7.4.1 "Marginal Cost" were $20, what would be the cost of producing 10 units of the good?

This page titled 7.4: The Costs of a Firm is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

7.5: Markup Pricing- Combining Marginal Revenue and Marginal Cost 7.5: Markup Pricing- Combining Marginal Revenue and Marginal CostEdit Update titlecancel

Learning Objectives

1. What is the optimal price for a firm?
2. What is markup?
3. What is the relationship between the elasticity of demand and markup?

Let us review the ideas we have developed in this chapter. We know that changes in output lead to changes in both revenues and costs. Changes in revenues and costs lead to changes in profits (see Figure 7.5.1). We have a measure of how much revenues change if output is increased—called marginal revenue, which you can calculate if you know price and the elasticity of demand. We also have a measure of how much costs change if output is increased—this is called marginal cost. Given information on current marginal revenue and marginal cost, a marketing manager can then decide if a firm should change its price. In this section, we derive a rule that tells us how a manager should make this decision.

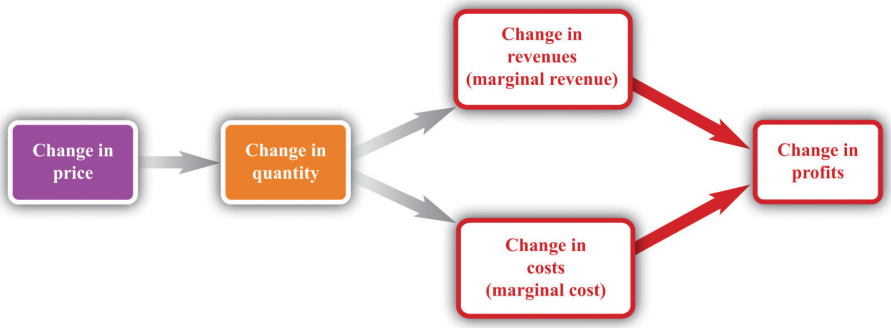


Figure 7.5.1: *Changes in Revenues and Costs Lead to Changes in Profits*. When a firm changes its price, this leads to changes in revenues and costs. The change in a firm's profit is equal to the change in revenue minus the change in cost—that is, the change in profit is marginal revenue minus marginal cost. When marginal revenue equals marginal cost, the change in profit is zero, so a firm is at the top of the profit hill.

In the real world of business, firms almost always choose the price they set rather than the quantity they produce. Yet the pricing decision is easier to analyze if we think about it the other way round: a firm choosing what quantity to produce and then accepting the price implied by the demand curve. This is just a matter of convenience: a firm chooses a point on the demand curve, and it doesn't matter if we think about it choosing the price and accepting the implied quantity or choosing the quantity and accepting the implied price (Figure 7.5.2).

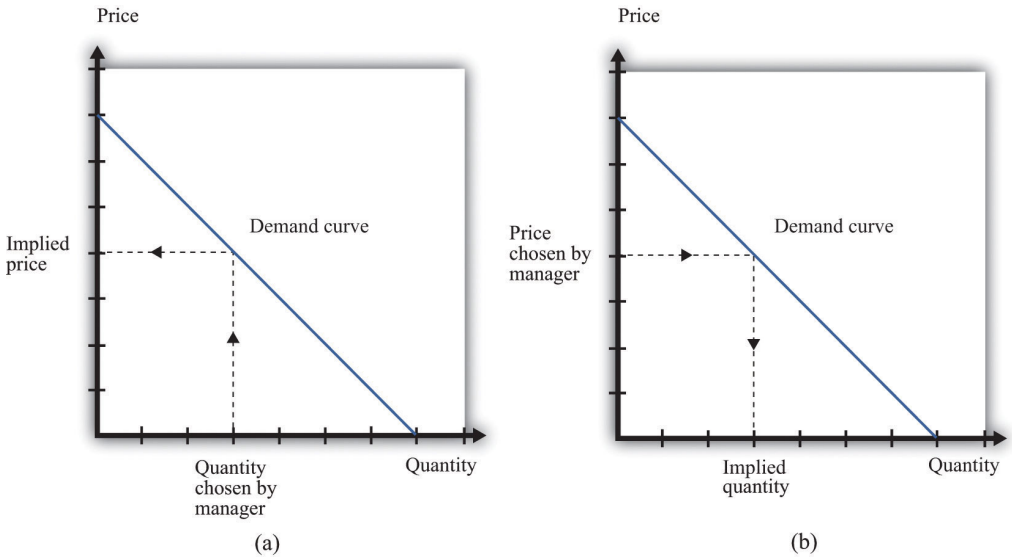


Figure 7.5.2: Setting the Price or Setting the Quantity. It doesn't matter if we think about choosing the price and accepting the implied quantity or choosing the quantity and accepting the implied price

Suppose that a marketing manager has estimated the elasticity of demand, looked at the current price, and used the marginal revenue formula to discover that the marginal revenue is $5. This means that if the firm increases output by one unit, its revenues will increase by $5. The marketing manager has also spoken to her counterpart in operations, who has told her that the marginal cost is $3. This means it would cost an additional $3 to produce one more unit. From these two pieces of information, the marketing manager knows that an increase in output would be a good idea. An increase in output leads to a bigger increase in revenues than in costs. As a result, it leads to an increase in profits: specifically, profits will increase by $2. This tells the marketing manager that it is a good idea to increase output. From the law of demand, she should think about decreasing the price.

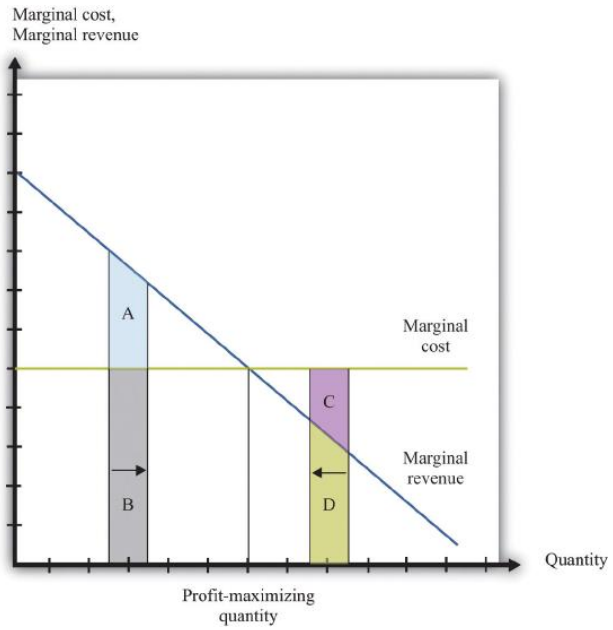


Figure 7.5.3: Optimal Pricing. To the left of the point marked "profit-maximizing quantity," marginal revenue exceeds marginal cost so increasing output is a good idea. The opposite is true to the right of that point.

Figure 7.5.3 shows this idea graphically. To the left of the point marked "profit-maximizing quantity," marginal revenue exceeds marginal cost. Suppose a firm is producing below this level. If it increases its output, the extra revenue it obtains will exceed the extra cost. We see that an increase in output yields extra revenue equal to the areas $A + B$ and extra costs equal to B. The increase in output yields extra profit, which is equal to A. Increasing its output is thus a good idea. Conversely, to the right of the profit-maximizing point, marginal revenue is less than marginal cost. If a firm reduces its output, the decrease in costs ($C + D$) exceeds the decrease in revenue (D). Decreases in output lead to increases in profit.

Profits are greatest when

$$marginal\ revenue\ =\ marginal\ cost.$$

This is the point where a change in price leads to no change in profits, so we are at the very top of the profit hill that we drew in Figure 7.5.3. See also Figure 7.5.2.

The Markup Pricing Formula

Think about Ellie's company. If it became more expensive for the company to produce each pill, it seems likely they would respond by raising their costs. Also, we said earlier that their customers are not very sensitive to changes in the price, which should allow them to set a relatively high price. In other words, the profit-maximizing price is related to the elasticity of demand and to marginal cost. These are the two critical ingredients of the pricing decision.

> **Toolkit: Section 31.15**
>
> Firms should set the price as a markup over marginal cost: This expression comes from combining the formula for marginal revenue and the condition that marginal revenue equals marginal cost. See the toolkit for more details.
>
> and

There are three facts about **markup**:

1. Markup is greater than or equal to zero—that is, the firm never sets a price below marginal cost.
2. Markup is smaller when demand is more elastic.
3. Markup is zero when the demand curve is perfectly elastic: −(elasticity of demand) = •.

Ellie's team looked at their numbers. At the current price, −(elasticity of demand) = 1.47. They learned that the marginal cost was $0.28 per pill, and they were charging $0.50 per pill. Their current markup, in other words, was about 79 percent: 0.5 = (1+ 0.79) × 0.28. But if they applied the markup pricing formula based on the *current* elasticity of demand, they could charge a markup of 1/0.47 = 2.12—that is, more than a 200 percent markup, leading to a price of $0.87. It was clear that they could do better by increasing their price

A Pricing Algorithm

To summarize, a manager needs two key pieces of information when determining price:

1. **Marginal cost.** We have shown that the profit-maximizing price is a markup over the marginal cost of production. If a manager does not know the magnitude of marginal cost, she is missing a critical piece of information for the pricing decision.
2. **Elasticity of demand.** Once a manager knows marginal cost, she should then set the price as a markup over marginal cost. But this should not be done in an ad hoc manner; the markup must be based on information about the elasticity of demand.

Given these two pieces of information, a manager can then use the markup formula to determine the optimal price. Be careful, though. The markup formula looks deceptively simple, as if it can be used in a "plug-and-play" manner—given marginal cost and the elasticity of demand, plug them into the formula and calculate the optimum price. But if you change the price, both marginal cost and the elasticity of demand are also likely to change. A more reliable way of using this formula is in the algorithm shown in Figure 7.5.4, which is based on our earlier idea that you should find your way to the top of the profit hill. The five steps are as follows:

1. At your current price, estimate marginal cost and the elasticity of demand.
2. Calculate the optimal price based on those values.
3. If the optimal price is greater than your actual price, increase your price. Then estimate marginal cost and the elasticity again and repeat the process.
4. If the optimal price is less than your actual price, decrease your price. Then estimate marginal cost and the elasticity again and repeat the process.
5. If the current price is equal to this optimal price, leave your price unchanged.

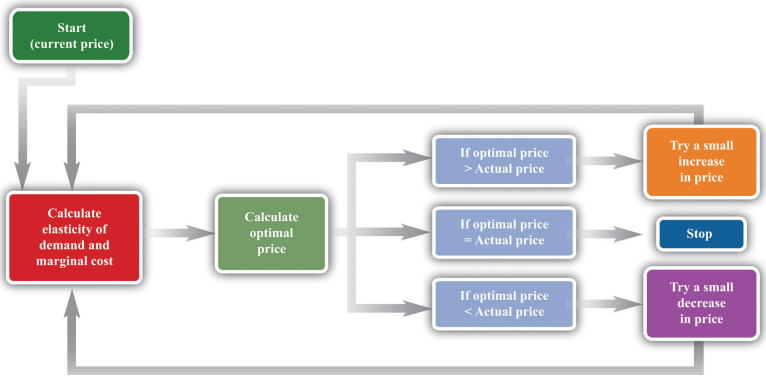


Figure 7.5.4: A Pricing Algorithm. This ***pricing algorithm*** *shows how to get the best price for a product.*

Ellie's team members were aware that, even though demand for the drug was apparently not very sensitive to price, they should not immediately jump to a much higher markup. They had found that *based on current marginal cost and elasticity*, the price could be raised. But as they raised the price, they knew that the elasticity of demand would probably also change. Looking more closely at their market research data, they found that at a price of $0.56 (a 100 percent markup), the elasticity of demand would increase to about 2. An elasticity of 2 means that the markup should be 100 percent to maximize profits. Thus—at least if their market research data were reliable—they knew that a price of $0.56 would maximize profits. Ellie recommended to senior management that the price of the drug be raised by slightly over 10 percent, from $0.50 per pill to $0.56 per pill.

Shifts in the Demand Curve Facing a Firm

So far we have looked only at *movements along* the demand curve—that is, we have looked at how changes in price lead to changes in the quantity that customers will buy. Firms also need to understand what factors might cause their demand curve to shift. Among the most important are the following:

- **Changes in household tastes.** Starting around 2004 or so, low-carbohydrate diets started to become very popular in the United States and elsewhere. For some companies, this was a boon; for others it was a problem. For example, companies like Einstein Bros. Bagels or Dunkin' Donuts sell products that are relatively high in carbohydrates. As more and more customers started looking for low-carb alternatives, these firms saw their demand curve shift inward.
- **Business cycle.** Consider Lexus, a manufacturer of high-end automobiles. When the economy is booming, sales are likely to be very good. In boom times, people feel richer and more secure and are more likely to purchase a luxury car. But if the economy goes into recession, potential car buyers will start looking at cheaper cars or may decide to defer their purchase altogether. Many companies sell products that are sensitive to the state of the business cycle. Their demand curves shift as the economy moves from boom to recession.
- **Changes in competitors' prices.** In a business setting, this is a critical concern. If a competitor decreases its price, this means that the demand curve you face will shift inward. For example, suppose that British Airways decides to decrease its price for flights from New York to London. American Airlines will find that its demand curve for that route has shifted inward. Ellie certainly has to worry about this because her company's product has only a small number of competitors. A change in price of a competing blood pressure drug might make a big difference in the sales and profits of Ellie's product.

If the demand curve shifts, should a firm change its price? The answer is yes *if the shift in the demand curve also leads to a change in the elasticity of demand*. In practice, this is likely to be the case, although it is certainly possible for a demand curve to shift without a change in the elasticity of demand. The correct response to a shift in the demand curve is to reestimate the elasticity of demand and then decide if a change in price is appropriate.

Complications

Pricing is a difficult and delicate job, and there are many factors that we have not yet considered:We address some of them in other chapters of the book; others are topics for more advanced classes in economics and business strategy.

- By far the most important problem that we have neglected is as follows: When making pricing decisions, *firms may need to take into account how other firms will respond to their decisions*. For example, a manager might estimate her firm's elasticity of demand and marginal cost and determine that she could make more money by decreasing price. That calculation presumes that competing firms keep their prices unchanged. In markets with a small number of competitors, it is instead quite likely that other

firms would respond by decreasing their prices. This would cause a firm's demand curve to shift inward and probably leave it worse off than before.

- We have assumed throughout that a firm has to charge the same price for every unit that it sells. In many cases, this is an accurate description of pricing behavior. When a grocery store posts a price, that price holds for every unit on the shelf. But sometimes firms charge different prices for different units—by either charging different prices to different customers or offering individual units at different prices to the same customer. You have undoubtedly encountered examples. Firms sometimes offer quantity discounts, so the price is lower if you buy more units. Sometimes they offer discounts to certain groups of customers, such as cheap movie tickets for students. We could easily fill an entire chapter with other examples—some of which are remarkably sophisticated.
- Firms can have pricing strategies that call for the price to change over time. For example, firms sometimes engage in a strategy known as penetration pricing, whereby they start off by charging a low price in an attempt to develop or expand the market. Imagine that Kellogg's develops a new breakfast cereal. It might decide to offer the cereal at a low price to induce people to try the product. Only after it has developed a group of loyal customers would it start setting their prices according to the markup principle.
- Pricing plays a role in the overall marketing and branding strategy of a firm. Some firms position themselves in the marketplace as suppliers of high-end offerings. They may choose to set high prices for their products to ensure that customers perceive them appropriately. Consider a luxury hotel that is contemplating setting a very low price in the off-season. Even though such a strategy might make sense in terms of its profits at that time, it might do long-term damage to the hotel's reputation. For various reasons, customers often use the price of a product as an indicator of that product's quality, so a low price can adversely affect a firm's image.
- Psychologists who study marketing have found that demand is sensitive at certain price points. For example, if a firm increases the price of a product from $99.98 to $99.99, there might be very little effect on demand. But if the price increases from $99.99 to $100.00, there might be a much bigger effect because $100.00 is a psychological barrier. Such consumer behavior does not seem completely rational, but there is little doubt that it is a real phenomenon.
- Throughout this chapter, we have said that there is no difference between a firm choosing its price and taking as given the implied quantity or choosing its quantity and taking as given the implied price. Either way, the firm is picking a point on the demand curve. This is true, but there is a footnote that we should add. A firm's demand curve depends on what its competitors are doing and, oddly enough, it *does* make a difference if those competitors are choosing quantities or prices.See Chapter 15 for discussion of this. We should also note that firms often do not know their demand curves with complete certainty. Suppose, for example, that the true demand curve for a firm's product is actually further outward than a firm expects. If the firm sets the price, it will end up with an unexpectedly large quantity being demanded. If the firm sets the quantity, it will end up with an unexpectedly high price.
- We have focused our attention on the market power of firms as sellers, as reflected in the downward-sloping demand curves they face. Firms can also have market power as buyers. Walmart is such an important customer for many of its suppliers that it can use its position to negotiate lower prices for the goods it buys. Governments are also often powerful buyers and may be able to influence the prices they pay for goods and services. For example, government-run health-care systems may be able to negotiate favorable prices with pharmaceutical companies.

key takeaways

- At the profit-maximizing price, marginal revenue equals marginal cost.
- Markup is the difference between price and marginal cost, as a percentage of marginal cost.
- The more elastic the demand curve faced by a firm, the smaller the markup.

check your understanding

1. We said that markup is always greater than zero. Look at the formula for markup. If markup is greater than zero, what must be true about −(elasticity of demand)? Can you see why this must be true? Look back at Figure 7.5.10 for a hint.
2. If price is a markup over marginal cost, then how does marginal revenue influence the pricing decision of a firm?
3. Starting at the profit-maximizing price, if a firm increases its price, could revenue increase?

This page titled 7.5: Markup Pricing- Combining Marginal Revenue and Marginal Cost is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

7.6: The Supply Curve of a Competitive Firm 7.6: The Supply Curve of a Competitive FirmEdit Update titlecancel

Learning Objectives

1. What is a perfectly competitive market?
2. In a perfectly competitive market, what does the demand curve faced by a firm look like?
3. What happens to the pricing decision of a firm in a perfectly competitive market?

In this chapter, we have paid a great deal of attention to demand, but we have not spoken of supply. There is a good reason for this: a firm with market power does not have a supply curve. A supply curve for a firm tells us how much output the firm is willing to bring to market at different prices. But a firm with market power looks at the demand curve that it faces and then chooses a point on that curve (a price and a quantity). Price, in this chapter, is something that a firm chooses, not something that it takes as given. What is the connection between our analysis in this chapter and a **market supply curve**?

Perfectly Competitive Markets

If you produce a good for which there are few close substitutes, you have a great deal of market power. Your demand curve is not very elastic: even if you charge a high price, people will be willing to buy the good. On the other hand, if you are the producer of a good that is very similar to other products on the market, then your demand curve will be very elastic. If you increase your price even a little, the demand for your product will decrease a lot.

The extreme case is called a perfectly **competitive market**. In a perfectly competitive market, there are numerous buyers and sellers of exactly the same good. The standard examples of perfectly competitive markets are those for commodities, such as copper, sugar, wheat, or coffee. One bushel of wheat is the same as another, there are many producers of wheat in the world, and there are many buyers. Markets for financial assets may also be competitive. One euro is a perfect substitute for another, one three-month US treasury bill is a perfect substitute for another, and there are many institutions willing to buy and sell such assets.

Toolkit: Section 31.9

You can review the market supply curve and the definition of a perfectly competitive market in the toolkit.

An individual seller in a competitive market has no control over price. If the seller tries to set a price above the going market price, the quantity demanded falls to zero. However, the seller can sell as much as desired at the market price. When there are many sellers producing the same good, the output of a single seller is tiny relative to the whole market, and so the seller's supply choices have no effect on the market price. This is what we mean by saying that the seller is "small." It follows that a seller in a perfectly competitive market faces a demand curve that is a horizontal line at the market price, as shown in Figure 7.6.1. This demand curve is *infinitely elastic*: −(elasticity of demand) = ∞. Be sure you understand this demand curve. As elsewhere in the chapter, it is the *demand faced by an individual firm*. In the background, there is a market demand curve that is downward sloping in the usual way; the market demand and market supply curves together determine the market price. But an individual producer does not experience the market demand curve. The producer confronts an infinitely elastic demand for its product.

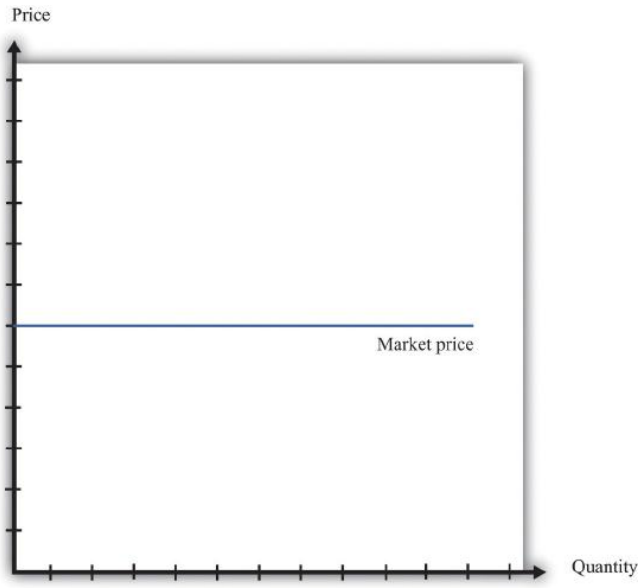


Figure 7.6.1: The Demand Curve Facing a Firm in a Perfectly Competitive Market. The demand curve faced by a firm in a perfectly competitive market is infinitely elastic. Graphically, this means that it is a horizontal line at the market price.

Everything we have shown in this chapter applies to a firm facing such a demand curve. The seller still picks the best point on the demand curve. But because the price is the same everywhere on the demand curve, picking the best point means picking the best quantity. To see this, go back to the markup formula. When demand is infinitely elastic, the markup is zero:

so price equals marginal cost:

$$price \ = \ (1 + markup) \times marginal \ cost \ = \ marginal \ cost.$$

This makes sense. The ability to set a price above marginal cost comes from market power. If you have no market power, you cannot set a price in excess of marginal cost. A perfectly competitive firm chooses its level of output so that its marginal cost of production equals the market price.

We could equally get this conclusion by remembering that

marginal revenue = marginal cost

and that when −(elasticity of demand) is infinite, marginal revenue equals price. If a competitive firm wants to sell one more unit, it does not have to decrease its price to do so. The amount it gets for selling one more unit is therefore the market price of the product, and the condition that marginal revenue equals marginal cost becomes

price = marginal cost.

For the goods and services that we purchase regularly, there are few markets that are truly perfectly competitive. Often there are many sellers of goods that may be very close substitutes but not absolutely identical. Still, many markets are close to being perfectly competitive, in which case markup is very small and perfect competition is a good approximation.

The Supply Curve of a Firm

Table 7.6.1 "Costs of Production: Increasing Marginal Cost" shows the costs of producing for a firm. In contrast to Table 7.4.1 "Marginal Cost", where we supposed marginal cost was constant, this example has higher marginal costs of production when the level of output is greater.Total cost in Table 7.6.1 "Costs of Production: Increasing Marginal Cost" is 50 + 10 × quantity + 2 × quantity^2.

Table 7.6.1: Costs of Production: Increasing Marginal Cost

Output	Total Costs ($)	Marginal Cost ($)
1	22	12
2	38	16

Output	Total Costs ($)	Marginal Cost ($)
3	58	20
4	82	24
5	110	28
…		

Figure 7.6.2 shows how we derive the supply curve of an individual firm given such data on costs. The supply curve tells us how much the firm will produce at different prices. Suppose, for example, that the price is $20. At this price, we draw a horizontal line until we reach the marginal cost curve. At that point, we draw a vertical line to the quantity axis. In this way, you can find the level of output such that marginal cost equals price. Looking at the figure, we see that the firm should produce 3 units because the marginal cost of producing the third unit is $20. When the price is $30, setting marginal cost equal to price requires the firm to produce 5.5 units. When the price is $40, setting marginal cost equal to price requires the firm to produce 8 units.

The supply curve shows us the quantity that a firm will produce at different prices. Figure 7.6.2 reveals something remarkable: *the* **individual supply curve** *of the firm is the marginal cost curve*. They are the same thing. As the price a firm faces increases, it will produce more. Note carefully how this is worded. We are *not* saying that if a firm produces more, it will charge a higher price. Firms in a competitive market must take the price as given. Instead, we think about the response of a firm to a change in the price.The individual firm's supply curve is an exact counterpart to something we show in Chapter 4, where we derive the demand curve for an individual. We show that an individual buys a good up to the point where marginal valuation equals price. From this we can conclude that the demand curve for an individual is the same as the individual's marginal valuation curve. In Chapter 8, we use an individual firm's supply curve as the basis for the **market supply curve**. Likewise, we use the individual demand curve as the basis for the **market demand curve**. By combining these curves, we obtain the **supply-and-demand** framework, which we can use to understand changing prices in an economy.

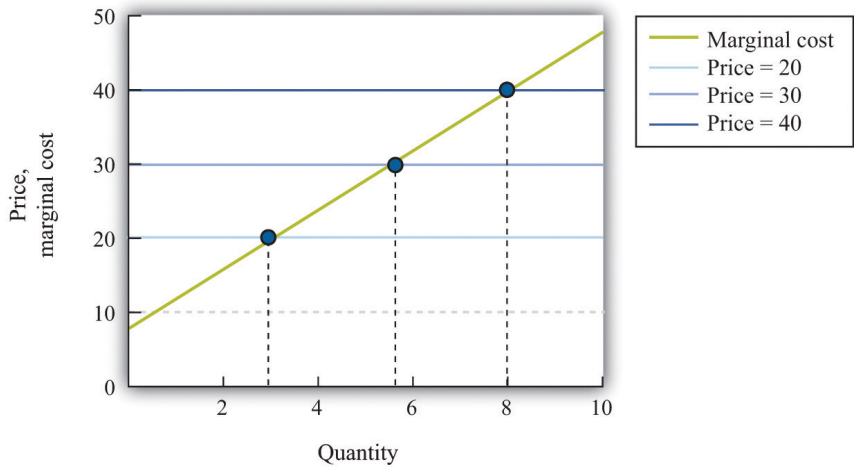


Figure 7.6.2: The Supply Curve of an Individual Firm. The firm chooses its quantity such that price equals marginal cost, which implies that *the marginal cost curve of the firm is the supply curve of the firm.*

key takeaways

- A perfectly competitive market has a large number of buyers and sellers of exactly the same good.
- In a perfectly competitive market, an individual firm faces a demand curve with infinite elasticity.
- In a perfectly competitive market, the firm does not set a price but chooses a level of output such that marginal cost equals the market price.

check your understanding

1. Explain why the demand curve a firm faces in a perfectly competitive market is horizontal even though the market demand curve is not horizontal.
2. Why is the cost of one unit $60 in Table 7.4.1 "Marginal Cost" but only $22 in Table 7.6.1 "Costs of Production: Increasing Marginal Cost"?

This page titled 7.6: The Supply Curve of a Competitive Firm is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

7.7: End-of-Chapter Material 7.7: End-of-Chapter MaterialEdit Update titlecancel

In Conclusion

Choosing the right price is one of the hardest problems that a manager faces. It is also one of the most consequential: few other decisions have such immediate impact on the health and success of a firm. It is hardly surprising that firms devote considerable resources to deciding on the price to set. Even though firms operate in many different market settings, our analysis of markup pricing is very general: it applies to firms in all sorts of different circumstances. It is thus a powerful tool for understanding the behavior of firms in an economy.

One goal of this textbook is to help you make good decisions, both in your everyday life and in your future careers. In this chapter, we have set out the principles of how prices *should* be set, assuming that the goal of a manager is to make as much profit for a firm as possible. It does not necessarily follow, however, that this is how managers actually behave in real life. Does this chapter just describe a make-believe world of economists or does it also describe how prices are set in the real business world?

The answer is a bit of both. Managers must think carefully about costs and demand when setting prices. Market research firms routinely investigate consumers' price sensitivities and estimate elasticities. At the same time, pricing decisions are sometimes more haphazard than this chapter might suggest. In practice, managers often use rules of thumb or standard markups that are not necessarily solidly based on the elasticity of demand.

There is one reason to think that managers do not stray too far from the prices that maximize their firms' profits, however. Business is a competitive affair, and firms that make poor decisions will often not survive in the marketplace. If a firm consistently sets the wrong price, it will make less money than its competitors and will probably be forced out of business or taken over by another firm that can do a better job of management. The marketplace imposes a harsh discipline on badly managed firms, but the end result is —usually at least—a more efficient and better-functioning economy.

Key Link

- College pricing: http://www.collegeboard.com/student/pay/add-it-up/4494.html

exercises

1. We suggested that a grocery store could conduct an experiment to find a demand curve by charging a different price each week for some product.
    1. Do you think that technique would be more accurate for a perishable good, such as milk, or for a nonperishable good, such as canned tuna? Why?
    2. Do you think the technique would be more accurate if the firm announced the price each week in advance or if it just let customers discover the different prices when they came to the store? Why?
2. Extend Table 7.6.1 "Costs of Production: Increasing Marginal Cost" for output levels 6–10. What does Table 7.6.1 "Costs of Production: Increasing Marginal Cost" look like if the fixed cost is $100?
3. Suppose your company is selling a product that is an inferior good. What do you think will happen to the demand curve facing your firm when the economy goes into recession?
4. Suppose you are a producer of DVDs and imagine that producers of DVD players decrease their prices. What do you think will happen to the demand curve you face?
5. If you were running a fast-food restaurant, what factors would you take into account in setting a price for burgers?
6. Suppose a monopolist could produce an extra unit at zero marginal cost and, at the current price, faces a demand curve with an −(elasticity of demand) of 2. Should the monopolist raise or lower its price to make more profit?
7. Suppose that instead of maximizing profit, the monopolist in Question 6 wants to maximize revenues. Would it behave any differently? What if the marginal cost was positive?
8. If the price of steak is $25.00 a pound and the −(elasticity of demand) is 2, what decrease in price would lead the quantity sold to increase by 4 percent?
9. Explain why marginal revenue must be less than the price when a firm faces a downward-sloping demand curve.
10. A monopolist is maximizing profit. Perhaps due to an innovation in some other product line, he finds that the elasticity of demand for his product is lower. What will this change in the elasticity of demand due to the profit of the monopolist? How will the monopolist respond to this change?
11. The following is an excerpt from an article in the Singaporean newspaper, the *Straits Times*:

Singaporeans with a sweet tooth could soon find themselves paying more for their favourite treats, as bakers and confectioners buckle under soaring sugar prices.

Since March last year, the price of white sugar has shot up by 70 per cent, according to the New York Board of Trade. As if that didn't make life difficult enough for bakers, butter and cheese prices have also risen, by 31 per cent and 17 per cent respectively.

The increases have been caused by various factors: a steep drop in Thailand's sugarcane production due to drought, higher sea freight charges, increasing demand from China's consumers for dairy products and the strong Australian and New Zealand dollar.

For the consumer in Singapore, what this may eventually boil down to is a more expensive bag of cookies, with prices at some bakeries expected to rise between 10 and 20 per cent.

[The owner of a Singapore bakery, Mr. Leong Meng Pock], said that he intends to raise prices possibly as early as next month. A sugared doughnut at his shop sells for 50 cents [about US$0.30] and a slice of Black Forest cake for $1.80 [about US$1.13], prices that have remained unchanged since 1990. Next month, the doughnuts may go up to 60 cents and the Black Forest cake to $2.

Said Mr Leong: "In Singapore, you have bread and cake prices that are at least 10 years old. This is especially true for the HDB [government-subsidized housing] neighborhoods, where customers are very price-sensitive."See http://straitstimes.asia1.com.sg.

1. Do you think bakers face a demand curve that is relatively elastic or relatively inelastic? Why?
2. What has happened to their marginal cost?

12. Explain the difference between a shift in the demand curve and movement along a demand curve.
13. If you observe the price of a product, then you can infer the marginal cost of the product if and only if the market is competitive. Explain.

Spreadsheet Exercise

1. Suppose that the cost function for a product is given by total costs = 100 + 2,000 × quantity. Create a spreadsheet to calculate the costs for different levels of output and use it to produce a graph like Figure 7.7.1.

Economics Detective

1. What prices are your local gas stations currently charging for gas? Do the stations generally have the same price for gas? If not, what would explain the differences in prices they set? Do the stations charge the same price all the time or does the price change? When the price changes, what might be the reason for that change?
2. Think about the college you are attending. What determines the profit of the college—what are its revenues and what are its costs? What is tuition at your college? Would you advocate an increase or a decrease in tuition rates to increase revenue at your college? What factors determine the elasticity of demand faced by your college?

This page titled 7.7: End-of-Chapter Material is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

8: Why Do Prices Change?

8.1: Prices in the News
8.2: Market Supply and Market Demand
8.3: Using the Supply-and-Demand Framework
8.4: Another Perspective on Changing Prices
8.5: Three Important Markets
8.6: Beyond Perfect Competition
8.7: End-of-Chapter Material
8.8: Appendix- Algebraic Presentation of Supply and Demand

This page titled 8: Why Do Prices Change? is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

8.1: Prices in the News 8.1: Prices in the NewsEdit Update titlecancel

Here are two recent headlines. The first discusses beer prices in England.

Price of a Pint "Could Rise 60%"

The average price of a pint of beer could hit £4 [about $8]…

Scottish & Newcastle today forecast "material price increases" next year. The brewer, which sells three of the top 10 beer brands in Europe including Kronenbourg and Foster's, is also reviewing its supply chain in a bid to cut costs.

Industry experts say the cost of an average pint will rise by at least 15p, although some are now predicting rises of up to 60%.…

"It is a bleak time for everyone," said Iain Lowe, research and information manager at Camra [The Campaign for Real Ale]. "These price rises have been predicted for a long time."Teena Lyons, "The Price of a Pint Could Rise 60%," *Guardian Unlimited*, November 20, 2007, accessed January 27, 2011, http://www.guardian.co.uk/business/2007/nov/20/fooddrinks.foodanddrink?gusrc=rss&feed=networkfront.

The second concerns sales of baseball merchandise in Detroit at the start of the 2010 season.

Tigers' Merchandise Off to Roaring Start

Opening Day is still a day away for Detroit's baseball fans, but its impending arrival already is generating its share of Detroit Tiger retail hits.

Thousands of fans have flooded Comerica Park's pro shop and other Metro Detroit sporting outlets in anticipation of Friday afternoon's home game against the Cleveland Indians, snapping up Tigers jerseys, T-shirts and hats bearing the surnames Cabrera, Verlander and Damon, even Granderson—Tigers old and new.

…

Though it's too soon to tell which Tigers will prove most popular at the checkout line, former players have been relegated to the discount bin.

"Retailers take a pretty aggressive stand," Powell said. Most shops have marked down jerseys and T-shirts branded with ex-Tigers between 25 and 50 percent.

"Granderson and Polanco—we discount the price," said Brian King, owner of Sports Authentics in Rochester Hills. "Unfortunately, you can't take the names off the back.""Tigers' Merchandise off to Roaring Start," *The Detroit News*, April 8, 2010, accessed January 27, 2011, http://www.detnews.com/article/20100408/BIZ/4080349/1129/Tigers-merchandise-off-to-roaring-start.

We could have picked thousands of other examples. If you search Google's news aggregator on any day with a string such as "an increase in the price of" you will find dozens, perhaps hundreds, of recent news articles that contain this phrase. Our task in this chapter is to see where all these price changes come from and what they imply for other economic variables, such as the quantity of these goods traded.

To see how good you are at this, think about these two stories. Can you explain why the price of beer increased? Can you explain why "Granderson" T-shirts are being sold at discounted prices? What do you think happened to the quantity of beer sold as the price increased? What do you think happened to the quantity of T-shirts sold as the price decreased?

Understanding the sources and consequences of changing prices in the economy is one of the most important tasks of an economist. There is an almost endless list of such analyses in economics. In fact, most of the applications in this textbook ultimately come down to understanding, explaining, and predicting changes in prices. The question that motivates this chapter is so important that we have chosen it as the title:

Why do prices change?

Road Map

All prices in the economy are ultimately chosen by someone. Sometimes they are chosen by marketing or pricing managers in big companies. Sometimes they are chosen by bidders in an auction. Sometimes they are agreed on by the buyer and the seller after bargaining. We discuss these choices in Chapter 6 and Chapter 7. Yet we can often make good predictions about prices without looking closely at the individual decision making of buyers and sellers by summarizing their decisions with demand curves and supply curves. Building on the ideas of the individual demand curve and a firm's supply curve for a good or service, we develop the ideas of supply and demand for an entire market. Individual demand and supply curves are introduced in Chapter 4 and Chapter 7. In this chapter, we look at the trade that occurs between firms and households or among different firms in the economy. In the business world, these are called business-to-consumer (B2C) and business-to-business (B2B) trade, respectively. The market demand and supply curves that we derive allow us to predict what will happen to prices and quantities traded when there are changes that influence the market. Chapter 6 also looks at supply and demand in the context of trade between individuals.

An old joke says that you can ask an economist any question, and he will always give the same answer: supply and demand. Yet—strictly speaking—we are supposed to use the supply-and-demand framework only when we are talking about a competitive market—a market in which a homogeneous good is traded by a large number of buyers and sellers. In practice, economists and others use the framework all the time in settings where these assumptions do not hold. Perhaps surprisingly, this can be a completely reasonable thing to do, and we explain why.

Once we understand why prices change, we consider the implications of these price changes for the functioning of the economy. Prices convey information to both producers and consumers. When the price of a good or a service increases, it encourages consumers to consume less and producers to produce more. As we will see, this means that prices play a crucial role in allocating resources in the economy.

We finish this chapter by looking at three very significant markets in the economy: the labor market, the credit market, and the foreign exchange market. Understanding how these three markets work is necessary for a good understanding of both microeconomics and macroeconomics.

This page titled 8.1: Prices in the News is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

8.2: Market Supply and Market Demand 8.2: Market Supply and Market DemandEdit Update titlecancel

Learning Objectives

1. How is the market demand curve derived?
2. What is the slope of the market demand curve?
3. How is the market supply curve derived?
4. What is the slope of the market supply curve?
5. What is the equilibrium of a perfectly competitive market?

We begin the chapter with the **individual demand curve**—sometimes also called the household demand curve—that is based on an individual's choice among different goods. (In this chapter, we use the terms *individual* and *household* interchangeably.) We show how to build the market demand curve from these individual demand curves. Then we do the same thing for supply, showing how to build a market supply curve from the supply curves of individual firms. Finally, we put them together to obtain the market equilibrium.

Market Demand

Figure 8.2.1 is an example of a household's demand for chocolate bars each month. Taking the price of a chocolate bar as given, as well as its income and all other prices, the household decides how many chocolate bars to buy. Its choice is represented as a point on the household's demand curve. For example, at \$5, the household wishes to consume five chocolate bars each month. The remainder of the household income—which is its total income minus the \$25 it spends on chocolate—is spent on other goods and services. If the price decreases to \$3, the household buys eight bars every month. In other words, the quantity demanded by the household increases. Equally, if the price of a chocolate bar increases, the quantity demanded decreases. This is the law of demand in operation.

One way to summarize this behavior is to say that the household compares its **marginal valuation** from one more chocolate bar to price. The marginal valuation is a measure of how much the household would like one more chocolate bar. The household will keep buying chocolate bars up to the point where

$$marginal\ valuation\ =\ price.$$

Toolkit: Section 31.1

You can review the foundations of individual demand and the idea of marginal valuation in the toolkit.

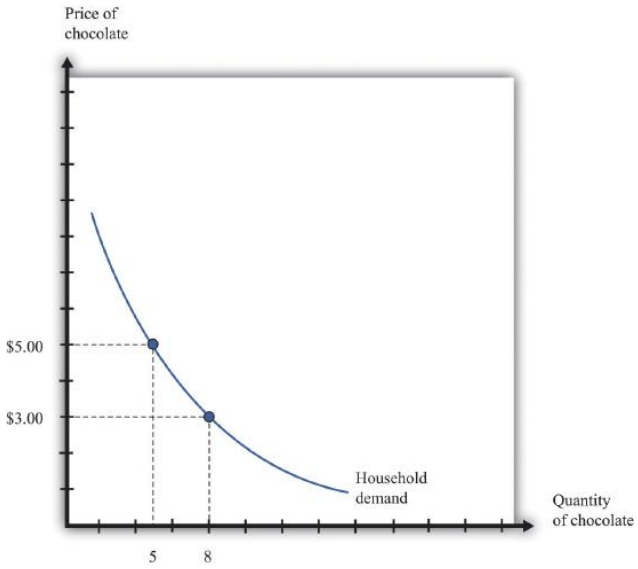


Figure 8.2.1: The Demand Curve of an Individual Household. The household demand curve shows the quantity of chocolate bars demanded by an individual household at each price. It has a negative slope: higher prices lead people to consume fewer chocolate bars.

Table 8.2.1: Individual and Market Demand

Price ($)	Household 1 Demand	Household 2 Demand	Market Demand
1	17	10	27
3	8	3	11
5	5	2	7
7	4	1.5	5.5

In most markets, many households purchase the good or the service traded. We need to add together all the demand curves of the individual households to obtain the **market demand curve**. To see how this works, look at Table 8.2.1 "Individual and Market Demand" and Figure 8.2.2. Suppose that there are two households. Part (a) of Figure 8.2.2 shows their individual demand curves. Household 1 has the demand curve from Figure 8.2.1. Household 2 demands fewer chocolate bars at every price. For example, at $5, household 2 buys 2 bars per month; at $3, it buys 3 bars per month. To get the market demand, we simply add together the demands of the two households at each price. For example, when the price is $5, the market demand is 7 chocolate bars (5 demanded by household 1 and 2 demanded by household 2). When the price is $3, the market demand is 11 chocolate bars (8 demanded by household 1 and 3 demanded by household 2). When we carry out the same calculation at every price, we get the market demand curve shown in part (b) of Figure 8.2.2.

Toolkit: Section 31.9

You can review the market demand curve in the toolkit.

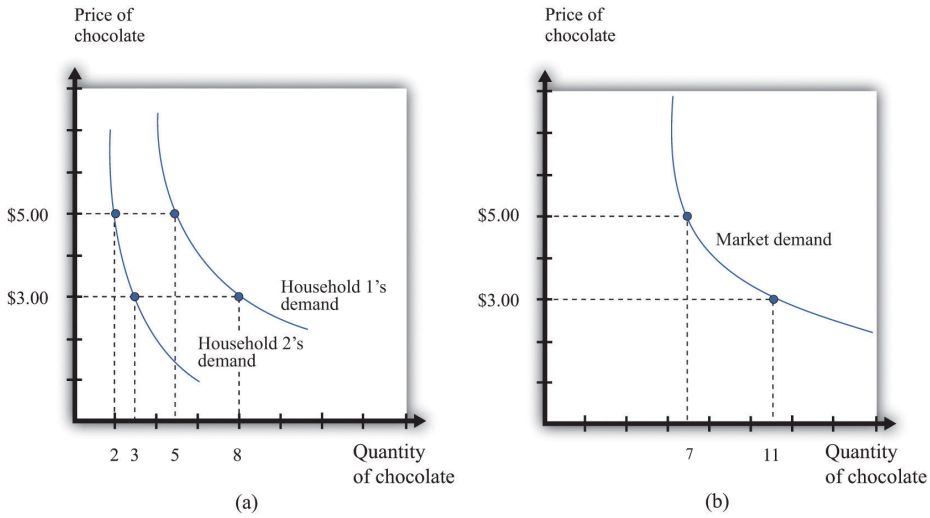


Figure 8.2.2: Market Demand. Market demand is obtained by adding together the individual demands of all the households in the economy.

Because the individual demand curves are downward sloping, the market demand curve is also downward sloping: the law of demand carries across to the market demand curve. As the price decreases, each household chooses to buy more of the product. Thus the quantity demanded increases as the price decreases. Although we used two households in this example, the same idea applies if there are 200 households or 20,000 households. In principle, we could add together the quantities demanded at each price and arrive at a market demand curve.

There is a second reason why demand curves slope down when we combine individual demand curves into a market demand curve. Think about the situation where each household has a **unit demand curve**: that is, each individual buys at most one unit of the product. As the price decreases, the number of individuals electing to buy increases, so the market demand curve slopes down.See Chapter 4 and Chapter 6 for discussions of unit demand. In general, both mechanisms come into play.

- As price decreases, some households decide to enter the market; that is, these households buy some positive quantity other than zero.
- As price decreases, households increase the quantity that they wish to purchase.

When the price decreases, there are more buyers, and each buyer buys more.

Market Supply

In a **competitive market**, a single firm is only one of the many sellers producing and selling exactly the same product. The demand curve facing a firm exhibits **perfectly elastic demand**, which means that it sets its price equal to the price prevailing in the market, and it chooses its output such that this price equals its **marginal cost** of production.At the end of Chapter 7, we derive the supply curve of a firm in a competitive market. If it were to try to set a higher price, it could not sell any output at all. If it were to set a lower price, it would be throwing away profits. Thus, for a competitive firm, the quantity produced satisfies this condition:

$$price = marginal\ cost.$$

Toolkit: Section 31.2

For more information on elasticity, see the toolkit.

We typically expect that marginal cost will increase as a firm produces more output. Marginal cost is the cost of producing one *extra* unit of output. The cost of producing an additional unit of output generally increases as firms produce a larger and larger quantity. In part, this is because firms start to hit constraints in their capacities to produce more product. For example, a factory might be able to produce more output only by running extra shifts at night, which require paying higher wages.

If marginal cost is increasing, then we know the following:

- Given a price, there is only one level of output such that price equals marginal cost.
- As the price increases, a firm will produce more.

Indeed, the supply curve of an individual firm is the same as its marginal cost curve.

Figure 8.2.3 illustrates the supply curve for a firm. A firm supplies seven chocolate bars at $3 and eight chocolate bars at $5. From this we can deduce that the marginal cost of producing the seventh chocolate bar is $3. Similarly, the marginal cost of producing the eighth chocolate bar is $5.

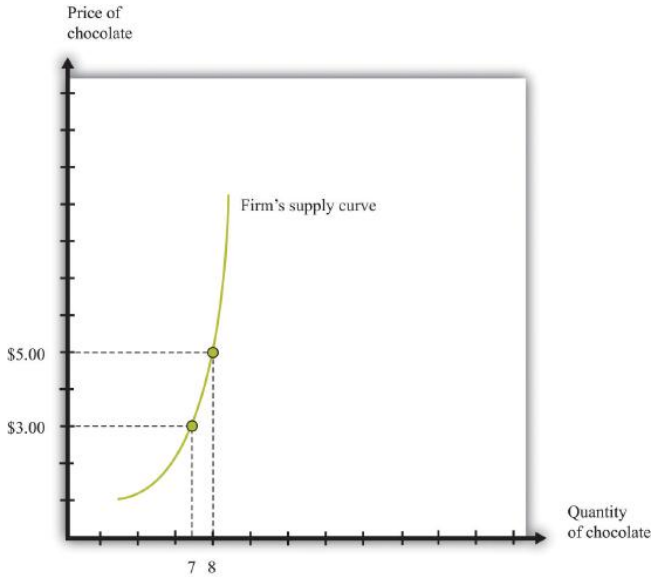


Figure 8.2.3: The Supply Curve of an Individual Firm. A firm's supply curve, which is the same as its marginal cost curve, shows the quantity of chocolate bars it is willing to supply at each price.

Just as the market demand curve tells us the total amount demanded at each price, the **market supply curve** tells us the total amount supplied at each price. It is obtained analogously to the market demand curve: at each price we add together the quantity supplied by each firm to obtain the total quantity supplied at that price. If we perform this calculation for every price, then we get the market supply curve. Figure 8.2.4 shows an example with two firms. At $3, firm 1 produces 7 bars, and firm 2 produces 3 bars. Thus the total supply at this price is 10 chocolate bars. At $5, firm 1 produces 8 bars, and firm 2 produces 5 bars. Thus the total supply at this price is 13 chocolate bars.

The market supply curve is increasing in price. As price increases, each firm in the market finds it profitable to increase output to ensure that price equals marginal cost. Moreover, as price increases, firms who choose not to produce and sell a product may be induced to enter into the market.A similar idea is in Chapter 6, where we show how to add together unit supply curves to obtain a market supply curve.

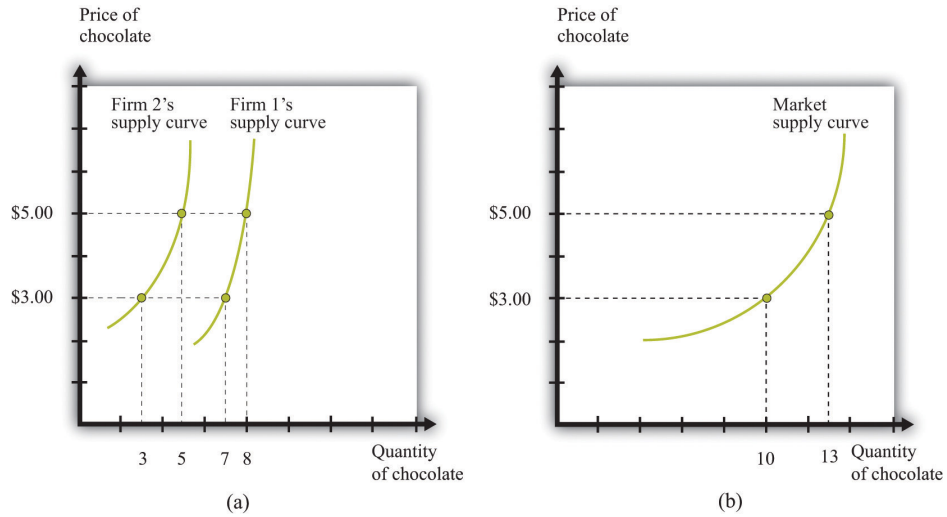


Figure 8.2.4: Market Supply. Market supply is obtained by adding together the individual supplies of all the firms in the economy.

In general, both mechanisms come into play. The market supply curve slopes up for two reasons:

1. As the price increases, more firms decide to enter the market—that is, these firms produce some positive quantity other than zero.
2. As the price increases, firms increase the quantity that they wish to produce.

When the price increases, there are more firms in the market, and each firm produces more.

Market Equilibrium

In a perfectly competitive market, we combine the market demand and supply curves to obtain the **supply-and-demand** framework shown in Figure 8.2.5. The point where the curves cross is the market **equilibrium**.The definition of equilibrium is also presented in Chapter 6. At this point, there is a perfect match between the amount that buyers want to buy and the amount that sellers want to sell. The term *equilibrium* refers to the balancing of the forces of supply and demand in the market. At the equilibrium price, the suppliers of a good can sell as much as they wish, and demanders of a good can buy as much of the good as they wish. There are no disappointed buyers or sellers.

Toolkit: Section 31.9

You can review the definition and meaning of equilibrium in the supply-and-demand framework in the toolkit.

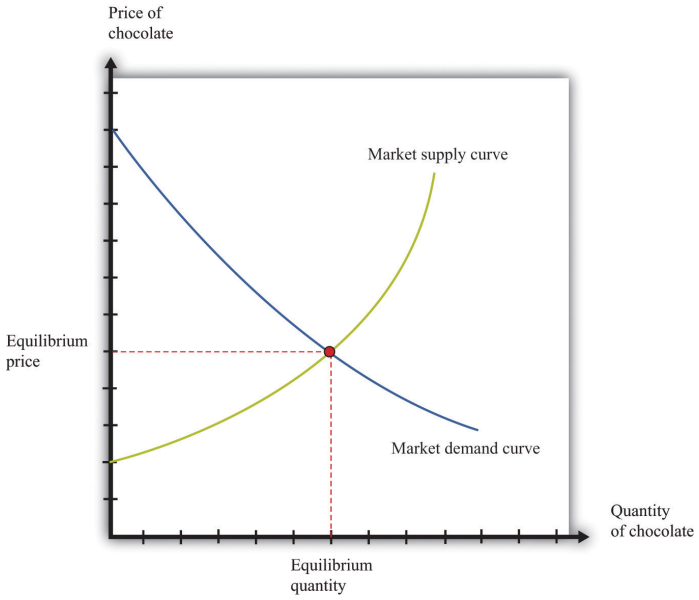


Figure 8.2.5: Market Equilibrium. In a competitive market, the equilibrium price and the equilibrium quantity are determined by the intersection of the supply and demand curves.

Because the demand curve has a negative slope and the supply curve has a positive slope, supply and demand will cross once. Both the equilibrium price and the equilibrium quantity will be positive. (More precisely, this is true as long as the vertical intercept of the demand curve is larger than the vertical intercept of the supply curve. If this is not the case, then the most that any buyer is willing to pay is less than the least any seller is willing to accept and there is no trade in the market.)

Table 8.2.2: Market Equilibrium: An Example

Price ($)	Market Supply	Market Demand
1	5	95
5	25	75
10	50	50
20	100	0

Table 8.2.2 "Market Equilibrium: An Example" shows an example of market equilibrium with market supply and market demand at four different prices. The equilibrium occurs at $10 and a quantity of 50 units. The table is based on the following equations:

$$market\ demand\ =\ 100 - 5 \times price$$

and

$$market\ supply\ = 5 \times price.$$

Equations such as these and diagrams such as Figure 8.2.5 are useful to economists who want to understand how the market works. Keep in mind, though, that firms and households in the market do not need any of this information. This is one of the beauties of the market. An individual firm or household needs to know only the price that is prevailing in the market.

Reaching the Market Equilibrium

Economists typically believe that a perfectly competitive market is likely to reach equilibrium for several reasons.

- If the prevailing price is different from the equilibrium price, then there will be an imbalance between demand and supply, which gives buyers and sellers an incentive to behave differently. For example, if the prevailing price is less than the equilibrium price, demand will exceed supply. Disappointed buyers might start bidding the price up, or sellers might realize they could charge a higher price. The opposite is true if the prevailing price is too high: suppliers might be tempted to try decreasing prices, and buyers might look for better deals. These are informal stories because the supply and demand curves are based on the idea that firms and consumers take prices as given. Still, the idea that there will be pressure on prices away from equilibrium is a plausible one.
- There is strong support for market predictions in the evidence from experimental markets.In Chapter 6, we explain that a double oral auction, in which buyers and sellers meet individually and bargain over prices, typically yields results very close to the market outcome in Figure 8.2.5.
- The supply-and-demand framework generally provides reliable predictions about the movement of prices.

key takeaways

- The market demand curve is obtained by adding together the demand curves of the individual households in an economy.
- As the price increases, household demand decreases, so market demand is downward sloping.
- The market supply curve is obtained by adding together the individual supply curves of all firms in an economy.
- As the price increases, the quantity supplied by every firm increases, so market supply is upward sloping.
- A perfectly competitive market is in equilibrium at the price where demand equals supply.

check your understanding

1. In Table 8.2.2 "Market Equilibrium: An Example", market supply was equal to 5 × price. Suppose instead that market supply = 15 × price. Would the equilibrium price still be $10? If not, construct a new column in the table and find the new equilibrium price.
2. Explain why supply and demand cross only once. Do they always cross at a positive price?

This page titled 8.2: Market Supply and Market Demand is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

8.3: Using the Supply-and-Demand Framework 8.3: Using the Supply-and-Demand FrameworkEdit Update titlecancel

Learning Objectives

1. Why do market prices increase and decrease?
2. How can I predict what is going to happen to prices?

Economists are often asked to make predictions about the effects of events on economic outcomes. They do so by using the supply-and-demand framework. To use this framework, we must first distinguish between those things that we take as given (**exogenous** variables) and those that we seek to explain (**endogenous** variables).

Toolkit: Section 31.16

An exogenous variable is something that comes from outside a model and is not explained in our analysis. An endogenous variable is one that is explained within our analysis. When using the supply-and-demand framework, price and quantity are endogenous variables; everything else is exogenous.

A Shift in Supply: Beer Prices in Britain

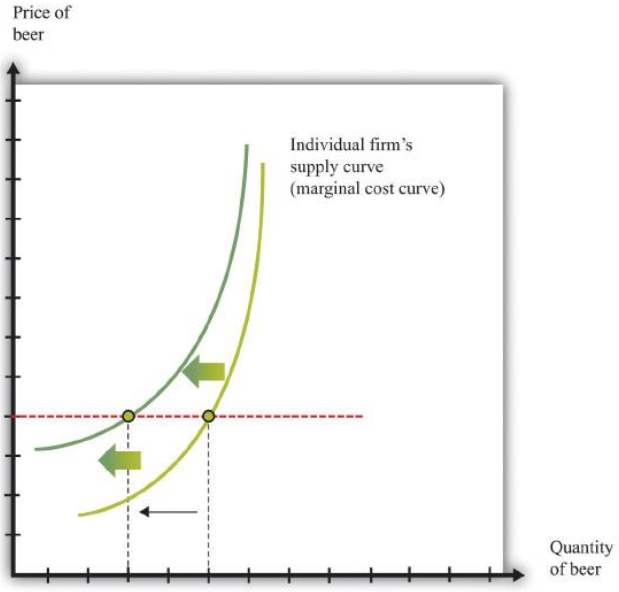


Figure 8.3.1: A Shift in the Supply Curve of an Individual Firm. An increase in marginal cost leads a firm to produce less output at any given price. This means that a firm's supply curve shifts upward and to the left.

When we quoted the British newspaper article about beer prices in the chapter introduction, we omitted some sentences. The first sentence reads, in full: "The average price of a pint of beer could hit £4 [pounds sterling] after poor weather forced up the price of hops." A few sentences later the article states: "Hop farmers have not seen any price rises for years, but the appalling summer has finally forced the prices up." According to the article, the price of beer is increasing because the price of hops has increased.

This makes intuitive sense, but it is worth understanding the exact chain of reasoning here. Hops are a key ingredient in the production of beer. An increase in the price of hops means an increase in the cost of producing beer. More precisely, the marginal cost of producing beer increases. The typical beer producer decides how much to produce by observing this decision rule:

$$price = marginal\ cost.$$

If marginal cost increases, then, at the existing price, the producer will find that price is now less than marginal cost. To bring price back in line with marginal cost, the producer will have to produce a smaller quantity. In fact, at any given price, an increase in marginal cost leads to a reduction in output (Figure 8.3.1). The supply curve of an individual firm shifts to the left.

The increase in the price of hops affects all firms in the market. Each firm sees an increase in its marginal cost of production, so each firm produces less output at a given price: the shift in supply shown in Figure 8.3.1 applies to all firms in the market. Figure 8.3.2 shows the outcome in the market. Because all the individual supply curves shift to the left, the market supply curve likewise shifts to the left. At any given price, firms supply less beer to the market. From the figure, we see that the higher price of hops leads to an increase in the price of beer and a reduction in the quantity of beer produced and sold.

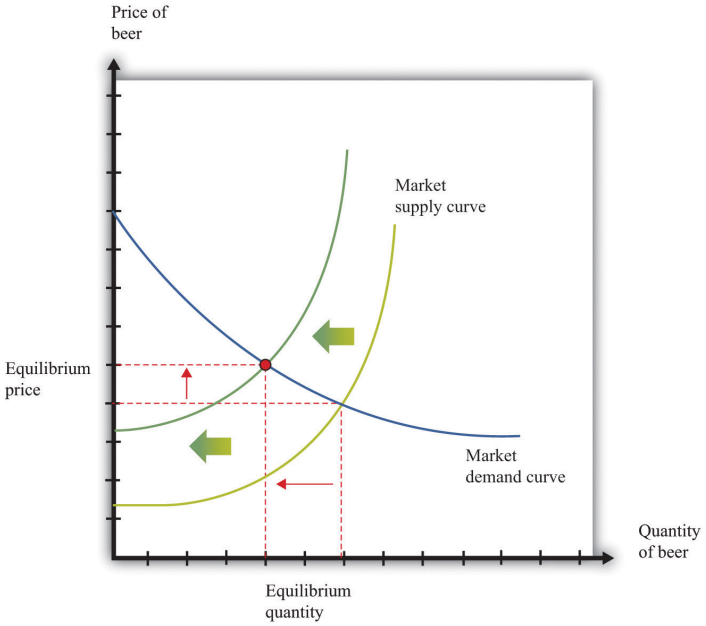


Figure 8.3.2: A Shift in Market Supply

An increase in the price of hops causes all beer producers to produce less at any given price. This means that the market supply curve shifts to the left. The consequence is an increase in the equilibrium price and a decrease in the equilibrium quantity.

For the individual producer, what does this mean? The producer sees an increase in marginal cost. In the new equilibrium, the producer also obtains a higher price. However, the increase in price is not as big as the increase in marginal cost. Because the producer sets price equal to marginal cost, each individual brewer still produces less. We show this in Figure 8.3.3.

1. The price of hops, an input into beer production, has increased, which increases the marginal cost of producing beer.
2. At each given price, beer producers want to supply less beer: the firm supply curve shifts to the left.
3. Because all the individual supply curves shift to the left, the market supply curve also shifts to the left.
4. The beer market reaches a new equilibrium with a higher price and smaller quantity of beer produced and consumed.

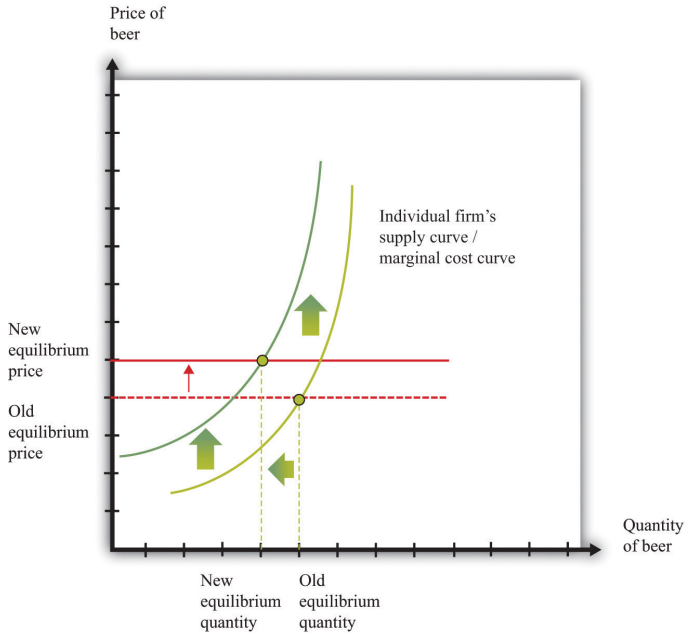


Figure 8.3.3: The New Equilibrium from the Perspective of an Individual Firm

Following the increase in the price of hops, the equilibrium price of beer increases. An individual firm ends up with higher marginal cost but also receives a higher price for beer. Because the increase in price is smaller than the increase in marginal cost, beer production still decreases.

Comparative Statics

The approach that we used here is an illustration of a general technique used by economists to explain changes in prices and quantities and to make predictions about what will happen to market prices.

Toolkit: Section 31.16

Comparative statics is a technique that allows us to describe how market equilibrium prices and quantities depend on exogenous events. As such, much of economics consists of exercises in comparative statics. In a comparative statics exercise, you must do the following:

1. Begin at an equilibrium point where the quantity supplied equals the quantity demanded.
2. Based on the description of the event, determine whether the change in the exogenous factor shifts the market supply curve or the market demand curve.
3. Determine the direction of this shift.
4. After shifting the curve, find the new equilibrium point.
5. Compare the new and old equilibrium points to predict how the exogenous event affects the market.

The most difficult part of a **comparative statics** exercise is to determine, from the description of the economic problem, which curve to shift—supply or demand. Once you determine which curve is shifting, then it is only a matter of using the supply-and-demand framework to find the new equilibrium. The final step is to compare the new equilibrium point (the new crossing of supply and demand) with the original point. With this comparison, you can predict what will happen to equilibrium prices and quantities when something exogenous changes.

A Shift in Demand

Let us try this technique again. Recall the second story from the chapter introduction about Detroit Tiger merchandise. In that story, we learned that "most shops have marked down jerseys and T-shirts branded with ex-Tigers between 25 and 50 percent." Figure 8.3.4 shows the demand of a typical Detroit household's demand for Granderson shirts. Now that Granderson has left the Tigers for the New York Yankees, the household's marginal valuation for these shirts is lower. At any given price, a household wants to purchase fewer shirts, so the household's demand curve shifts to the left.

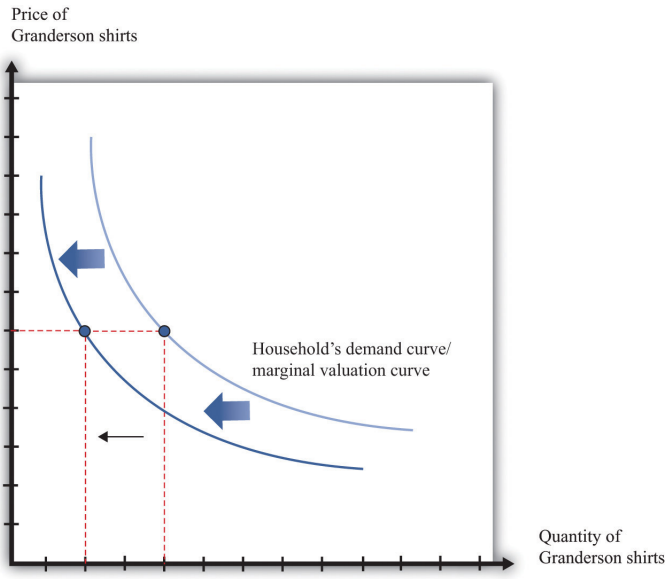


Figure 8.3.4: Shifts in Household Demand

A decrease in the marginal valuation of Granderson T-shirts leads a household to demand a smaller quantity at any given price. This means that a household's demand curve shifts to the left.

We would expect that this shift in demand would apply to most households that contain Detroit Tigers fans. If we now add all the demand curves together, we get the market demand curve. The market demand curve shifts to the left (Figure 8.3.5). The end result is that we expect to see a decrease in the price of T-shirts—that is, the retailers put them in the discount bins—and also a decrease in demand.

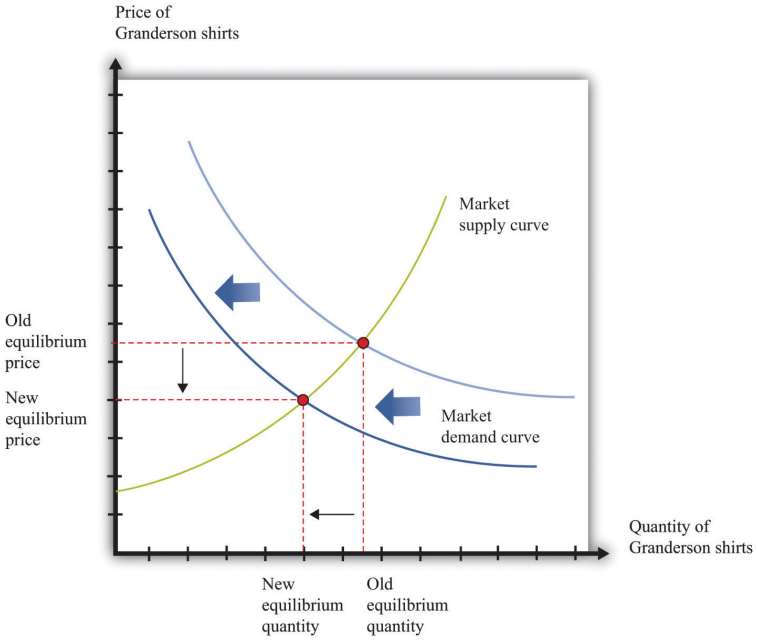


Figure 8.3.5: Shifts in Market Demand. The decrease in demand causes both the equilibrium price and the equilibrium quantity of T-shirts to decrease.

Learning about the Slopes of the Supply and Demand Curves

Comparing our beer and T-shirt examples, we see that the quantity demanded decreased in both examples. In the first case, price increased; in the second case, price decreased. We can understand the difference by using the supply-and-demand framework. In the Detroit Tigers example, there is a decrease in the price of shirts and in the quantity sold. This might seem like a violation of the law of demand, which tells us that when price decreases, the quantity demanded increases. The explanation comes directly from Figure 8.3.5.

- Market demand is downward sloping and obeys the law of demand.
- Both equilibrium price and equilibrium quantity decrease after the departure of Granderson to the New York Yankees.

Curtis Granderson's move leads to a *shift in the demand curve* and a *movement along the supply curve*. The law of demand, by contrast, applies to the movement along a demand curve.

Shifts in a Curve versus Movements along a Curve

Understanding the distinction between moving along a curve (either supply or demand) and shifting a curve is the hardest part about learning to use the supply-and-demand framework. Journalists and others frequently are confused about this—and no wonder. It requires practice to learn how to use supply and demand properly.

Let's look at another example. An article in the British newspaper the *Guardian* reported about sales of beef when the news came out that eating beef might carry a risk of bovine spongiform encephalopathy (BSE), better known as mad cow disease. On November 1, 2000, the newspaper wrote, "Beef sales did drop after the link between BSE and deaths in humans was circumstantially established in 1996, but they have recovered as prices have fallen."[1] "First Beef—Now Lamb to the Slaughter?" Analysis, *Guardian*, November 1, 2000, accessed February 4, 2011, http://www.guardian.co.uk/uk/2000/nov/01/bse?INTCMP=SRCH.

The exogenous event here is the medical news about beef and mad cow disease. Presumably, this primarily affects the demand for beef: consumers decide to eat less beef and more of other products—such as chicken and pork. The demand curve for beef shifts to the left. As we saw in the T-shirt example, a leftward shift of the demand curve has two consequences: price decreases, and the quantity demanded and supplied also decreases. Thus the conclusion that the news should lead to a decrease in beef sales is perfectly consistent with our supply-and-demand analysis, as well as with common sense.

But what about the second part of the sentence? The article claims that beef sales "have recovered as prices have fallen." This is not consistent with our supply-and-demand analysis. The decrease in prices is intimately connected with the decrease in quantity: both were caused by the health news. They are two sides of the same coin, so it does not make sense to use the decrease in prices to explain a recovery in beef sales.

In fact, you should be able to convince yourself that an increase in beef sales together with a decrease in prices (as asserted by the article) would require a *rightward shift of the supply curve*. (Draw a diagram to make sure you understand this.) It seems unlikely that health concerns about beef led cattle farmers to *increase* their production of beef. It is hard to escape the conclusion that the journalist became confused about shifts in the demand curve and movements along the curve.

Estimating Demand and Supply Curves

Comparative statics allows us to make qualitative predictions about prices and quantities. Given an exogenous shock in a market, we can determine whether (1) the price is likely to increase or decrease and (2) the quantity bought and sold is likely to increase or decrease. Often, though, we would like to be able to do more. We would like to be able to make some predictions about the *magnitudes* of the changes.

Figuring out what will happen to equilibrium prices and quantities requires economists to know the shapes of supply and demand. When the supply curve shifts, we need to know about the slope of the demand curve to predict the impact on price and quantity. When the demand curve shifts, we need to know about the slope of the supply curve to predict the impact on price and quantity. More precisely, we need measures of the **elasticity** of demand and of supply.

How do economists learn about these elasticities? The answer, perhaps surprisingly, is through the logic of comparative statics. For example, suppose the supply curve does not move, but the demand curve shifts around a lot. As the demand curve shifts, we observe different combinations of prices and quantities. Part (a) of Figure 8.3.6 shows this in a supply-and-demand diagram. The different points that we observe are points on the supply curve. If the demand curve shifts but the supply curve does not, we eventually gather data on the supply curve. We can use these data to come up with estimates of the **price elasticity of supply**.

> **Toolkit: Section 31.2**
>
> The price elasticity of supply equals the percentage change in the quantity supplied divided by the percentage change in price.

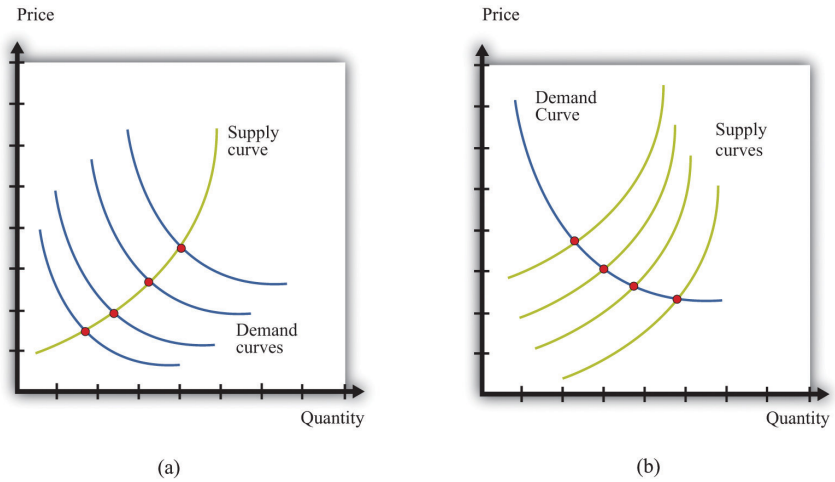


Figure 8.3.6: Finding the Elasticities of the Supply and Demand Curves

Economists estimate the elasticities of supply and demand curves by looking for situations in which one curve is relatively stable while the other one is moving. What we actually observe are the equilibrium points. Movements in the demand curve (a) mean that the equilibrium points trace out the supply curve; movements in the supply curve (b) allow us to observe the demand curve. In most real-life cases, both curves move, and economists use sophisticated statistical techniques to tease apart shifts in supply from shifts in demand.

Part (b) of Figure 8.3.6 shows the opposite case, where demand is stable and the supply curve is shifting. In this case, the data that we observe are different points on the demand curve. We can use this information to estimate the **price elasticity of demand**, which is the percentage change in the quantity demanded divided by the percentage change in price. It is important to note that we are speaking here about the elasticity of the *market* demand curve, not the elasticity of the demand curve facing an individual firm.

This sounds straightforward in theory, but it is difficult in practice. Economic data are messy. Typically, both the demand curve and the supply curve are shifting simultaneously. If economists had access to controlled environments, perhaps like a biochemist does, we could "shift the demand curve" and see what happens in the laboratory. Occasionally, we get lucky. Sometimes we can isolate a particular event that we know is likely to shift only one of the curves. This is sometimes called a **natural experiment**. Most of the time, however, we are not so lucky. Economists and statisticians have come up with sophisticated statistical techniques to disentangle shifts in demand and supply in these circumstances. Chapter 7 discusses how a firm can use a similar technique to learn about the demand curve that it faces. Chapter 11 discusses the difficulties of measuring the demand curve for labor.

key takeaways

- Changes in prices come from shifts in market supply, market demand, or both.
- Economists use comparative statics to predict changes in prices. This technique explains how changes in exogenous variables cause shifts in supply and/or demand curves, which lead to changes in prices.

check your understanding

1. Suppose coffee crops in Brazil are destroyed by inclement weather. What happens to the supply curve for coffee? What happens to the price of coffee and the equilibrium quantity of coffee?
2. The discussion of the Detroit Tigers states that "it's too soon to tell which Tigers will prove popular at the checkout line." Suppose that Miguel Cabrera has an excellent season and breaks the home run record. What do you expect will happen to the price and quantity of T-shirts with his name on the back?
3. In our discussion of the demand for beef and mad cow disease, we said that an increase in quantity and a decrease in price require a rightward shift of the supply curve. Draw a diagram to illustrate this case.

This page titled 8.3: Using the Supply-and-Demand Framework is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

8.4: Another Perspective on Changing Prices 8.4: Another Perspective on Changing PricesEdit Update titlecancel

Learning Objectives

1. Are price changes good for the economy?
2. How is information conveyed among households and firms in an economy?

Think back to our story of increasing beer prices. In Figure 8.4.1, we saw that an increase in the marginal cost of beer production led to an increase in the price and a decrease in the quantity supplied. In that explanation, we focused on what was happening to supply. But as the supply curve shifted, we moved along the demand curve to a new equilibrium. What was happening to the quantity demanded as the quantity supplied decreased? The answer is that as firms started decreasing their supply, the price in the market began to increase. Consumers of beer, confronted by these higher prices, bought less beer. Perhaps they switched to wine or spirits instead. The higher prices induced the quantity demanded to decrease in line with the decline in supply.

Something remarkable is happening in this story, however. Bad weather has affected the hops harvest, making beer more expensive to produce, relative to other goods and services. Because it is more expensive to make beer, it makes sense—from the point of view of society as a whole—to shift resources away from the production of beer and toward the production of other goods. And it makes sense—from the point of view of society as a whole—for people to consume less of the expensive-to-produce beer and more of other goods and services. If we imagine an all-knowing, all-powerful *central planner*, whose job is to allocate resources in the economy, we would expect this person to respond to the decrease in the hops harvest by ordering the production and consumption of less beer.

But this is exactly what happens in an economy, simply through the mechanism of supply and demand. The automatic adjustment of prices, resulting from shifts in supply and demand, brings about desirable shifts in production and consumption. Nobody orders producers to produce less or consumers to consume less. These outcomes result from the working of supply and demand.

Similarly, think about our T-shirt example. Consumers decide that they would like to consume fewer Granderson T-shirts. This change in their preferences shows up in the market as a shift in the demand curve, which causes the price of T-shirts to decrease. This decrease in the price encourages producers in the economy to adjust their behavior to fit the changed tastes of households. Firms stop producing Granderson shirts. Again, this is not because anyone has instructed them to do so. The changed tastes of households generate the price signal that induces firms to produce less.

So far, we have answered the question of the chapter by saying that prices change because of shifts in supply and/or demand. This answer is correct. But we could give a different answer from another perspective: prices change in order to provide signals to firms and households about what to produce and what to consume. In a market economy, households and firms decide what to consumer by considering the prices they face. Prices change in response to changes in costs and tastes, and these changes lead firms and households to adjust their decisions in line with the new economic reality.

It is fair to ask whether we should trust prices to play this role. Economics provides a very direct answer to this question: when markets are competitive, the price system delivers an efficient allocation of resources. In the following subsections, we develop the idea that markets deliver efficient outcomes by looking at a single market.

Buyer Surplus

Consider the market for chocolate bars, as shown in Figure 8.4.5. At the market clearing price, suppliers and demanders of chocolate bars trade the equilibrium quantity of chocolate bars. Imagine first that each household purchases no more than a single chocolate bar at the equilibrium price. For example, if 200 chocolate bars are sold, then 200 separate households bought a chocolate bar. Not all these households are alike, however: some like chocolate bars more than others. Most of them would have, in fact, been willing to pay more than the equilibrium price for the chocolate bar. Their **valuation** of a chocolate bar is greater than the price.

Any household that would have been willing to pay more than the equilibrium price gets a good deal. For example, suppose the equilibrium price is $5, but a household would have been willing to pay $7. Then that household receives a **buyer surplus** of $2.See Chapter 6 for more discussion. This logic extends to the case where households consume more than one unit. The demand curve of a household indicates the maximum amount that a person would pay for each successive unit of a good. The demand curve

shows the household's **marginal valuation** of a good. The individual household's demand curve slopes downward because the household is willing to pay less and less for each successive unit—the marginal unit—as the total quantity consumed increases.

In general, we know that a household purchases chocolate bars up to the point where

$$marginal\ valuation\ =\ price.$$

The household receives no surplus on the very last bar that it purchases because the marginal valuation of that bar equals price. But it receives surplus on all the other bars because its marginal valuation exceeds price for those bars. Diminishing marginal valuation means that the household obtains surplus from all the chocolate bars except the very last one.

Table 8.4.1 "Calculating Buyer Surplus for an Individual Household" gives an example of a household facing a price of $5. The first column is the quantity, the second is the price, the third is the marginal valuation (the extra value from the last chocolate), the fourth column measures the marginal surplus, and the last column is the total surplus.

Table 8.4.1: Calculating Buyer Surplus for an Individual Household

Quantity (Bars)	Price	Marginal Valuation	Surplus for Marginal Unit	Total Surplus
1	5	10	5	5
2	5	8	3	8
3	5	5	0	8
4	5	3	–2	6

The household is willing to buy three chocolate bars because the marginal value of the third bar is exactly equal to the price of $5. (In fact, the household would be equally happy buying either two or three bars. It makes no substantive difference to the discussion, but it is easier if we suppose that the household buys the last bar even though it is *indifferent* about making this purchase.) The household would not buy four bars because the marginal valuation of the last unit is less than the price, which means the surplus from a fourth chocolate bar would be negative.

The household obtains surplus from the first and second bars that it purchases. The household would have been willing to pay $10 for the first bar but only had to pay $5. It gets $5 of surplus from this first bar. The household would have been willing to pay $8 for the second bar but only had to pay $5. It gets $3 of surplus for this second bar. It gets no surplus from the third bar. So the total buyer surplus for this household is $5 + $3 = $8. Notice that by following the rule "buy until marginal valuation equals price," the household maximizes its total surplus from the purchase of chocolate bars.

More generally, the buyer surplus for this household is measured by the area under its demand curve (Figure 8.4.1). For each unit, the vertical difference between the price actually paid for each unit and the price the household would have been willing to pay measures the surplus earned for that unit. If we add the surplus over all units, we get the area between the demand curve and the price.

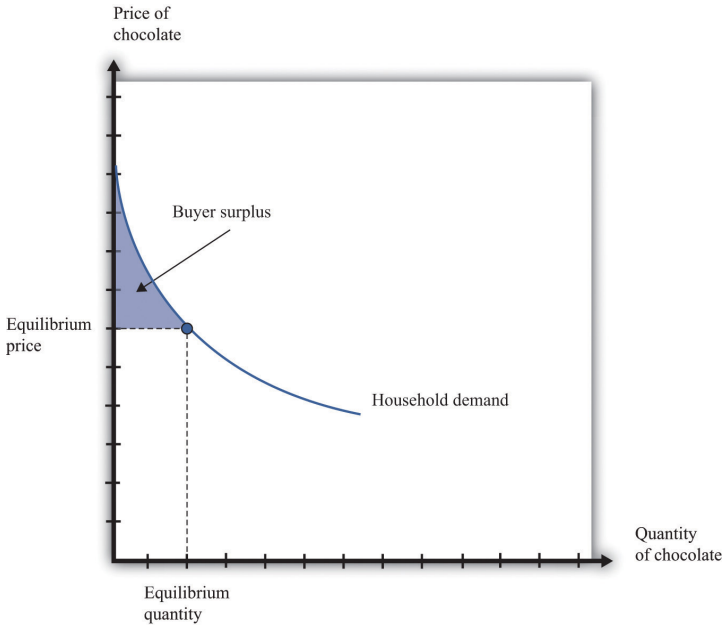


Figure 8.4.1: Buyer Surplus for an Individual Household

The buyer surplus is equal to the area between the demand curve and the price.

Seller Surplus

Sellers as well as buyers obtain surplus from trade. Suppose you won a used bicycle that you value at $20. If you can sell that bicycle for $30, you receive a **seller surplus** of $10—the difference between the price and your valuation of the good. It is worth your while to sell as long as the price is greater than your valuation. When a firm is producing a good for sale, the situation is analogous. If a firm can produce one more unit of a good at a marginal cost of $20, then the firm's valuation of the good is effectively equal to $20. If the firm can sell that unit for $30, it will receive a surplus of $10. The seller surplus earned by a firm for an individual unit is the difference between price and the marginal cost of producing that unit.

Given the price prevailing in a market, an individual firm in a competitive market will supply output such that the marginal cost of producing the last unit equals the price. The firm follows the rule: increase production up to the point where

$$price = marginal\ cost.$$

The example in Table 8.4.2 "Calculating Seller Surplus for an Individual Firm" gives the marginal cost of production for each unit and the surplus earned by a firm from producing that unit. If the firm produced only one unit, it would incur a marginal cost of $1, sell the unit for $5, and obtain a surplus of $4. The second unit costs $3 to produce, providing the firm with a surplus of $2. The third unit provides surplus of $1. The fourth unit costs $5 to produce, so the firm earns no surplus on this final unit. So the firm produces four units and obtains a total seller surplus of $7.

Table 8.4.2: Calculating Seller Surplus for an Individual Firm

Quantity	Price	Marginal Cost	Marginal Surplus	Total Surplus
1	5	1	4	4
2	5	3	2	6
3	5	4	1	7
4	5	5	0	7
5	5	6	–1	6

This difference between the price of a good and the marginal cost of producing the good is the basis of the seller surplus obtained by a firm. Exactly analogously to a household's buyer surplus, we measure the seller surplus by looking at the benefit a firm gets from selling each unit, and then we add them together. For each unit, the seller surplus is the difference between the price and the

supply curve (remember that the supply curve and the marginal cost curve are the same thing). When we add the surplus for all units, we obtain the area above the supply curve and below the price (Figure 8.4.2).

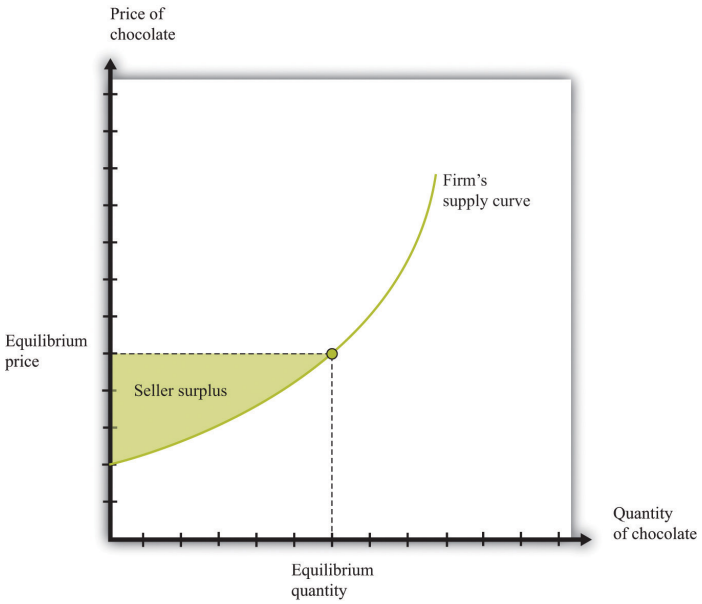


Figure 8.4.2: Seller Surplus for an Individual Firm. The seller surplus is the area between the equilibrium price and the firm's supply curve.

Toolkit: Section 31.1 and Section 31.10

You can review the concepts of valuation, marginal valuation, buyer surplus, and seller surplus in the toolkit.

Buyer Surplus and Seller Surplus for the Entire Market

So far we have considered the buyer surplus and seller surplus for an individual household and an individual firm. Because the market demand and supply curves are obtained by adding together the individual demand and supply curves, the same result holds if we look at the entire market. We illustrate this in Figure 8.4.3, which shows the total surplus flowing to all households and firms in the market equilibrium. The area below the market demand curve and above the price level is the total buyer surplus. The area above the market supply curve and below the price is the total seller's (producer's) surplus.There is one slightly technical footnote we should add. In some circumstances, the seller surplus may not all go to the firm. Instead, it may be shared between the firm and its workers (or other suppliers of inputs to the firm). Specifically, this occurs when increases in the market supply are large enough to cause input prices to change.

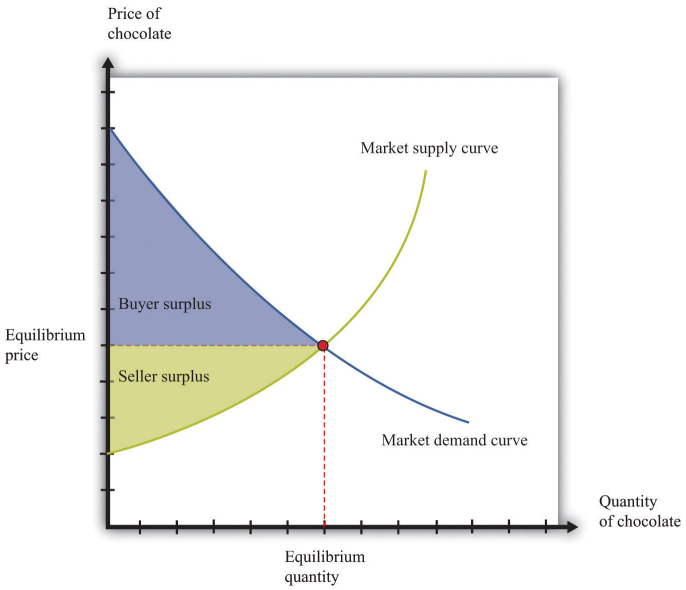


Figure 8.4.3: Surplus in the Market Equilibrium

The total surplus generated in a market is the sum of the buyer surplus and the seller surplus. It is therefore equal to the area below the demand curve and above the supply curve.

Markets and the Gains from Trade

The buyer surplus and the seller surplus tell us something remarkable about market outcome. If we add together the surplus for all buyers and sellers, we obtain the total surplus (gains from trade) in the market. In a competitive market, this is the maximum amount of surplus that it is *possible* to obtain—that is, exchange in a competitive market exhausts all the gains from trade.

There are two ways of seeing why this is true. First, we can ask what level of output would give us the largest total surplus. You might be able to see by looking at Figure 8.4.3, where the equilibrium quantity yields the largest total surplus. Figure 8.4.4 explains why in more detail. If there are fewer trades, then some surplus goes unrealized: some transactions that would yield positive surplus do not take place. To put it another way, there are buyers whose marginal valuation exceeds the marginal cost of production but who are unable to purchase the good. By contrast, if there are more trades than the equilibrium quantity, then some trades generate a negative surplus. The marginal cost of producing output beyond the competitive level is *less* than the goods are worth to consumers.

Second, the following things are true at the market equilibrium:

- For each household, the marginal valuation for the last unit equals the price.
- For each firm, price equals the marginal cost for the last unit.

Combining these two pieces of information, we know that each household's marginal valuation of the last unit is the marginal cost of producing that unit. As quantity increases, marginal valuation decreases and marginal cost increases. Therefore, if more of the good were produced, the marginal cost of the extra units would be higher than the marginal valuation. By the same argument, if fewer units were produced, the reduction in the household's valuation would be higher than the reduction in cost. So producing one unit more or one unit less would not be beneficial to households and producers. Remember that it is the adjustment of prices that ensures that an economy trades at the point where supply and demand are equal. Price adjustment allows buyers and sellers to obtain all the gains from trade.

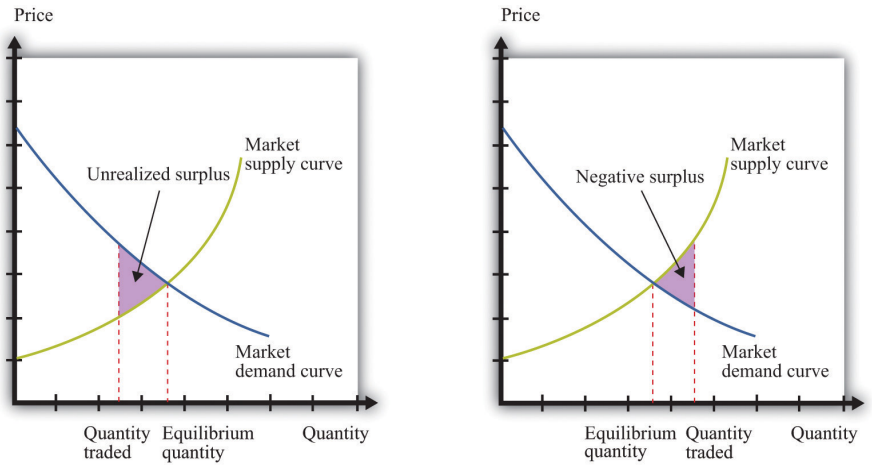


Figure 8.4.4: Surplus Away from the Market Equilibrium. If the quantity traded is less than the equilibrium quantity (a), then some gains from trade go unrealized. If the quantity traded is greater than the equilibrium quantity (b), then some trades generate negative surplus.

Why Markets?

We have been highlighting one of the principal messages in economics: markets are a mechanism to achieve the gains from trade. But are there other ways of achieving the same result? We previously introduced the fiction of an all-knowing, all-powerful central planner. Such a planner tells everyone what they should produce, takes all those goods, and distributes them throughout the economy. The planner tells everybody how much to work at each technology and decides exactly how to distribute all the goods and services that the economy produces.

Should we take this idea seriously or is it only a device to help us think about our theory? The answer is a bit of both. No economy has ever literally been run by a central planner. Historically, though, there have been many examples of so-called planned economies, where government bureaucracies played a major role in deciding what goods and services should be produced. For much of the 20th century, the economy of the Soviet Union operated under such a regime. China also used to be a largely planned economy. North Korea still operates as a largely planned economy.

Neither the Soviet Union nor China enjoyed much economic success under this system. The collapse of the Soviet Union's economy was a key reason why the country itself collapsed. China eventually changed its system of economic organization to one that gives more primacy to markets. Today, there are very few economies that operate under central planning and none that are significant in global economic terms. However, there are still several economies in which the government plays a significant role in the allocation of resources; so the analysis of the planner remains relevant.

Why were planned economies so unsuccessful? Books have been written on this topic, but there is one key insight. In order to make good decisions—decisions in the interest of individuals in the economy—the planner would need a lot of information. It is simply inconceivable that a planner could have sufficient knowledge about the abilities and skills of different individuals to make good decisions about where and how much they should work. Moreover, the planner also needs to know the tastes of everyone in the economy. Without that knowledge, the planner might instruct them to produce too many chocolate bars or not enough beer. If we think of an economy with millions of inhabitants, all with their own preferences and abilities, it is surely impossible that a planner could be sufficiently well informed to make decisions that are in the interest of all an economy's inhabitants.

key takeaways

- The response of prices and quantities to exogenous events is key for the efficient allocation of resources in the economy.
- Information about the tastes of households and the costs of production for a firm is conveyed through the price system.
- Through price adjustments in competitive markets, all potential gains from trade are realized.

check your understanding

1. If the demand for a good increases but the price is fixed, what gains from trade are lost?
2. Where in this section did we use the assumption of competitive markets?

This page titled 8.4: Another Perspective on Changing Prices is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

8.5: Three Important Markets 8.5: Three Important MarketsEdit Update titlecancel

Learning Objectives

1. How do the tools of comparative statics extend to other markets, such as the market for credit or labor?
2. How do markets interact with one another?

In other chapters in this book, we use the supply-and-demand framework to look at how goods and services are traded.We focus on labor in Chapter 9 and Chapter 11. Chapter 10 looks at both the loan/credit market and the foreign exchange market. Here we give a brief overview of these markets.

Credit Market

The **credit market (or loan market)**—we use the terms *loans* and *credit* interchangeably—is where suppliers and demanders of credit meet and trade (Figure 8.5.1). On the supply side are households and firms that, for various reasons, have chosen to save some of their current income. On the demand side are other households, firms, and (in some cases) the government. Households buy houses and cars, so they often need to borrow funds to finance those purchases. Firms seek credit to finance investment, such as the construction of a new production plant. Finally, governments borrow to finance some of their expenditures.

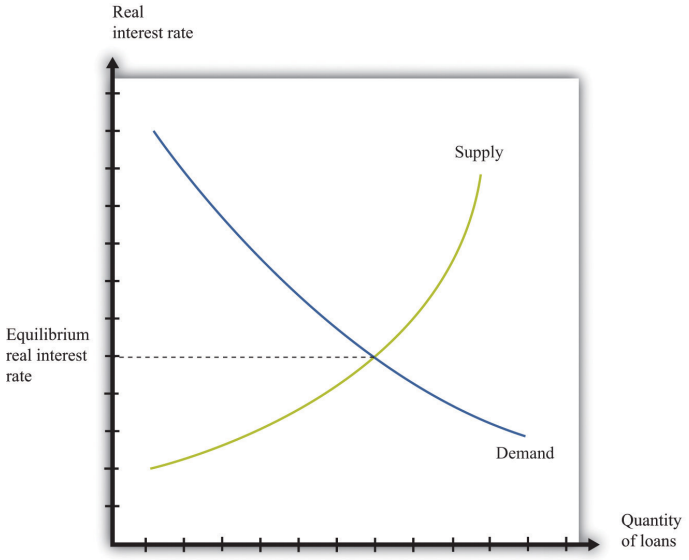


Figure 8.5.1: Credit Market Equilibrium. This diagram represents the loan or credit market.

The price of credit is the **real interest rate**, which is a measure of the value of the interest charged on a loan, adjusted for inflation. There are many different markets for credit because there are different kinds of loans in the economy. Chapter 10 discusses these. Associated with these different credit markets are different interest rates. For simplicity, though, we often suppose that there is a single market for credit.

Toolkit: Section 31.6

You can review the credit market and the real interest rate in the toolkit.

The demand for credit decreases as the real interest rate increases. When it becomes more expensive to borrow, households, firms, and even governments want fewer loans. The supply of credit by households increases with the real interest rate. When the return on savings increases, households and firms will typically save more and so supply more loans to the market.The response of savings to changes in the real interest rate is discussed more fully in Chapter 5.

The news is filled with stories about interest rates increasing and decreasing. You can always use some version of Figure 8.5.1 to understand why interest rates are changing. Ultimately, any change in the interest rate is due to a shift in either the supply of credit or the demand for credit. For example, if construction firms anticipate high future demand for housing, they will think that building new homes is a good use of investment funds. They will borrow to finance such construction. The increased demand for credit will

shift the demand curve in Figure 8.5.1 outward, and interest rates will increase. As another example, if individuals in other countries wish to increase their investment in US assets, this will shift the supply of credit outward, and interest rates will decrease.

Two of the most important players in the credit market are the government and the monetary authority. If the US federal government borrows more, this shifts the demand for credit outward and increases the interest rate. (The government is such a big player in this market that its actions affect the interest rate.) The monetary authority, meanwhile, buys and sells in credit markets to influence the real interest rate in the economy.The actions of the Federal Reserve and other monetary authorities are studied in detail in macroeconomics courses.

Foreign Exchange Market

If you travel abroad, you need to acquire the currency used in that region of the world. If you take a trip to Finland, Russia, and China, for example, you will undoubtedly buy euros, rubles, and yuan along the way. To do so, you need to participate in the **foreign exchange market**, trading one currency for another. Foreign exchange markets operate like other markets in the economy. The price—which in this case is called the **exchange rate**—is determined by the interaction of supply and demand.

Toolkit: Section 31.20

The foreign exchange market is the market where currencies are traded. The price in this market is the price of one currency in terms of another and is called the exchange rate.

Figure 8.5.2 shows an example of a foreign exchange market—the market in which US dollars are exchanged for euros. On the horizontal axis, we show the number of euros bought and sold on a particular day. On the vertical axis is the exchange rate: the price of euros in dollars. This market determines the dollar price of euros just like the gasoline market in the United States determines the dollar price of gasoline.

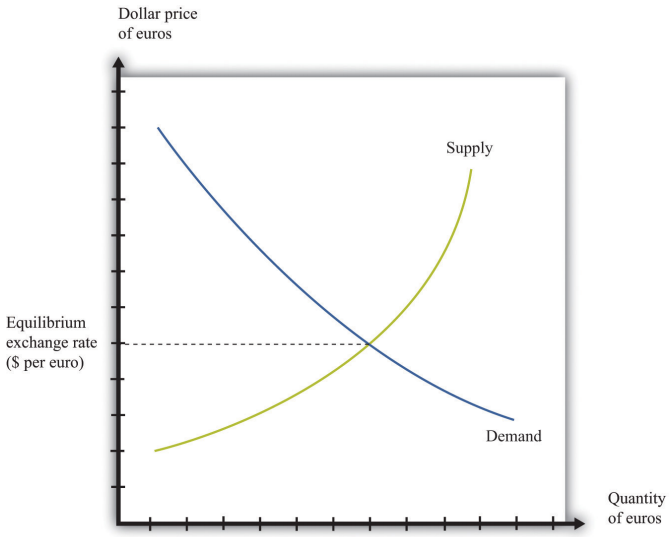


Figure 8.5.2: Foreign Exchange Market Equilibrium. This diagram represents the foreign exchange market in which euros are bought and sold with US dollars.

On the demand side, there are traders (households and firms) who want to buy European goods and services. To do so, they need to buy euros. This demand for euros expressed in dollars need not come from US households and firms. Anyone holding dollars is free to purchase euros in this market. On the supply side, there are also traders (households and firms) who are holding euros and who wish to buy US goods and services. They need to sell euros to obtain US dollars.

There is another source of the demand for and the supply of different currencies. Households and, more importantly, firms often hold assets denominated in different currencies. You could, if you wish, hold some of your wealth in Israeli government bonds, in shares of a South African firm, or in Argentine real estate. But to do so, you would need to buy Israeli shekels, South African rand, or Argentine pesos. Likewise, many foreign investors hold US assets, such as shares in Dell Computer or debt issued by the US government. Thus the demand and the supply for currencies are influenced by the portfolio choices of households and firms. In

practice, the vast majority of trades in foreign exchange markets are conducted by banks and other financial institutions that are adjusting their asset allocation.

In addition to households and firms, monetary authorities also participate in foreign exchange markets. For example, the US Federal Reserve Bank monitors the value of the dollar and may even intervene in the market, buying or selling dollars in an attempt to influence the exchange rate.

If you open a newspaper or browse the Internet, you can quickly find the current price of euros. This price changes all the time in response to changes in the currency's demand and supply. For example, if you read that the euro is getting stronger, this means that the euro is becoming more expensive: you must give up more dollars to buy a euro. This increase in the price of the euro could reflect either an outward shift in the demand for euros, say as US households demand more goods from Europe, or an inward shift of the euro supply curve if holders of euros are not as willing to sell them for dollars.

Labor Market

In the markets for goods and services, the supply side usually comes from firms. In some cases, buyers are other firms (businesspeople call these B2B transactions), whereas in other cases buyers are households (often called B2C transactions). For the most part, though, households are not on the supply side of these markets. In the **labor market**, by contrast, firms and households switch roles: firms demand labor and households supply it.

Supply and demand curves for labor are shown in Figure 8.5.3. Here the price of labor is the **real wage**. The real wage measures how much in the way of goods and services an individual can buy in exchange for an hour's work. It equals the nominal wage (the wage in dollars) divided by the general price level.

> **Toolkit: Section 31.3**
>
> You can review the labor market and the real wage in the toolkit.

The demand for labor comes from the fact that workers' time is an input into the production process. This demand curve obeys the law of demand: as the price of labor increases, the demand for labor decreases. The supply of labor comes from households that allocate their time between work and leisure activities. In Figure 8.5.3, the supply of labor is upward sloping. As real wage increases, households supply more labor. There are two reasons for this: (1) higher wages induce people to work longer hours, and (2) higher wages induce more people to enter the labor force and look for a job. Chapter 11 explains more about nominal wages and real wages, and we study the individual demand for labor in Chapter 9. The decisions underlying labor supply are explained more fully in Chapter 4.

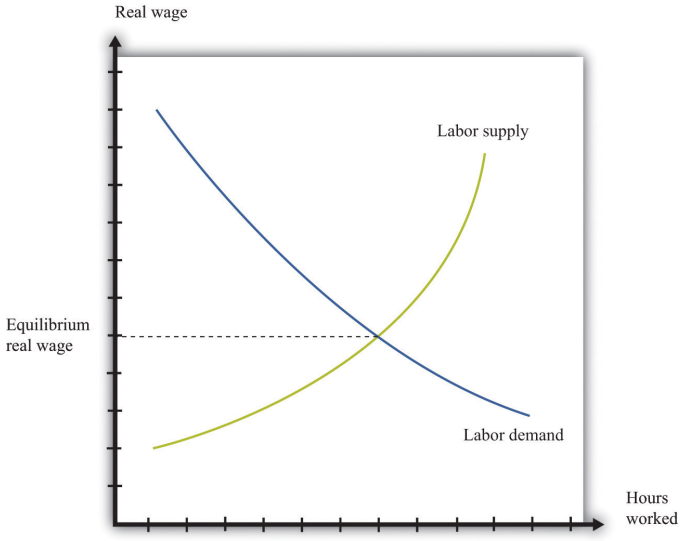


Figure 8.5.3: Labor Market Equilibrium. The equilibrium real wage is the price where supply equals demand in the labor market.

As with the other markets, we can use Figure 8.5.3 to study comparative statics. For example, if an economy enters a boom, firms see more demand for their products, so they want to buy more labor to produce more product. This shifts the labor demand curve outward, with the result that real wages increase and employment is higher.

Multiple Markets

You have now seen equilibrium in a wide variety of markets: goods (chocolate), loans, foreign exchange, and labor. Actual economies contain hundreds of thousands of markets. Analyzing a single market would be enough if the markets in an economy were not connected, but markets are interrelated in many ways:A topic in advanced studies of economics is the simultaneous equilibrium of all markets. Because all markets are linked, it is necessary to find prices for all goods and services and all inputs simultaneously such that supply equals demand in all markets. This is an abstract exercise and uses lots of mathematics. The bottom line is good news: we can usually expect an equilibrium for all markets.

- Factors that shift the demand in one market may affect other markets as well. For example, an increase in energy costs will raise the marginal costs of firms in all sectors of an economy.
- The demand for one good depends on the prices of others. In the market for coffee, for example, the demand curve depends on the price of goods that are complements to coffee, such as milk, and the prices of goods that are substitutes for coffee, such as tea.
- The supply curve for most goods depends on the prices of inputs, such as labor. The real wage—the price of labor—is determined by the supply and demand for labor. Thus the outcome of the labor market influences the position of the supply curve in almost every other market.
- The income level of households affects the position of the demand curve for most goods and services. But the level of income comes, in part, from the labor market outcome because labor income is part of the income households have to spend.

The following newspaper story from the Singaporean newspaper the *Straits Times* nicely illustrates linkages across markets.

Singaporeans with a sweet tooth could soon find themselves paying more for their favourite treats, as bakers and confectioners buckle under soaring sugar prices.

Since March last year, the price of white sugar has shot up by 70 per cent, according to the New York Board of Trade. As if that didn't make life difficult enough for bakers, butter and cheese prices have also risen, by 31 per cent and 17 per cent respectively.

The increases have been caused by various factors: a steep drop in Thailand's sugarcane production due to drought, higher sea freight charges, increasing demand from China's consumers for dairy products and the strong Australian and New Zealand dollar.See http://straitstimes.asia1.com.sg. We also discuss this quote in Chapter 6.

Look at the last paragraph. First, we are told that a drought has caused a drop in sugarcane production in Thailand. Part (a) of Figure 8.5.4 shows this market. We can see that a decrease in sugar production will increase the price of sugar. In this picture we are showing the market in Thailand, so the price is measured in Thai baht.

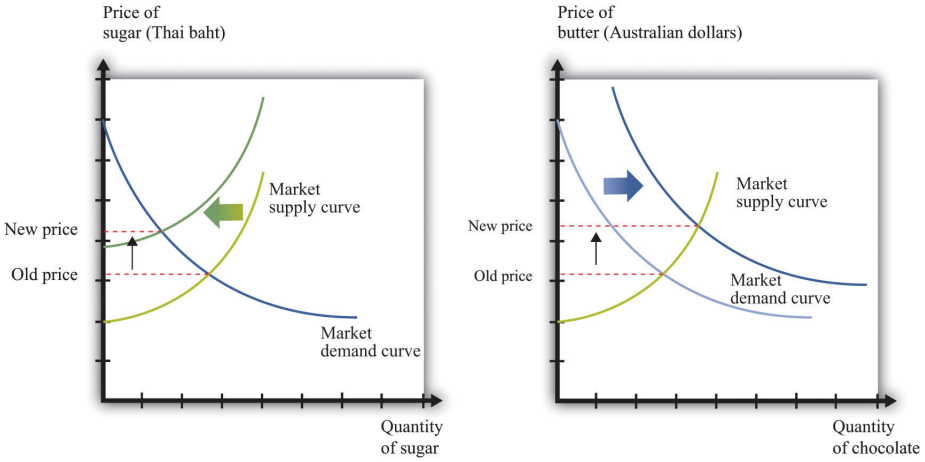


Figure 8.5.4: The Sugar Market in Thailand and the Butter Market in Australia. (a) The market for sugar in Thailand is affected by a drought, which has decreased the sugar supply, causing an increase in sugar prices measured in Thai baht. (b) In the Australian butter market, increased demand from China causes the demand curve to shift outward, increasing the price of butter measured in Australian dollars.

We are also told that there has been increased demand for dairy products coming from China. Australia and New Zealand are the major suppliers of dairy products in Southeast Asia. Part (b) of Figure 8.5.4 shows the market for butter in Australia. Increased demand from China shifts the demand curve outward, leading to an increase in the price of butter. For this market, we measure the price in Australian dollars.

From the perspective of Singapore bakers, both of these price changes show up as increases in their marginal cost. Moreover, the article reveals that these changes are exacerbated by other factors. Shipping costs have increased, so it also costs more to obtain sugar from Thailand and butter from Australia. And the Australian dollar has appreciated relative to the Singapore dollar, making goods imported from Australia even more expensive.

These are examples of B2B transactions. In fact, there is likely a whole chain of such transactions between, say, the Australian dairy farmer and the Singaporean baker. Farmers sell milk to butter producers, butter producers sell to wholesalers, and wholesalers sell to Singaporean importers and bakeries.

This story also illustrates again the powerful way in which market prices provide information that helps us understand the efficient allocation of resources. Drought in Thailand has reduced the amount of sugar available in the world. Through the magic of a series of prices, one of the results is that people in Singapore are less likely to eat cake for dessert.

key takeaways

- The supply-and-demand framework can be used to understand the markets for labor, credit, and foreign currency.
- Comparative statics can be used to study price and quantity changes in these markets.
- As markets interact with one another, sometimes comparative statics requires us to look at effects across markets.

exercises

1. Figure 8.5.2 shows the market where euros are bought and sold using dollars. We could equivalently think of this as the market where dollars are bought and sold using euros. Draw the graph for this market. How are the supply and demand curves in the two markets related to each other?
2. Using supply and demand, explain how an increase in Chinese demand for Australian butter might be a factor that causes the Australian dollar to appreciate.
3. What might be the effect of the financial crisis in the United States in 2008–9 on the income of lawyers? How does your answer depend on the specialization of the lawyer?

This page titled 8.5: Three Important Markets is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

8.6: Beyond Perfect Competition 8.6: Beyond Perfect CompetitionEdit Update titlecancel

Learning Objective

1. How can I predict what will happen to prices when markets are not competitive?

Everything that we have discussed in this chapter applies, strictly speaking, only to perfectly competitive markets. Yet the conditions for perfect competition are quite stringent. For a market to be perfectly competitive, there must be a large number of sellers of an identical product. There also must be a large number of buyers. Each buyer and seller must be "small" relative to the market, meaning that they cannot influence market price.

There are certainly some markets that fit these criteria. Markets for commodities, such as wheat or gold, are one example. Markets for certain financial assets are another. Such examples notwithstanding, the vast majority of markets are *not* perfectly competitive. In most markets, firms possess some market power, meaning that the demand curve they face is not perfectly elastic.

You might think this greatly weakens the usefulness of the supply-and-demand framework. A firm with market power chooses a point on the demand curve that it faces. It sets a price as a markup over marginal cost and then produces enough to meet demand at that price.We explain how firms set these prices in Chapter 7. A firm with market power does not take the price as given and then determine a quantity to supply. In fact—strictly speaking—there is no such thing as a supply curve when a firm has market power.

Economists understand this very well. Yet suppose you ask an economist to predict the likely effect of a worsening conflict in the Middle East on oil prices. The mental model she will use is almost certainly to imagine a supply curve for oil shifting to the left. Based on this model, she will predict higher prices and lower consumption. If you were to ask another economist to predict the effects of an economic recession on purchases of automobiles, he would imagine a demand curve shifting to the left and thus predict lower prices and lower output.

The first economist would use a supply-and-demand framework even though oil producers have market power. The second economist would use a supply-and-demand framework even though not all cars are identical. Although economists understand that many markets do not satisfy the strict conditions of perfect competition, they also know that the intuition from comparative statics carries over to more general market structures.

To see why, let us go back to our beer example again. We all know that not all beer is the same, and the beer companies spend a lot of money to convince us of this fact. Different beers have different tastes, and there are customers who are loyal to different beer brands. Breweries possess market power, meaning that we cannot—strictly speaking—draw a supply curve for individual beer producers or for the market as a whole.

Yet our comparative statics story, which supposes that the beer market is competitive, gives us an answer that makes sense. When the price of hops increases, this increases the marginal cost of production for all beer producers. Because they set prices based on a markup over marginal cost, the price of beer will increase, and less will be consumed. Output will be lower for all producers, and prices will be higher. Our comparative statics technique gives the right answer. Let us go through this more formally, first for a change in production costs and then for a change in demand.

Shifts in the Marginal Cost of Production

Figure 8.6.1 shows how a firm with market power sets its price.This figure is explained more fully in Chapter 7. To maximize its profits, a firm wants to produce the quantity where **marginal revenue** equals marginal cost. It sets the appropriate price as a **markup** over marginal cost.

Toolkit: Section 31.15

You can review the details of pricing with market power, including marginal revenue and markup, in the toolkit.

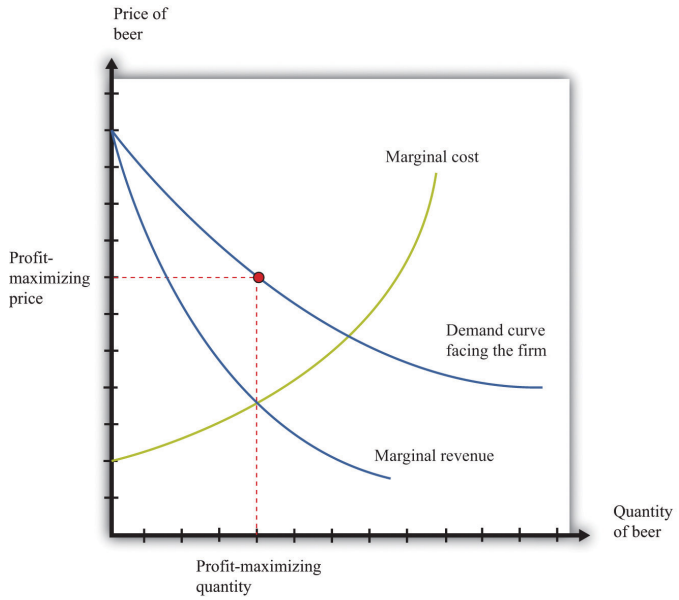


Figure 8.6.1: Finding the Profit-Maximizing Price and Quantity When a Firm Has Market Power. A firm with market power faces a downward-sloping demand curve and earns maximum profit at the point where marginal revenue equals marginal cost.

Now what happens if marginal cost increases? Think of a single beer producer and then imagine that the price of hops increases, so the marginal cost of producing an extra unit of output increases. This change in the marginal cost of production leads the brewer to decrease production (Figure 8.6.2). Marginal cost decreases and marginal revenue increases until the two are again equal.

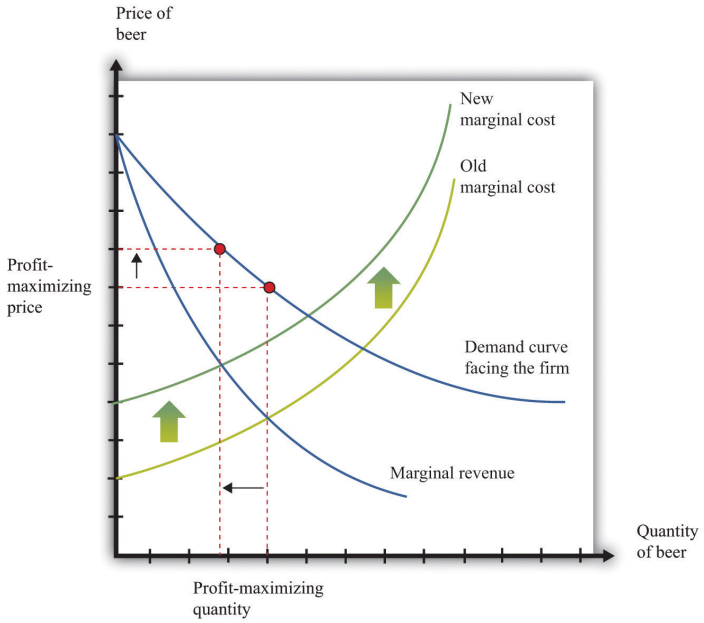


Figure 8.6.2: An Increase in Marginal Cost. In response to an increase in marginal cost, a firm now finds it optimal to set a higher price and produce a smaller quantity of beer.

Although Figure 8.6.2 is drawn for a single seller, it captures the common experience and response of *all* sellers. The increase in the price of hops causes the marginal cost to increase for all brewers, and they all respond by producing less and increasing the price of beer. And *this is exactly what we predicted about prices and quantities when we considered an increase in marginal cost in a perfectly competitive market.*

Shifts in Demand

Now reconsider the T-shirt example. There may be only a small number of producers who are licensed to produce Detroit Tigers T-shirts. Although there are many different kinds of replica sporting shirts in the world—at least one for each major team in most of the sports you can imagine—these shirts are not all the same. So the market for replica T-shirts—and certainly the market for

Granderson T-shirts—is not competitive. The producers of Granderson shirts for the Tigers choose quantities and set prices (see Figure 8.6.1).

Figure 8.6.3 shows the market as seen by one of these producers. Granderson's move from the Tigers shifts the demand curve inward for the shirts that they produce. This shift in the demand curve also shifts the marginal revenue curve inward. In response, the firm adjusts its output so that marginal revenue again equals marginal cost, choosing its price to match the point on the demand curve at the new quantity produced. Output and price both fall. Again, this is the same prediction that we obtained from the comparative statics of a competitive market.

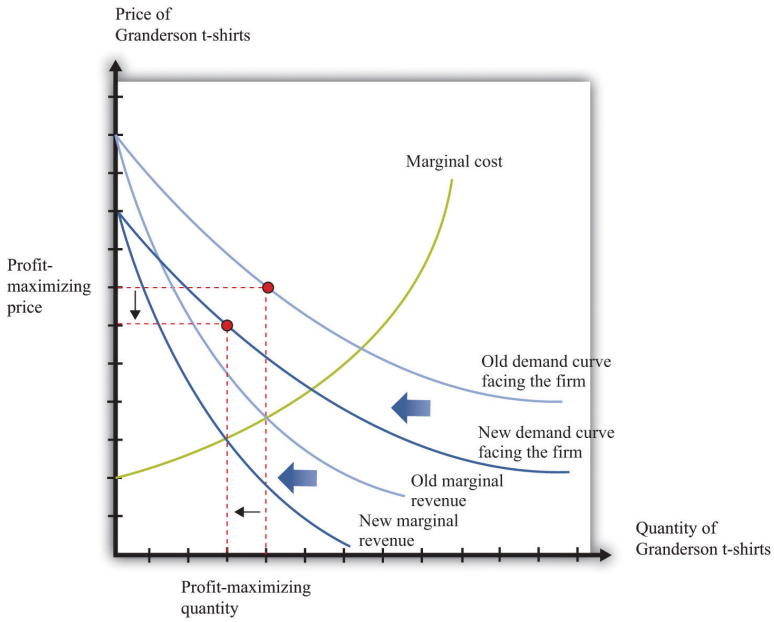


Figure 8.6.3: Shifts in Demand and Marginal Revenue. A decrease in demand causes marginal revenue to shift to the left. Marginal cost and marginal revenue intersect at a lower level of output. This lower level of output means marginal cost is lower, so the firm will also decrease its price.

When Is Using Supply and Demand Misleading?

Even when markets are not perfectly competitive, the supply-and-demand framework is usually a good device for predicting what will happen to prices and quantities in a market following a shock. Even if firms have market power, an increase in marginal cost will typically lead to an increase in the price and a decrease in the quantity supplied, just as supply and demand predict. Similarly, an increase in demand will typically lead to an increase in the price and an increase in the quantity supplied, again as predicted by basic supply-and-demand reasoning.

Although the supply-and-demand framework can be used for most situations where markets are not perfectly competitive, we still need to know when it might mislead us.

- Though unlikely, it is possible that an increase in demand will lead to an increase in the price and a *decrease* in the quantity supplied. This is because it is possible, though unlikely, that an increase in demand will cause marginal revenue to decrease.
- It is possible that an increase in demand will lead to an increase in quantity but a *decrease* in the price. There are two ways this could happen, both stemming from the fact that a firm sets its price as a markup over marginal cost. First, even if demand is greater, the elasticity of demand could change so as to make the optimal markup. Second, some firms with market power may have decreasing marginal cost.

We also need to recognize that while we may be able to use supply-and-demand intuition for qualitative predictions, it is more difficult to make *quantitative* predictions when markets are not competitive. We cannot determine supply and demand elasticities as easily when firms have market power. The reason is that one firm's decisions depend on what other firms are doing. Consider, for example, an increase in marginal cost in the beer industry. We have said that each firm in the market will respond by increasing its price and decreasing its quantity (Figure 8.6.2). For example, Miller in the United States will set higher prices in response to an increase in the price of hops. But when markets are not perfectly competitive, the story does not stop there. Firms also look at the prices set by their competitors. Miller's decisions on pricing depend also on the price chosen by Budweiser. If Budweiser sets a

higher price as well, then Miller may want to increase its price still further, and so on.How exactly this plays out is a complicated problem, requiring some of the ideas that we introduce in Chapter 15 and Chapter 17.

To sum up, the supply-and-demand framework can occasionally mislead when markets are not perfectly competitive. Yet most economists still begin with supply and demand when trying to explain a change in prices or quantities. Then they consider if there are reasons to expect either changes in the elasticity of demand or decreasing marginal cost. If neither of these seems likely, then the simple intuition of supply and demand will almost certainly give the right answer.

key takeaways

- Even if markets are not competitive, the qualitative predictions from comparative statics in a competitive market remain.
- The prediction of the supply-and-demand framework could be misleading if a shift of the demand curve does not lead the marginal revenue curve to shift in the same direction.

check your understanding

1. Suppose there is a decrease in marginal cost in some industry. What will happen to price and quantity if the industry is competitive? What will happen to price and quantity if firms set the price as a markup over marginal cost?

This page titled 8.6: Beyond Perfect Competition is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

8.7: End-of-Chapter Material 8.7: End-of-Chapter MaterialEdit Update titlecancel

In Conclusion

The supply-and-demand framework is the most powerful framework in the economist's toolkit. Armed with an understanding of this framework, you can make sense of much economic news, and you can make intelligent predictions about future changes in prices.

A true understanding of this framework is more than just an ability to shift curves around, however. It is an understanding of how markets and prices are one of the main ways in which the world is interlinked. Markets are, quite simply, at the heart of economic life. Markets are the means by which suppliers and demanders of goods and services can meet and exchange their wares. Because exchange creates value—it makes both buyers and sellers better off—markets are the means by which our economy can prosper. Markets are the means by which economic activity is coordinated in our economy, allowing us to specialize in what we do best and buy other goods and services.

Economists wax lyrical about these features of markets, but this should not blind us to the fact that markets can go wrong. There are many ways in which market outcomes may not be the most desirable or efficient. In other chapters, we look in considerable detail at all the ways that markets can fail us as well as help us.

exercises

1. Fill in the missing values in Table 8.7.1 "Individual and Market Demand". What can you say about the missing price in the table?

Price of Chocolate Bars ($)	Household 1 Demand	Household 2 Demand	Market Demand
1	7		22
2		11	16
10	0.5	3	3.5
	0.75	4	4.75

TABLE 8.7.1: INDIVIDUAL AND MARKET DEMAND

2. If the income levels of all households increase, what happens to the individual demand curves? What happens to market demand?
3. Suppose the price of coffee increases. Household 1 always eats chocolate bars while drinking coffee. What will happen to household 1's demand for chocolate bars when the price of coffee increases? Household 2 either has coffee or a chocolate bar for dessert. What happens to household 2's demand for chocolate bars when the price of coffee increases? What happens to the market demand for chocolate bars when the price of coffee increases?
4. In Figure 8.7.4, list the factors that would imply that firm 1 produces fewer chocolate bars than firm 2 when the price is $5. The figure is drawn so that firm 1 produces less than firm 2 at all prices. Does this have to be the case? Could the firms' supply curves cross?
5. (Advanced) Draw a version of Figure 8.7.3 if there is an outward shift in demand but no shift in marginal revenue. What would happen to the market price?
6. Consider the operation of a café. Describe the types of trades in terms of whether they are B2C or B2B. In what ways do you think that B2B trades are different from B2C trades?
7. Economists often say that individual decisions are "made at the margin." How do you see that in the determination of market supply and market demand?
8. If there are fewer sellers in a market, what will happen to total output? What will happen to the output of each seller?
9. Explain why an increase in the mortgage rate, which reduces the demand for new houses, can teach researchers about the elasticity of the supply curve.
10. (Advanced) Using the credit market, show how governmental borrowing increases interest rates. Could governmental borrowing also lead to an outward shift in the supply of credit as households save more to pay off the future debt? How would you show this in a supply-and-demand diagram?

11. If the US Federal Reserve Bank takes actions to lower interest rates in the United States relative to other countries, what will happen to the euro price of the dollar? Explain.
12. Draw a figure showing an outward shift in a demand curve along with a reduction in marginal revenue. Explain what is going on in the diagram and how a monopolist would respond to the situation.

Spreadsheet Exercise

1. Using a spreadsheet, construct a version of Table 8.2.2 "Market Equilibrium: An Example" assuming that market demand = 50 − 5 × price. Fill in all the prices from 1 to 100. What is the equilibrium price and the equilibrium quantity in the market? How would you explain the difference between this equilibrium and the one displayed in Table 8.2.2 "Market Equilibrium: An Example"?

Economics Detective

1. Find a newspaper article that describes a price change for a good or service. Why did the price change? What happened to the quantities produced and sold?

This page titled 8.7: End-of-Chapter Material is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

8.8: Appendix- Algebraic Presentation of Supply and Demand 8.8: Appendix- Algebraic Presentation of Supply and DemandEdit Update titlecancel

The supply-and-demand framework can be analyzed with algebra. We start with supply and demand and then talk about market equilibrium. This presentation uses some notation rather than only words:

- p is the price of a chocolate bar.
- I measures the income of individuals in a market.
- q^d is the quantity demanded.
- q^s is the quantity supplied.
- A measures the technology of chocolate bar production.

Market Demand

With this notation, we represent the demand curve as follows:

$$q^d = a_d - \beta_d p + \mathrm{V}_d I$$

α_d, β_d, and γ_d are constants that characterize the effects of prices and income on the quantity demanded. With the restriction that $\beta_d > 0$, the demand curve is downward sloping because an increase in p implies a reduction in the quantity demanded. It is natural to assume $\gamma_d > 0$, so an increase in income leads to an increase in the quantity demanded. This is represented as a shift in the demand curve.

Market Supply

With this notation, we represent the supply curve as follows:

$$q^5 = a_s - \beta_s p + \mathrm{Y}_s \mathcal{A}$$

α_s, β_s, and γ_s are constants that characterize the effects of prices and income on the quantity demanded. With the restriction that $\beta_s > 0$, the supply curve is upward sloping because an increase in p leads to an increase in the quantity supplied by all firms. It is natural to assume $\gamma_s > 0$, so an increase in the productivity of the current technology leads to an increase in the quantity produced at a given price. This is represented as a shift in the supply curve.

Market Equilibrium

The market is in equilibrium if there is a price and quantity combination, denoted (p*, q*) such that at the price p*, the quantity demanded, and the quantity supplied equal q*. Equilibrium is the simultaneous solution of supply and demand and can be found using the substitution method outlined in the toolkit.

Using $q^s = q^d = q^*$, we can substitute Equation 8.8.2 into Equation 8.8.1 yielding:

$$a_d - \beta_d p^* + \mathrm{Y}_d I = a_s - \beta_s p^* + \mathrm{Y}_s \mathcal{A}$$

This is a single equation in a single unknown, p*. Solving the equation for p* implies

Equation 8.4

The denominator is positive because we have assumed that both β_d and β_s are positive. The numerator is positive as long as the vertical intercept of the demand curve is greater than the vertical intercept of the supply curve: $(a_d + \mathrm{y}_s I) > (a_s + \mathrm{Y}_d A)$ This condition, combined with the restrictions on the slopes of supply and demand, is sufficient to guarantee that an equilibrium price exists in the market.

Using this calculation of p* in, say, the supply curve, we find

Equation 8.5

Grouping the terms into a constant, $\gamma_d I$ and $\gamma_s A$, this becomes

Equation 8.6

Looking at Equation 8.4 and Equation 8.6, these expressions determine the equilibrium price and the equilibrium quantity depending on the two (exogenous) factors that impact supply and demand: income level I and state of technology A. Though

income influences only the position of the demand curve, variations in income influence both the equilibrium price and the equilibrium quantity. The same is true for variations in technology that shift only the supply curve.

This page titled 8.8: Appendix- Algebraic Presentation of Supply and Demand is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

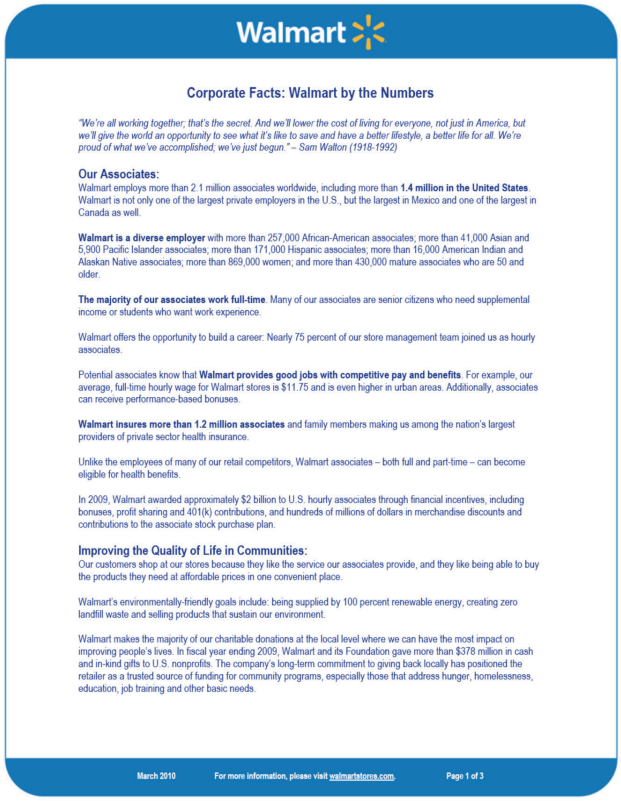
CHAPTER OVERVIEW

9: Growing Jobs

9.1: Changing Jobs
9.2: How Do Firms Decide How Many Hours of Labor to Hire?
9.3: Entry and Exit
9.4: Search
9.5: Government Policies
9.6: End-of-Chapter Material

This page titled 9: Growing Jobs is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

9.1: Changing Jobs 9.1: Changing JobsEdit Update titlecancel



Walmart

Corporate Facts: Walmart by the Numbers

"We're all working together; that's the secret. And we'll lower the cost of living for everyone, not just in America, but we'll give the world an opportunity to see what it's like to save and have a better lifestyle, a better life for all. We're proud of what we've accomplished; we've just begun." – Sam Walton (1918-1992)

Our Associates:
Walmart employs more than 2.1 million associates worldwide, including more than **1.4 million in the United States**. Walmart is not only one of the largest private employers in the U.S., but the largest in Mexico and one of the largest in Canada as well.

Walmart is a diverse employer with more than 257,000 African-American associates; more than 41,000 Asian and 5,900 Pacific Islander associates; more than 171,000 Hispanic associates; more than 16,000 American Indian and Alaskan Native associates; more than 869,000 women; and more than 430,000 mature associates who are 50 and older.

The majority of our associates work full-time. Many of our associates are senior citizens who need supplemental income or students who want work experience.

Walmart offers the opportunity to build a career: Nearly 75 percent of our store management team joined us as hourly associates.

Potential associates know that **Walmart provides good jobs with competitive pay and benefits**. For example, our average, full-time hourly wage for Walmart stores is $11.75 and is even higher in urban areas. Additionally, associates can receive performance-based bonuses.

Walmart insures more than 1.2 million associates and family members making us among the nation's largest providers of private sector health insurance.

Unlike the employees of many of our retail competitors, Walmart associates – both full and part-time – can become eligible for health benefits.

In 2009, Walmart awarded approximately $2 billion to U.S. hourly associates through financial incentives, including bonuses, profit sharing and 401(k) contributions, and hundreds of millions of dollars in merchandise discounts and contributions to the associate stock purchase plan.

Improving the Quality of Life in Communities:
Our customers shop at our stores because they like the service our associates provide, and they like being able to buy the products they need at affordable prices in one convenient place.

Walmart's environmentally-friendly goals include: being supplied by 100 percent renewable energy, creating zero landfill waste and selling products that sustain our environment.

Walmart makes the majority of our charitable donations at the local level where we can have the most impact on improving people's lives. In fiscal year ending 2009, Walmart and its Foundation gave more than $378 million in cash and in-kind gifts to U.S. nonprofits. The company's long-term commitment to giving back locally has positioned the retailer as a trusted source of funding for community programs, especially those that address hunger, homelessness, education, job training and other basic needs.

March 2010 For more information, please visit walmartstores.com. Page 1 of 3

Figure 9.1.1: Walmart Fact Sheet. Source: *http://walmartstores.com/pressroom/factsheets*.

Figure 9.1.1 is a fact sheet from Walmart. The fact sheet tells you—if you didn't know already—that Walmart is everywhere. It has over 7,000 retail units in the world, with over 3,000 outside the United States. Walmart employs about 1.36 million people in the United States alone, which is about 1 percent of the total number of workers employed in the United States. It seems that Walmart means jobs.



Figure 9.1.2: Walmart: Growing Jobs?. Source: This image is taken from en.Wikipedia.org/wiki/Image:Walmartbizarro.png. The image is copyright Dan Piraro 2006.

Despite the fact that Walmart provides so many jobs, the announcement of a new Walmart store is often greeted with trepidation or outright opposition. There are websites and even a film (http://www.walmartmovie.com) dedicated to criticism of Walmart. It is true that the arrival of Walmart in a town will mean the creation of new jobs, including checkout clerks, shelf packers, and many other positions. Yet the arrival of Walmart will also mean that its competitors will lose jobs. The overall effect on jobs is unclear.

The arrival of a Walmart in a town has implications beyond the effects on jobs. Consumers are likely to face a different menu of goods. Walmart will bring some goods that were previously unavailable, while at the same time other goods sold by unsuccessful specialty stores will disappear. The shopping experience will change for consumers because shopping in a Walmart is not like shopping in a series of small stores. Consumers will face different and—for the most part—lower prices: Walmart is able to obtain goods more cheaply from suppliers and also has a very efficient distribution system that decreases costs. As fewer consumers visit the smaller shops, other nearby businesses, such as local restaurants, may suffer a decline in demand. The patterns of life in the town will change in numerous yet subtle ways.

This scenario has played out in many countries with many different stores. In France, there are hypermarkets such as Carrefour that likewise have had major effects on local businesses and sometimes encounter opposition. In England, the same is true of the Asda supermarket chain. In this chapter, we look at the economics that lies behind a firm's decision to enter new markets (that is, open new stores) and exit from markets (that is, close down and go out of business). Although we begin here with Walmart, we tell a story that is about much more. Over the course of every year, jobs are created when existing firms expand and new firms enter. At the same time, jobs are lost when firms contract their workforces or close down completely.

A job is created when either an existing firm or a new firm hires workers. Jobs are destroyed when a firm fires some of its workers, some workers quit, or a firm exits a market. Figure 9.1.3 provides data on net job creation in the United States from 2000 to 2009. By net job creation, we mean the number of new jobs created minus the number of jobs destroyed. For example, if some firms expand their employment by 200 workers and other firms reduce their workforce by 150 workers, we say that 200 jobs are created, 150 are destroyed, and the net job creation is 50.

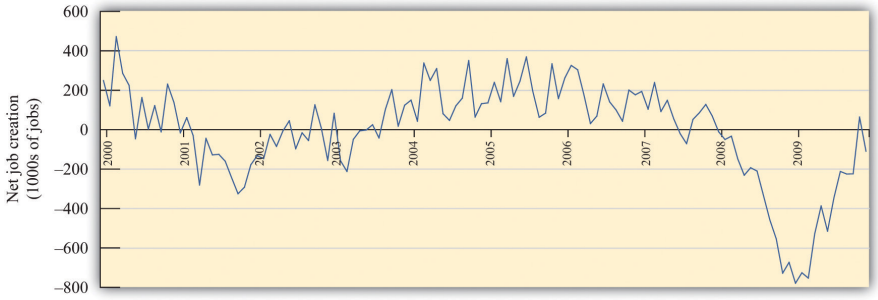


Figure 9.1.3: US Net Job Creation from 2000 to 2009

Job creation and destruction take place all the time. Within an industry, and sometimes even within a firm, we see job creation and destruction occurring simultaneously. The Bureau of Labor Statistics now regularly produces quarterly job creation and destruction rates for the US economy.See Bureau of Labor Statistics, Business Employment Dynamics, “Table 3. Private sector gross job gains and losses, as a percent of employment, seasonally adjusted,” accessed March 1, 2011, http://www.bls.gov/web/cewbd/table3_1.txt. For the US private sector over the period 1990–2010, the average quarterly job creation and destruction rates were 7.5 percent and 7.3 percent, respectively. To put it another way, if you looked at 1,000 typical private sector jobs right now, then 75 of them didn’t exist last quarter, and 73 won’t exist next quarter. This implies that about 15 percent of jobs are either destroyed or created in a given quarter. Steven Davis, John Haltiwanger, and Scott Schuh, who were perhaps the first economists to study these processes in detail, call this *job reallocation* because it reflects the reshuffling of jobs across production locations. Sometimes a car produced at one particular automobile factory is selling well, so new jobs are created there. In the same quarter, another automobile factory may be shut down because the models produced there are not selling. The picture you should take away from these numbers is one of a very fluid labor market. One of our goals in this chapter is to understand how this labor market works.

One way that jobs can be created is by expanding an existing plant or firm. Another way that jobs are created is by the entry of a new plant or firm. The situation is analogous for job destruction. Some jobs are lost when existing plants contract (such as a plant eliminating a shift). Others are lost through the exit of a plant or a firm. The Davis, Haltiwanger, and Schuh data suggest that, each quarter, 11.5 percent of the jobs destroyed come from plants closing. An additional 20 percent of jobs are destroyed when plants undertake large workforce adjustments of more than 50 percent. On the entry side, about 8.4 percent of jobs created in a quarter are due to start-ups. Of course, if these start-ups are successful, then they create more jobs in later years. Ultimately, we want to answer the following question:

What are the economic forces driving job creation and destruction?

Road Map

The first part of this chapter examines labor demand by firms. We begin by looking at the decision of how many hours to hire. Then we turn to the entry and exit decisions of firms. We find the decision rules that govern these decisions.

After that, we consider a different way of looking at the labor market by examining the process of search and bargaining. We look at how workers who supply labor and firms that demand labor actually interact. Finally, we examine the effects of various government policies on labor market outcomes.

This page titled 9.1: Changing Jobs is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

9.2: How Do Firms Decide How Many Hours of Labor to Hire? 9.2: How Do Firms Decide How Many Hours of Labor to Hire?Edit Update titlecancel

Learning Objectives

1. What is a production function?
2. How does a firm decide how much to produce?
3. What factors determine a firm's labor demand?

When economists are asked to explain the creation and destruction of jobs in an economy, they will typically begin with a diagram of supply and demand in the **labor market**. In the labor market, the **real wage** (on the vertical axis) and the total number of hours worked (on the horizontal axis) are determined by the interaction of labor supply and labor demand. As shown in Figure 9.2.1, equilibrium in the labor market occurs at the wage and employment level such that the number of hours supplied and demanded is equal.

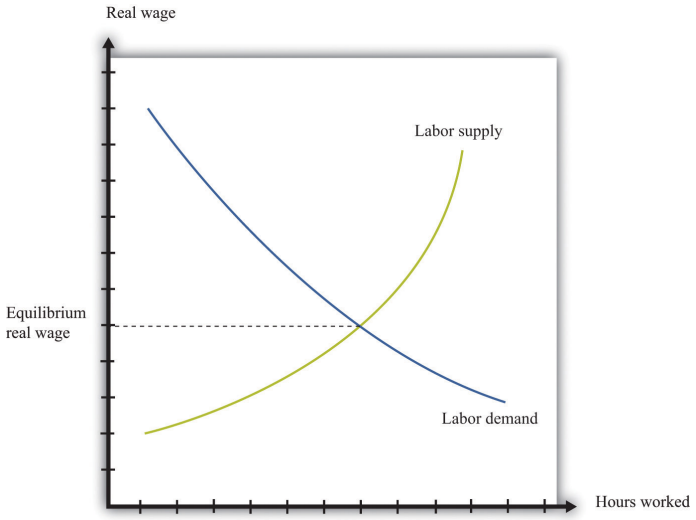


Figure 9.2.1: Labor Market Equilibrium

The equilibrium real wage in the labor market is the price where supply equals demand.

Toolkit: Section 31.3

See the toolkit for more discussion on the labor market.

The upward-sloping supply curve tells us that households will want to supply more labor time as wages increase.Labor supply is discussed in Chapter 4.

- As the wage increases, some people find it worthwhile to enter the labor force and look for a job.
- As the wage increases, some of those with jobs will find it worthwhile to work more hours.

Labor demand slopes downward for two analogous reasons:

- As the wage increases, some firms find that they are no longer profitable and close down.
- As the wage increases, those firms that are in business choose to hire fewer hours of labor.

Thus an increase in wages will induce job destruction, and a decrease in wages will induce job creation.

The Decisions of a Firm

Firms hire labor to help them produce output. The amount of labor that a firm needs depends on the amount of output that it wants to produce. At the same time, its decision about how much to produce depends on its costs of production, which include the cost of labor. Our task here is to combine these ideas. The decision about how much labor to hire is only one of a large number of choices made by a firm's managers. Of these, the most fundamental decisions are the following:

- Should a firm be in business at all?
- How much should a firm produce, and what price should its managers set?
- How should a firm produce its desired level of output?

A firm's managers should actually answer these questions in the reverse order:

1. Managers should first determine the best way to produce output.
2. Then managers need to make a price/output decision. A firm is fundamentally constrained by the desires of the market. If managers choose the price of output, they must accept whatever sales are demanded by consumers at that price. If they choose the level of output, they can only charge the price that the market will bear for that quantity. In other words, a firm's managers must choose a point on the **demand curve facing a firm**.
3. Then, and only then, can a firm's managers decide if it is worth being in business at all. An existing firm can stay in the market or exit, thus destroying jobs. Potential new firms can enter the market and create jobs.

We follow this logic in our discussion.

> **Toolkit: Section 31.15**
>
> You can review the demand curve facing a firm and the details of pricing with market power in the toolkit.

The Production Function

A firm possesses a means of turning inputs into outputs. For example, Starbucks produces, among other things, grande vanilla low-fat decaf lattes. This drink is an example of a Starbucks output. The list of inputs that Starbucks needs to produce this product is much too long to write out in full but includes the following:

- Water, coffee beans, milk, vanilla syrup, paper cups, lids, and electricity
- An espresso machine
- The time of the barista—the person making the drink
- Starbucks' "blueprints" (that is, the instructions for how to make a grande vanilla low-fat decaf latte)

This list doesn't include any of the "back office" aspects of Starbucks' operations, such as accounting, payroll, or the logistics of sourcing coffee beans and delivering them to individual stores. Other firms, of course, would have a very different list of inputs. So if we want to talk in abstract terms about the production of a firm, we need a description of production that could apply not only to Starbucks but also to General Motors, IKEA, your local computer repair store, and a manufacturer of paper clips. Therefore, we group inputs into broad categories called **factors of production**.

> **Toolkit: Section 31.17**
>
> Economists group the inputs of any firm into a small number of general categories: raw materials, capital, and labor. We call these inputs a firm's factors of production.

You can think of raw materials as the things that are transformed in the production process. In our Starbucks example, these include milk, coffee beans, and electricity. Labor refers to the time of the employees who work at a firm, so the time put in by a Starbucks' barista counts as labor. Capital refers to goods that are used to help with production but are not used up in the process. The espresso machine is one of Starbucks' capital goods; others are the tables and chairs in the café.

Starbucks' **technology**—which we also think of as a factor of production—is the knowledge that allows it to take all these inputs and turn them into an output—the final product that people actually want to buy. It is this knowledge that ultimately lies behind Starbucks' existence as a firm. Included in the technology are the managerial skills that allow Starbucks to operate effectively.

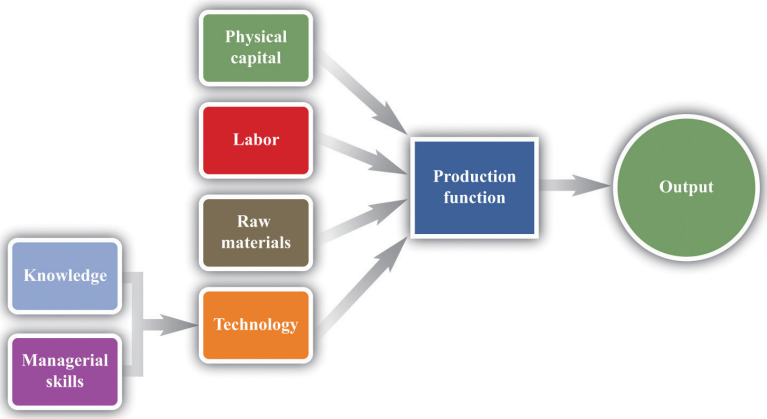


Figure 9.2.2: The Technology of a Firm

The production function combines a firm's physical capital stock, labor, raw materials, and technology to produce output. Technology is the knowledge (the blueprints) that a firm possesses, together with managerial skills.

We represent the production process of a firm schematically in Figure 9.2.2. Our description is quite general and can apply to nearly any kind of firm—for example, a lawyer's office, Walmart, a university, and a child's lemonade stand. Most people find it easiest to visualize a production function in terms of physical manufacturing, such as a production line for automobiles. Think of a firm's capital as factory buildings and machinery; its labor as the workers on the production line; and its raw materials as the steel, plastic, and glass that it purchases.

A Production Function That Uses Only Labor

We summarize the technological possibilities of a firm using a **production function**, which is a description of how much output a firm can produce as it varies its inputs. Even though a typical firm's production function contains many different inputs, we can understand most of the key features of the production function using an example where labor is the only factor of production. Although there are few goods or services that literally require no inputs other than labor, there are many services that are highly labor intensive, such as babysitting, housecleaning, and personal training at a gym.

To be concrete, think about housecleaning and suppose it has the following production function:

$$output \ = \ productivity \ \times \ hours \ of \ labor \ input,$$

where we think of productivity as just some number. If output measures clean houses, and if it takes 5 hours of labor to produce one clean house, then productivity is 0.2, and the production function is

$$output \ = \ 0.2 \times \ hours \ of \ labor \ input.$$

The production function tells us the level of output of a firm for given levels of labor input. Labor input is the total hours of labor time used by a firm. At this point, we are not distinguishing between hours worked per person and the number of people working, so a firm with 8 employees each working 20 hours per week has the same weekly labor input as a firm with 4 employees each working 40 hours per week. Table 9.2.1 "Production Function for Housecleaning" lists the amount of output that a housecleaning firm can obtain from various levels of input. We call this a *linear* production function because its graph is a straight line, as shown in Figure 9.2.3.

Table 9.2.1: Production Function for Housecleaning

Labor Input (Hours)	Output (Clean Houses)
0	0
1	0.2
2	0.4
3	0.6
4	0.8

Labor Input (Hours)	Output (Clean Houses)
5	1
10	2
15	3
20	4
25	5
30	6
35	7
40	8

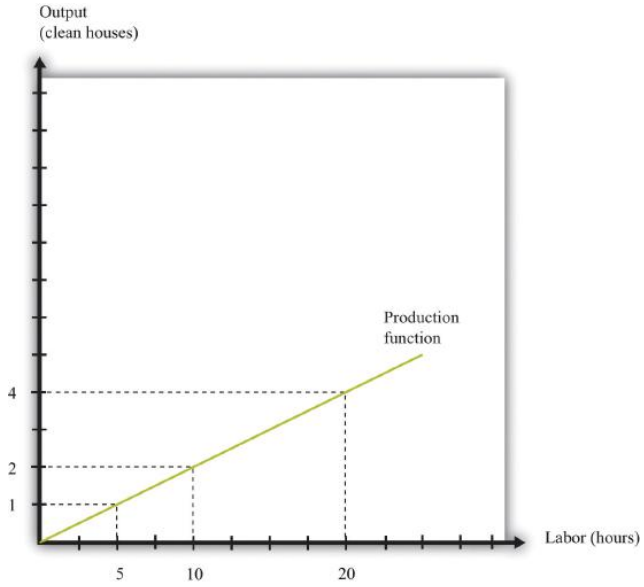


Figure 9.2.3: A Linear Production Function. A production function shows the maximum amount of output produced, given a level of labor input.

The **marginal product of labor** is the amount of extra output produced from one extra hour of labor input and is defined as

When the production function is linear, the marginal product of labor is constant. It is equal to the number we labeled *productivity* in our original production function.

In most cases, the marginal product of labor is not constant. To understand why, imagine you are managing a Starbucks outlet. You already have the machines to produce espresso, and you have lots of coffee beans on hand. You also have 500 square feet of space for making coffee and charging customers. But you still need labor. If you have no barista to operate the espresso machine, then you will have no output. If you hire one worker, you will be able to serve coffee to people. Adding the first worker will increase output considerably. However, that person must not only make the coffee but also clear the tables and handle the cash register. Adding a second worker to help with the register and clear tables will increase output even more. Now suppose you keep increasing the number of workers in the 500 square feet of space. After the third or fourth worker, they will start to bump into each other, and the barista will start to be very annoyed and unproductive. In other words, because one of your inputs—the amount of available space—is fixed, each additional worker contributes less and less to output. We call this the **diminishing marginal product of labor**.

Table 9.2.2 "Production Function for Coffee with a Diminishing Marginal Product of Labor" is an example of a production function with a diminishing marginal product of labor. In creating this table, the labor input is changed while holding all other inputs (the size of the café, the number of espresso machines, etc.) fixed.

Table 9.2.2: Production Function for Coffee with a Diminishing Marginal Product of Labor

Labor Input (Hours)	Output (Cups of Coffee)	Marginal Product of Labor

Labor Input (Hours)	Output (Cups of Coffee)	Marginal Product of Labor
0	0	
1	10	10
2	14.1	4.1
3	17.3	3.2
4	20	2.7
5	22.4	2.4
…		
10	31.6	1.6
15	38.7	1.3
20	44.7	1.1
25	50.0	1.0
30	54.8	0.9
35	59.2	0.9
40	63.2	0.8

The marginal product of labor is shown in the third column. For the first few entries, you can calculate it directly from the table because you can easily determine how much output changes from one row to the next. For example, the marginal product of the third hour of labor is 17.3 – 14.1 = 3.2. Finding the marginal product of, say, the 40th unit of labor from the table is trickier because the table doesn't tell us how much we can produce with 39 hours of labor. Looking back at the formula for the marginal product of labor, however, we can calculate it:

We illustrate this production function in Figure 9.2.4. Notice that while the slope of the production function is always positive, the slope decreases as the labor input increases.

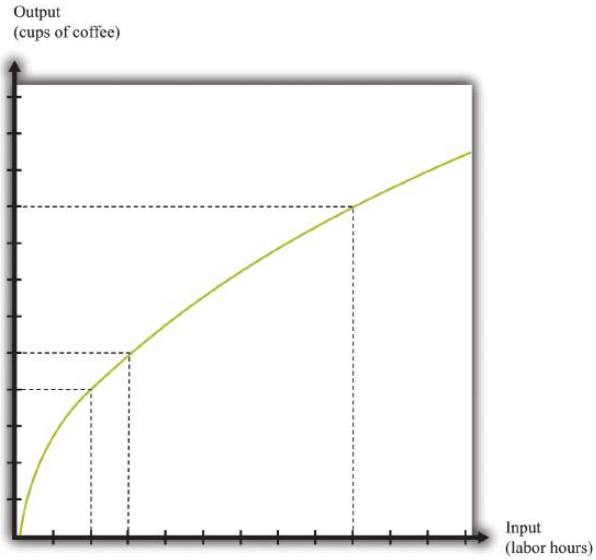


Figure 9.2.4: A Production Function with a Diminishing Marginal Product of Labor

This production function exhibits diminishing marginal product of labor: as more labor is added to a firm, output increases at a decreasing rate.

Toolkit: Section 31.17

The production function is a description of how much output a firm can produce as it varies its inputs. Typically, we suppose that the production function exhibits the following:

- Positive marginal product of labor
- Diminishing marginal product of labor

The first property means that adding more labor into production means more output—that is, the slope of the production function is positive. The second property explains how the marginal product of labor varies as labor input increases. Though the marginal product of labor is always positive, it will generally decrease as more labor is added to a production process. That is why the second property is called *diminishing* marginal product of labor. (It is technically possible that the marginal product of labor could even become negative. But because a firm would never pay for workers when they decrease output, we never expect to see a firm operating with a negative marginal product of labor.)

The Cost Function

Now that we have a way of describing a firm's ability to produce goods, we are well on our way to understanding how a firm produces output. This then allows us to understand how much it will cost a firm to produce different levels of output. Our next goal is to describe these costs. The total cost of producing some specified level of output represents the cost of acquiring all the inputs needed.

To see how this works, let us determine the costs for our earlier housecleaning example. Recall that the production function is

$$output \ = \ 0.2 \times \ number \ of \ hours \ of \ labor \ input.$$

Suppose that housecleaners can be hired at $10 per hour:

$$cost \ of \ one \ clean \ house \ = 5 \ hours \ \times \$10 \ per \ hour \ = \$50.$$

The cost of two clean houses is $100, the cost of three clean houses is $150, and so on.

More generally, suppose we take the linear production function and divide both sides by the level of productivity. We get

The cost of a single hour of labor is given by the wage. Thus we can write

This is the cost function of a firm, which is illustrated in Figure 9.2.5. Pay careful attention to the axes in Figure 9.2.3 and Figure 9.2.5. Figure 9.2.5 has hours of labor on the horizontal axis and output on the vertical axis. Figure 9.2.5 has output on the horizontal axis and costs ($= labor \ hours \ \times \ wage$) on the vertical axis.

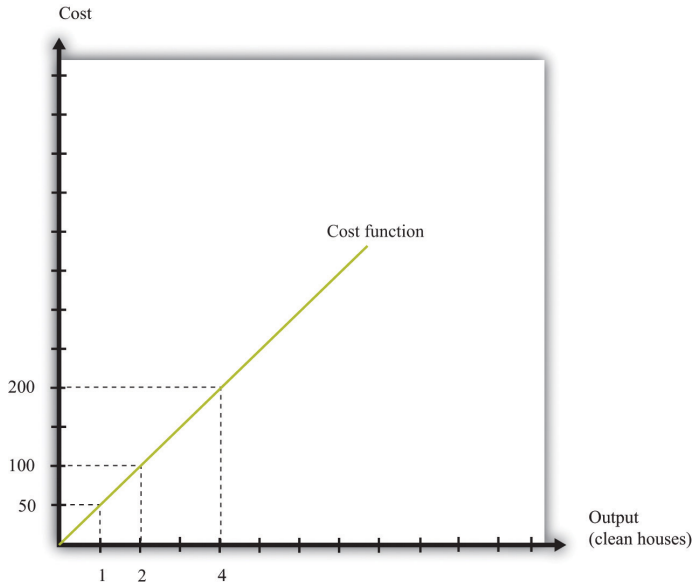


Figure 9.2.5: The Cost Function

The cost function shows the cost of producing different levels of output.

Marginal Cost

The cost function in Figure 9.2.5 is linear. Because the production function has a constant marginal product of labor, the cost function displays constant **marginal cost**. What about the case in which the production function has a diminishing marginal product? Then additional labor provides less and less output. Turning this around, it follows that producing each additional unit of output requires more and more labor. We show this in Figure 9.2.6. In this figure, the marginal cost is increasing, so the cost function gets steeper as we produce more output.

Toolkit: Section 31.14

You can review the definition of marginal cost in the toolkit.

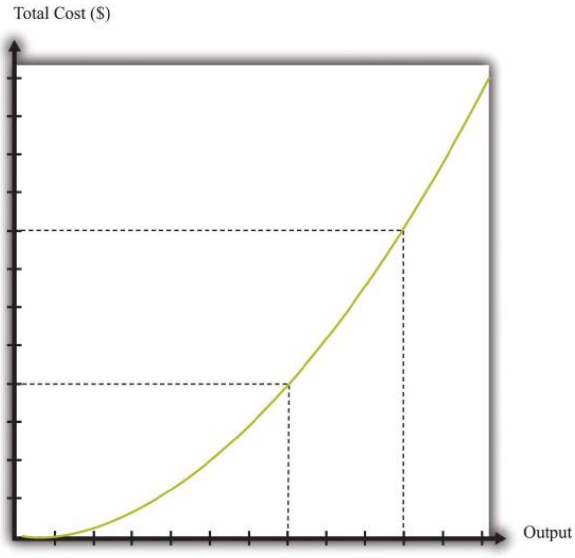


Figure 9.2.6: The Cost Function with a Decreasing Marginal Productivity of Lab

If a firm's technology exhibits a diminishing marginal product of labor, the cost function will increase at an increasing rate.

We show the marginal cost curve in Figure 9.2.7. In this example, marginal cost is a straight line, but this need not be the case in general.

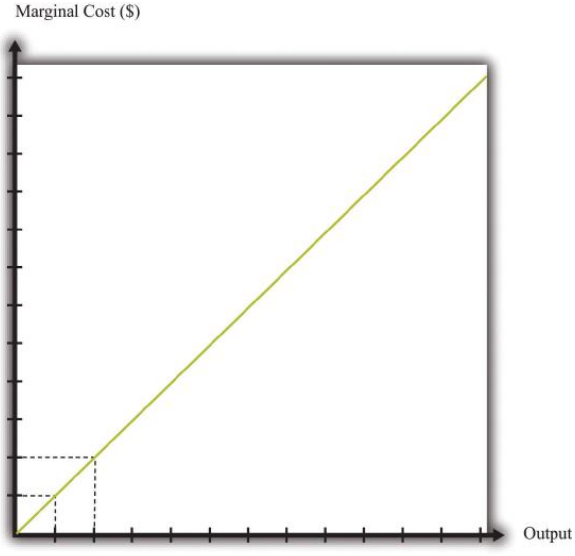


Figure 9.2.8: The Marginal Cost Function. If a firm's technology exhibits a diminishing marginal product of labor, then the marginal cost will increase as output increases.

Marginal cost depends on the following:

- The *cost* of inputs into the production process
- The *productivity* of the inputs into the production process

If the costs of inputs increase, then the marginal cost is higher is well. If the productivity of the inputs into the production function is higher, then the marginal cost is lower. In fact, marginal cost can be written as

We can see from this equation that when the marginal product of labor decreases, the marginal cost of production increases. We see also that an increase in the cost of inputs—in this case an increase in wages—leads to an increase in the marginal cost of production.

The Choice of Inputs

There is one issue that we are ignoring here. Firms typically have many different ways in which they can produce the same quantity of output. A firm might have a choice between two production processes: (1) a process that is simple and cheap to operate but wasteful of raw materials and (2) a process with recycling that uses fewer materials but is more complicated and costly to run. As another example, if a construction company needs to dig a ditch, it could employ 20 people and equip each with a shovel, or it could hire a single individual and a backhoe. Economists say that the first process is *labor intensive* because it requires a lot of labor relative to capital; they call the second process *capital intensive* because it requires a relatively large amount of capital.

In medium-sized or large firms, there is usually a specific functional area, called *operations*, that decides how to produce output. Operational decisions are governed in large part by technical or engineering considerations: what are the ways in which it is physically possible to transform inputs into the desired output? Operational decisions also have an economic component. Given that there may be many different ways to get the same final amount of output, which is the most cost effective? In economics, not surprisingly, we focus on the second of these questions and leave the first to engineers and other technical experts.

The basic principle is intuitive: operations managers tend to choose methods of production that economize on relatively expensive inputs. For example, much garment manufacture takes place in countries like China or Vietnam, where wages are low (that is, labor is relatively cheap). As a result, the production processes tend to be highly labor intensive, using a lot of workers relative to the amount of machinery. By contrast, garment manufacture in richer countries (where labor is much more expensive) tends to use methods of production that require fewer people and more machines.

We simply presume that the operations function of a firm is doing a good job and has succeeded in finding the cheapest way of producing the firm's output, taking into account both the technical aspects of production and the costs of different inputs. When we talk about the cost function of a firm, therefore, we are assuming that it gives us the *lowest cost* for producing each given level of output.

The Price/Output Decision

We have now completed our discussion of the first decision that managers must make: how to produce output. The production function tells us what a firm needs in terms of inputs—in this case, labor—to produce a given level of output. The more output a firm wants to produce, the more labor it will hire and the more jobs it will create. The cost function tells us the cost of producing different levels of output, and the marginal cost function tells us the cost of producing additional output.

Marginal cost is the critical ingredient in the next decision made by managers, which is selecting a point on the demand curve. We can think of managers as either choosing the price and then selling the quantity demanded at that price or choosing the level of output and selling it at the price that the market will bear. In either case, they are picking the point on the demand curve whereThis decision of the firm is also covered in detail in Chapter 7.

$$marginal\ revenue\ =\ marginal\ cost.$$

We show this decision graphically in Figure 9.2.9.

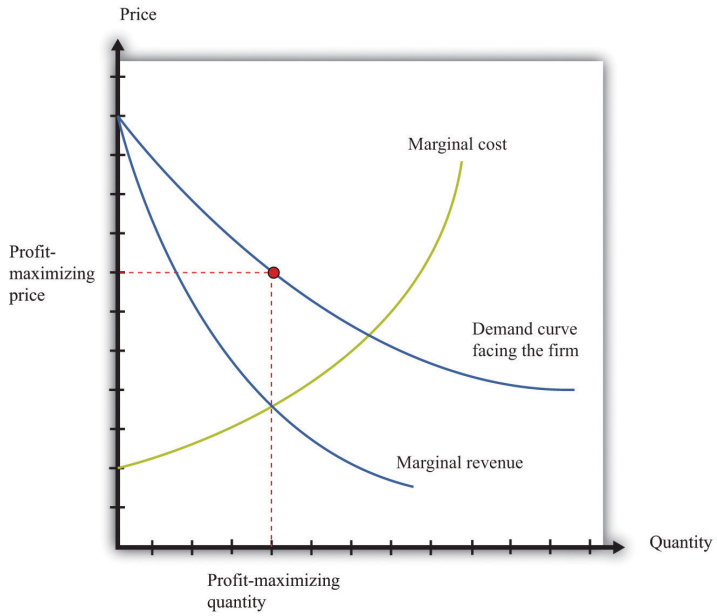


Figure 9.2.9: Output and Price Decisions of a Profit-Maximizing Firm

A profit-maximizing firm produces a quantity such that marginal revenue equals marginal cost, and the price is determined by the demand curve.

In our discussion of costs to this point, we have not specified whether we were talking about the nominal wage (that is, measured in dollars) or the real wage (that is, adjusted for inflation). The most important thing is being consistent. If we use the nominal wage when calculating our cost functions, then we end up with nominal costs. If we use the real wage, then we end up with real costs. And when we equate marginal revenue and marginal cost, we must be sure that we measure in nominal terms or real terms (not a mixture).

The distinction becomes important only when the general price level changes, so it is not central to our discussion here. When the price level is constant, we can always just suppose that it is equal to 1, in which case the nominal wage and the real wage are equal. Still, when we draw diagrams of the labor market, we typically put the real wage on the axis, so from here on we will explicitly suppose that we are measuring everything in real terms.

We can now explain labor demand by a firm. There are two steps:

1. As in Figure 9.2.9, a firm produces a level of output such that marginal revenue equals marginal cost.
2. Using Figure 9.2.3, a firm determines the amount of labor it needs to produce the output chosen in step 1.

We already know that the marginal cost of production depends on the real wage: decreases in the real wage lead to decreases in real marginal costs. Figure 9.2.10 shows how a decrease in the real wage affects the output and price decision of a firm. As the real wage decreases, the marginal cost of an additional unit of output decreases, so a firm will choose to produce more output. The price will decrease because the firm must lower the price to sell the additional output.

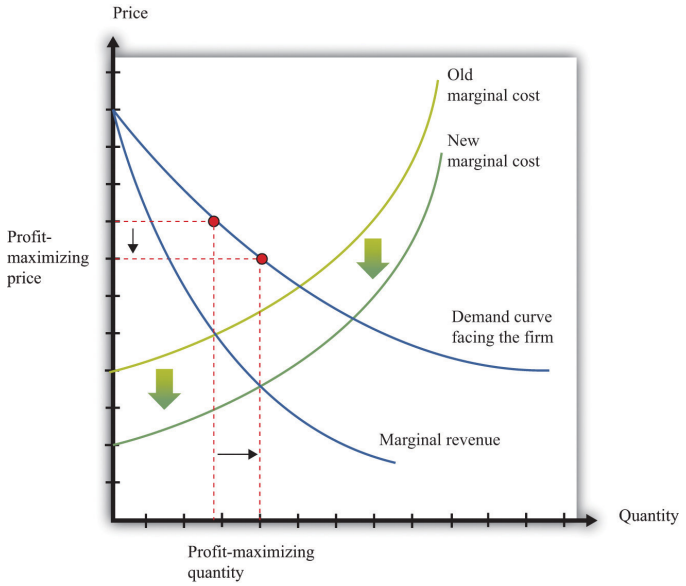


Figure 9.2.10: The Effect of a Change in Marginal Cost on a Firm's Choice of Output and Employment

When the real wage decreases, marginal cost decreases, so the firm reduces price, increases output, and creates jobs.

Because a firm wants to produce more output, it will demand more hours of labor. In other words, a decrease in wages leads to an increase in the quantity of labor demanded. The resulting inverse relationship between the real wage and the amount of labor demanded is shown in Figure 9.2.11.

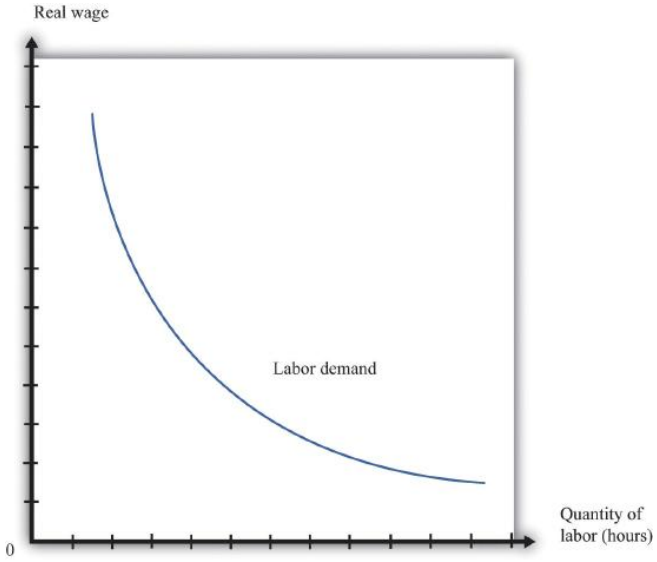


Figure 9.2.11: The Demand for Labor. As the real wage increases, the demand for labor input decreases.

The labor demand curve for a single firm is downward sloping. This is true for every firm in the labor market. The market demand curve for labor is obtained by adding together the demand curves of individual firms. So the market demand for labor is downward sloping as well.

Changes in Employment

We can now connect our understanding of the labor market with the data on net job creation that we showed in Figure 9.2.3. Based on what we have learned, there are three main reasons why jobs might be created or destroyed: (1) changes in the real wage, (2) changes in productivity, and (3) changes in demand.

Changes in the Real Wage

Changes in the cost of labor are one reason firms create or destroy jobs. Decreases in the real wage lead firms to produce more output and hire more workers, thus creating jobs. Increases in the real wage cause firms to produce less output and lay off workers. Going deeper, we can ask *why* the real wage might change. The answer comes from looking back at Figure 9.2.1. The real wage changes, causing a change in the quantity of labor demanded, if the labor supply curve shifts.

Population growth is one source of a shift in labor supply. As the number of workers in the economy grows, then total labor supply will shift. At a given wage, there will be more workers and hence the labor supply curve will shift to the right. Other things equal, this causes a decrease in the real wage.

Changes in labor market participation also shift the labor supply curve. A leading example of this is the increased participation of women in the labor market. In the United States, the fraction of women in the labor force rose from about 20 percent in 1950 to about 70 percent in 2000. Participation might also change because of workers' expectations about the *future* state of the labor market. If you are worried you won't have a job next year, you might want to work this year.

Changes in Productivity

Changes in productivity—more precisely, in the marginal product of labor—work exactly like changes in the real wage. Remember that marginal cost depends on both real wages and productivity. If productivity increases, perhaps because a firm has upgraded its capital equipment, then marginal cost decreases. Firms will produce more output and hire more labor. The opposite is true if productivity decreases: in this case, firms will produce less and destroy jobs.

Over long periods of time, productivity in an economy increases. This increase in productivity is driven largely by technological advances: firms get better at producing goods and services and so are able to produce them more cheaply. As workers' productivity increases, firms demand more labor at any given wage.

Changes in Demand

When a firm's product becomes more desirable in the market, the demand curve that it faces shifts outward. This shift in demand typically leads to an outward shift in marginal revenue, inducing a firm to produce more output and demand more labor. We show this in Figure 9.2.12: an outward shift of a firm's demand curve typically leads to an outward shift in labor demand.

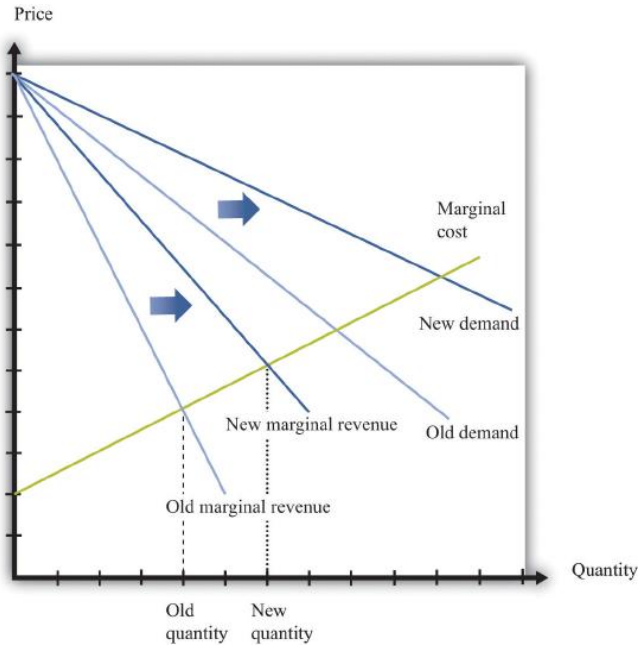


Figure 9.2.12: The Effect of a Change in Demand on a Firm's Choice of Output and Employment

An increase in demand typically leads to an increase in marginal revenue, which in turn induces firms to produce more output and create jobs.

Key Takeaways

- A firm produces a quantity such that the marginal cost of producing an extra unit of output equals the marginal revenue from selling that extra unit of output.
- The demand for labor depends on the level of productivity, the demand for a firm's product, and the cost of labor compared to the cost of other inputs in the production process.

Exercises

1. uppose the production function exhibits increasing marginal product of labor. What would it look like? What would the cost function look like in this case?
2. Marginal cost is defined as

   We also know that

   and

   $$change\ in\ cost = wage \times change\ in\ labor\ input.$$

   Show how you can use these three equations to derive the condition in the text that
3. Using Figure 9.2.12, what would happen to a firm's decision on prices, the quantity of output, and labor demand if the demand curve and marginal revenue curves shifted inward?

This page titled 9.2: How Do Firms Decide How Many Hours of Labor to Hire? is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

9.3: Entry and Exit 9.3: Entry and ExitEdit Update titlecancel

Learning Objectives

1. What is the difference between a fixed cost and a variable cost?
2. What factors determine if a firm should remain in business?
3. What is a sunk cost?

So far we have been thinking about a firm simply changing the number of labor hours that it wants to plug into its technology. Such job creation and destruction takes place at individual firms all the time. Some firms see an increase in productivity and hire more workers. Other firms see reduced demand for their output and destroy existing jobs. The joint creation and destruction of jobs underlies the net job creation we displayed earlier in Figure 9.3.3.

Yet the expansion and contraction of employment in existing plants is only one source of job creation and destruction. During an economic downturn, such as the severe recession that began in 2008, some firms closed factories, and other firms went completely out of business. For example, US automobile manufacturers, such as General Motors, responded to the decreased demand for cars by closing some of their existing manufacturing plants. This led to job destruction at these plants. At other times, when an economy is expanding, new firms enter into business and existing ones open new plants. Thus a complete picture of the job creation and destruction process requires us to understand the economics of entry and exit.

Only when the firm's managers know how much the firm is going to produce, the price at which it will sell it, and the cost of producing that output can they figure out profits and decide whether it is sensible to be in business at all. This logic applies to both managers of firms that are already in business and entrepreneurs who are considering starting a business. Firms also apply this logic to parts of their operations—for example, a firm may want to decide whether to shut down an existing plant or open a new one.

If a firm that is already in business discovers that its profits are too small to justify its other costs, then it should exit the market, shutting down its operations completely. If an entrepreneur is contemplating starting a new firm and calculates that the profits it will earn justify the costs of setting up operations, then we say that a firm enters the market.

In the previous section, we explained how job creation and destruction take place as firms expand and contract. When Walmart comes to town, however, much more is going on. The opening of a new Walmart means that some new jobs are created. Against that, existing stores may be forced to close down completely. Now that we have looked at a firm's price and output decisions, we are able to analyze entry and exit decisions.

Exit

Businesses do not stay around forever. At some point, they exit the market, destroying jobs in the process. Restaurants that were a big hit only a few years ago can quickly lose their luster and disappear from the scene. The same is true for many retail outlets. Manufacturing plants also close, taking jobs with them. Imagine, for example, that you own a small clothes retailing store. Then Walmart comes to your town. Now your customers have another place to buy their clothes, and you must decide whether to stay in business. You need to decide which is more profitable: staying in business or closing your business down and selling off any assets you possess.

To understand the factors influencing firm exit, we begin with a key distinction between different kinds of costs.

Toolkit: Section 31.14

Costs that are the same at all levels of production are called fixed costs. Costs that vary with the level of production are called variable costs. As an accounting identity, the total costs of a firm are divided up as follows:

$$total\ costs\ =\ fixed\ costs\ +\ variable\ costs.$$

Table 9.3.1 "Monthly Costs of Production" shows an example of **fixed costs**, **variable costs**, and **total costs** for your store. (To keep life simple, we treat all the different kinds of clothing you sell as if they were the same. Let us call them blue jeans.) The numbers in Table 9.3.1 "Monthly Costs of Production" are based on the following equation for costs:

$$total\ costs\ =\ 14,000 + 10 \times quantity.$$

Suppose your firm has fixed costs every month of $14,000. By definition, these fixed costs do not change as your level of output changes. Think of these as your *overhead* costs—for example, the cost of renting your retail space, utility bills, the wage of your sales clerk, and so on.

Table 9.3.1: Monthly Costs of Production

Quantity	Fixed Costs ($)	Variable Costs ($)	Total Costs ($)
0	14,000	0	14,000
1	14,000	10	14,010
2	14,000	20	14,020
…	…	…	…
200	14,000	2,000	16,000
400	14,000	4,000	18,000
600	14,000	6,000	20,000
800	14,000	8,000	22,000
1,000	14,000	10,000	24,000
1,200	14,000	12,000	26,000
1,400	14,000	14,000	28,000
1,600	14,000	16,000	30,000
1,800	14,000	18,000	32,000
2,000	14,000	20,000	34,000
2,200	14,000	22,000	36,000
2,400	14,000	24,000	38,000

By contrast, variable costs increase as the level of output increases. In this example, if output increases by one, variable costs increase by $10. You can think of this as the cost of purchasing your blue jeans from the wholesaler. For you, the cost of "producing"—that is, making available for sale—one more unit of output is $10. Figure 9.3.1 graphs the data from this table. Notice that even if your firm produces no output at all, it still incurs fixed costs.

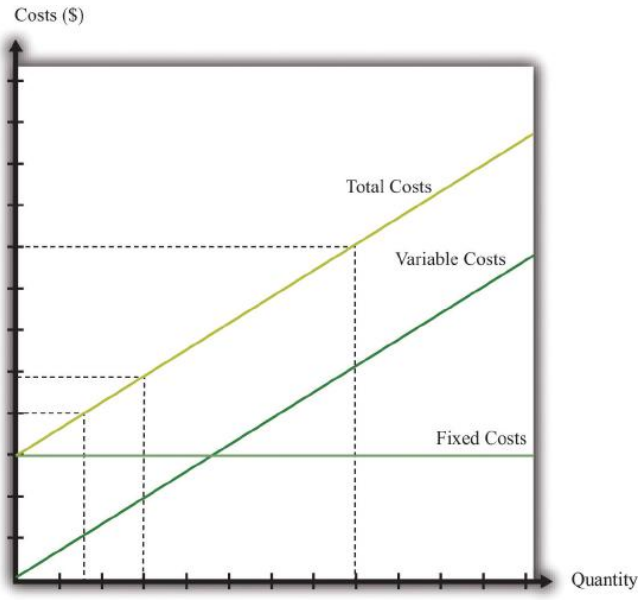


Figure 9.3.1: Total Costs, Fixed Costs, and Variable Costs. Fixed costs are the same at all levels of output. Variable costs increase as the quantity of output increases. Total costs equal fixed costs plus variable costs.

The Exit Decision

We are now ready to study the decision to continue in business or exit. You need to compare your revenues, defined as price times quantity, with the cost of running your business. Profit is the difference between revenues and costs, so

$$profit = total\ revenues - total\ costs = total\ revenues - variable\ costs - fixed\ costs.$$

Table 9.3.2: Demand and Profit before Walmart Comes to Town

Quantity	Price ($)	Total Revenues ($)	Variable Costs ($)	Fixed Costs ($)	Profits ($)
0	30	0	0	14,000	–14,000
200	29	5,800	2,000	14,000	–10,200
400	28	11,200	4,000	14,000	–6,800
600	27	16,200	6,000	14,000	–3,800
800	26	20,800	8,000	14,000	–1,200
1,000	25	25,000	10,000	14,000	1,000
1,200	24	28,800	12,000	14,000	2,800
1,400	23	32,200	14,000	14,000	4,200
1,600	22	35,200	16,000	14,000	5,200
1,800	21	37,800	18,000	14,000	5,800
2,000	20	40,000	20,000	14,000	6,000
2,200	19	41,800	22,000	14,000	5,800

Quantity	Price ($)	Total Revenues ($)	Variable Costs ($)	Fixed Costs ($)	Profits ($)
2,400	18	42,550	24,000	14,000	5,200

The demand for your blue jeans is shown in the first two columns of Table 9.3.2 "Demand and Profit before Walmart Comes to Town". Looking at this table, your profit is at its highest when you set a price at $20 and sell 2,000 pairs of jeans. In this case, you earn $6,000 per month. Your revenues are enough to cover your variable costs and your **fixed operating costs**.

After Walmart comes to town, the demand for your jeans shifts inward because shoppers start going to Walmart instead. Now your demand is as shown in Table 9.3.4 "Demand and Profit after Walmart Comes to Town".

Table 9.3.3: Demand and Profit after Walmart Comes to Town

Quantity	Price ($)	Revenues ($)	Variable Costs ($)	Fixed Costs ($)	Profit ($)
0	26	0	0	14,000	–14,000
200	25	5,000	2,000	14,000	–11,000
400	24	9,600	4,000	14,000	–8,400
600	23	13,800	6,000	14,000	–6,200
800	22	17,600	8,000	14,000	–4,400
1,000	21	21,000	10,000	14,000	–3,000
1,200	20	24,000	12,000	14,000	–2,000
1,400	19	26,600	14,000	14,000	–1,400
1,600	18	28,800	16,000	14,000	–1,200
1,800	17	30,600	18,000	14,000	–1,400
2,000	16	32,000	20,000	14,000	–2,000
2,200	15	33,000	22,000	14,000	–3,000
2,400	14	33,600	24,000	14,000	–4,400

In response to this decrease in demand, you should drop your price. Your profits are now maximized at $18. Unfortunately, at this price, you don't earn enough to cover your fixed costs. Your profits are –$1,200 a month. Should you remain in business? The answer to this question depends on when you ask.

Suppose you ask this question just *after* you have paid your monthly fixed operating costs. During the course of a month, you should stay in business because you are earning enough revenues to cover your variable costs. As soon as it is time to pay your monthly fixed cost again, though, you should choose to exit and close down your store. In this case, you would engage in job destruction by firing your sales clerk.

In this simple example, it is easy to see exactly when and why you should exit. A more general rule for when to exit is as follows. You should exit if

$$discounted\ present\ value\ of\ expected\ future\ profits\ <\ value\ of\ recoverable\ assets.$$

To make sense of this rule, we need to look at each part of it in turn.

- **Discounted present value.** Our previous example considered only a single month. In fact, you must think about the entire future. This means you must use the tool of **discounted present value** to add up profits earned in different months.For examples of discounted present value in action, look at Chapter 5 and Chapter 10.
- **Expected future profits.** Even though your price has decreased this month, it might not stay low forever. Perhaps your customers will decide, after they have tried Walmart, that they prefer your store after all. It would then be foolish to close down

your store immediately just because you fail to cover fixed costs in one month. This means you must make a decision in the presence of uncertainty: you don't know for sure if your customers will come back, and if they do, you don't know for sure that they will not go away again. As a simple example, suppose you think there is a 75 percent chance that the shift in your demand curve is permanent and a 25 percent chance that it will go back to its original position. Looking ahead and using the tool of **expected value**, you would calculate your expected profit as follows:

In this case, you still expect to make a small profit every month. Provided you were not too **risk-averse**, you would keep your store in business. Of course, after some months had gone by, you would probably have better information about whether your customers are truly coming back or not.

- **Value of recoverable assets.** If you are thinking of closing down your store, then you also need to look at your existing assets in the store. You may be able to sell off some of these assets. For example, you could perhaps sell your cash register or computers. We say that such assets are (partially) **recoverable assets** because you can get back a portion of what you originally paid for these assets.

Toolkit: Section 31.4, and Section 31.7

You can review the meaning and calculation of discounted present value, expected value, and risk-aversion in the toolkit.

Defining Fixed Costs

Our definition of fixed costs seems very straightforward. Unfortunately, it is not always easy to decide in practice whether a cost is fixed or variable. There are two main reasons for this:

1. **Time horizon.** Business planning must be carried out over multiple time horizons. You must decide what to do from one week to the next, from one month to the next, and from one quarter to the next. Costs that are fixed over short time horizons may be variable over longer time horizons. For example, suppose your contract with your employee says you must give her six weeks' notice prior to letting her go. Then her wages are a fixed cost when you are planning for the next six weeks but a variable cost over a longer horizon. Similarly, you may have to lease your store space yearly, in which case that cost is fixed until your lease next comes up for renewal. The bottom line is that whether you think of a cost as fixed or variable depends on your time horizon.
2. **"Lumpiness."** Some inputs are easier to vary than others. You can freely decide how many pairs of jeans to buy from your wholesaler, so your purchase of jeans to sell is a variable cost. Other inputs are "lumpy"; they are fixed over some ranges of output but variable over others. This means that some costs are fixed over some ranges of output but variable over others. For example, you might be able to sell up to, say, 10,000 pairs of jeans a month in your current store space. However, if you wanted to expand beyond that, you would no longer have enough room to store your inventory and provide an acceptable shopping space for your customers. Because the size of your store is not something you can vary smoothly, this is a lumpy input.

In fact, if we take a very long time horizon and very large ranges of output, there are few costs that are truly fixed.

Entry

We use very similar reasoning to think about a firm's decision to enter a market. When Walmart's senior management team contemplates opening a new store, they compare the costs of entry against the (discounted present value of the) profit they expect to earn once they enter.

What are some of Walmart's costs when it wants to open a new store?

- Searching for a suitable location
- Going through the necessary legal processes in the particular location
- Purchasing the land
- Dealing with public opposition (through lobbying, advertising, sending representatives to town council meetings, etc.)
- Designing the store
- Building the store
- Adjusting their supply chain logistics so as to be able to supply the store
- Hiring people to work there

You can probably think of many more. We call these Walmart's **entry costs**.

Toolkit: Section 31.14

Entry costs are the one-time fixed costs that a firm incurs when establishing a business. The toolkit has more discussion of other kinds of costs.

Such costs of establishing a business can be very substantial. Notice, by the way, that entry costs are for the most part truly fixed costs. Walmart must incur these costs before it can let a single customer inside; these costs are fixed no matter how long the time horizon; these costs are largely independent of Walmart's scale of operation.

The senior management team must also predict how much profit they expect the store to make. These forecasts are based on the idea that, once the store is opened, the store will set its prices and manage its operations to maximize its profits. Because the team will be uncertain about profits, they will need to use expected value calculations. They will also be counting on a profit flow for years if not decades and will need to use discounted present value calculations. Thus the appropriate decision rule for a firm is to enter ifIn Chapter 15, you will find an example of very similar economic reasoning. There we present a parallel rule for the situation where a firm is deciding whether or not to engage in innovation.

$$discounted\ present\ value\ of\ expected\ future\ profits\ >\ entry\ costs.$$

A firm is more likely to enter if

- the costs of entry are low,
- variable costs are low,
- the revenues from operating are high, and
- demand for its product is very inelastic.

Sunk Costs and Recoverable Costs

Firms that enter markets know that it is possible that they will exit again in the future. Because their profit flow is uncertain, they recognize that there may come a point where they will judge it better to close down their operations. If they close down their operations, they may be able to sell off some of their existing assets. Therefore, when deciding to enter, managers also take into account the extent to which their assets are recoverable. If they are likely to be able to reclaim most of the value of their assets, then entry is more likely to be profitable even if demand turns out to be lower than expected.

Specifically, we can divide entry costs into **sunk costs** and **recoverable costs**.

Toolkit: Section 31.14

A sunk cost is a cost that, once incurred, cannot be recovered. A recoverable cost is a cost that, once incurred, can be recovered.

Looking back at our list of Walmart's entry costs, we can see that many of these costs are sunk costs. All the planning and legal fees are completely tied to this store; if they end up not building the store, they cannot get any of the monies back. The building is a sunk cost as well. Other costs are at least partly recoverable. If they decide not to build the store, they can resell the land. If they have equipped the building with shelving and cash registers and then decide not to open the store, they can resell these assets or move them to another Walmart store instead.

Economic reasoning gives clear instructions about sunk costs: they should be irrelevant for any future decision. Whether it was a good or bad idea to build a store, any decisions made going forward should take into account only the future profitability of the store. For example, suppose that Walmart's entry costs were $100 million, of which $30 million were recoverable. Suppose also that Walmart's managers estimated the discounted present value of expected profits at $120 million and therefore decided to build the store. However, once it was built, they discover that they have badly overestimated demand. The managers revise their estimate of future profits by half to $60 million. They now regret having built the store; it was a bad investment. *But they should still keep the store open because it is earning more than they could obtain by closing the store and selling its assets.*

Even though the economic principle is clear, people frequently include sunk costs in their calculations instead of ignoring them. This is such a pervasive problem that it is given the name the **sunk cost fallacy**.

Toolkit: Section 31.14

The sunk cost fallacy is the mistake of including sunk costs in future-looking decisions, even though they should properly be ignored.

Key Takeaways

- A fixed cost is paid regardless of the level of output produced; a variable cost depends on the level of output produced.
- A firm should exit when the discounted present value of its future profits decreases below the value it can receive from selling its assets.
- A sunk cost is a cost that cannot be recovered, such as the cost of entry. This cost should have no effect on the decision to exit.

Exercises

1. Go back to our discussion of the data in Table 9.3.3 "Demand and Profit after Walmart Comes to Town". Explain why you should reduce your price after Walmart arrives in town.
2. Give an example of a fixed cost associated with taking this economics class. Is that cost sunk? How much of it can you recover if you stop taking the class?
3. Suppose the interest rate increases. Explain how that will lead more firms to exit the market. (Hint: think about discounted present value.)

This page titled 9.3: Entry and Exit is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

9.4: Search 9.4: SearchEdit Update titlecancel

Learning Objectives

1. What is the process of matching workers and vacancies?
2. What is the optimal strategy to follow when looking for a job?
3. How are wages determined in labor markets?

As we have seen, job creation and destruction occur because of the entry and exit of firms. Jobs are created when firms enter into an industry and destroyed when firms exit. Job creation and destruction also arise as a result of the hiring and firing decisions of existing firms. We have used the labor market as a device to help us understand these hiring and firing decisions.

If you have ever looked for a job, though, then you know there is more to the labor market than supply and demand. Several aspects of the way labor is traded do not fit neatly into this framework. Workers and firms devote time and money to finding one another: search is an important element of the job market. And wages are often determined by some type of bargaining process, perhaps between a single worker and a firm or between a firm and a union that represents many workers.

Internet Job Search

Internet job searches are now an established part of the way labor markets operate. If you are a worker looking for a job, you can go to a site like Monster.com (http://www.monster.com) or CareerBuilder.com (http://www.CareerBuilder.com) to search for vacancies posted by firms.The Department of Labor sponsors a website (www.careeronestop.org) filled with information, including compensation levels, for different occupations. Help with job search is available here as well. When you search on one of these sites, you are asked to provide information about the type of job you are looking for by providing the following:

- Keywords (the type of work you want to do)
- Categories (a description of the occupation)
- Location of the job
- Career level

In addition, you provide information about yourself, such as your work experience and education level. The search engine then provides a list of vacancies posted by firms matching these characteristics.

If there are potential matches for you, you are provided with a list of potential employers. Each will typically provide some information about the job. Sometimes this will include a salary range. These postings often include a description of the type of worker the firm is searching for, using phrases such as “team player,” “responsible,” “leadership skills,” or “people skills.” The next step is then up to you. Along with the job postings comes information about how to contact the firm. You can indicate your interest to the firm, and you may be called in for an interview. If that goes well, a job offer will follow. At this point, negotiation over compensation comes into play.

Eventually, you must decide whether or not to accept the job. What should you do? If you knew for sure that this was your dream job of a lifetime, the decision would be obvious: accept the job. But life is never that easy. In reality, you face considerable uncertainty over any job you are offered:

- What will the job really be like?
- What other options are there?

The first type of uncertainty has multiple dimensions. No matter how many brochures you read about a job, how many other workers you talk to, or how much time you spend watching someone at work, you still will not know everything about a job until you actually go to work. Even then, there are elements of a job that you will not know about until you have worked for many years. An example is promotion. When you consider a job, you will probably hear about opportunities for advancement if you stay with the firm. But whether or not you will be promoted is something that will be resolved in the future and is part of the uncertainty you face when you think about accepting a job.

The second type of uncertainty concerns the alternatives to the job you are considering. If you had a list of all possible jobs available to you, then you could consult that list and pick the best job. But, of course, there is no such list. Instead of being presented with a list, you have to search for a job. If you turn down the job you are offered today, you will not know for sure what

job will be available to you tomorrow. Uncertainty over how to respond to a job offer is very important for some workers but less so for others. The difference is determined by how easy it is to change jobs. We can illustrate the point with two extreme examples.

The first case is a job that offers lifetime employment. If you accept this job, it is yours forever. You will never be fired and—let us suppose—you can never quit either. Given this situation and faced with a job offer, what would you do? Presumably the first thought that comes to your mind is "be careful." You would not accept this job unless you were very sure it was a good match for you. If you are not sure, you should reject the job offer and search more.

The second case is a job that offers very short-term employment, on a week-by-week basis. If you accept this job, you are employed for the week; then you can choose to remain in the job (if it is still available) or leave to search for another one. Also, suppose that during your work time you can still keep an eye out for other jobs. It might be that you can check a computer that displays job ads, look at classified ads in the newspaper, or pass by a few shops advertising job openings during lunch. If you are offered this second type of job, there is no need to be very selective. Your employment is very temporary, and it is easy to change jobs.

The first kind of job is more descriptive of professional positions available to highly skilled workers, where employers are very selective about the type of person they hire. For these types of individuals, searching and changing jobs can be very lengthy and expensive. If they accept a job, it had better be right for them. The second kind of job is one you might be more familiar with as a student—a short-term job such as waiting tables, working as a secretarial temp, or selling in a retail outlet.

Search and Bargaining

The existence of Monster.com and similar search engines makes clear that the trade of labor services is quite different from the trade of, say, US government bonds. The return on a bond is the same regardless of who owns it. But the match between a firm and a worker is special. No two jobs are the same, and no two workers are the same. Also, if you want to buy a US government bond, you can simply call a broker to buy one for you. But if you run a restaurant and want to hire a worker with some very special skills in a particular location, there is no obvious person to call or place to go.

There are three stages of search and bargaining:

1. The *meeting* of workers looking for jobs with firms looking to fill vacancies
2. The *matching* of workers with certain characteristics with jobs requiring certain characteristics
3. The *determination of wages* through a bargaining process

These three elements correspond to the stages you might encounter when you look for a job. First, there is time spent looking for job opportunities. This might involve a recruiting program or search on the Internet. Once you have found a job opening, there is normally a second stage: an interview process. You will typically be interested in the characteristics of the job (such as wages, hours, benefits, promotion possibilities, and job security), and the firm will be interested in your characteristics (such as skills, experience, and trustworthiness). If both you and the firm think that the match is a good one, then the process moves to a third and final stage of bargaining to determine the compensation you will receive.

Searching with a Reservation Wage

We suppose that the bargaining process between a worker and a firm is very simple. The firm makes a take-it-or-leave-it offer. In other words, the firm gives the terms of its offer, including the compensation package and the working conditions, and the worker can then either accept or reject this offer. We also suppose that the offer can be summed up in terms of wages.

To see how this works, imagine that there are two firms each offering jobs at $10 per hour. One firm provides very flexible working hours, while the other requires you to work from 10 p.m. to 6 a.m. and sometimes on the weekend. The first job is evidently more desirable than the second one. If you would be willing to pay $4 an hour for the flexibility of the first job, then it is *as if* the second firm was offering a job at $6 per hour.

Once a worker has a wage offer in hand, should that person accept or reject it? The answer comes from balancing the benefits of having a job (and therefore a wage) right now versus waiting for another job to come along in the hope that it will pay a higher wage. To see how a worker would make this choice, here is a simple numerical example.

- There are only two possible wage offers: some firms offer $500 per week, and some offer $1,000 per week.
- It takes a week to search for a job. If a worker turns down the offer he gets this week, he will not get another offer until next week.
- If a worker accepts a job, he cannot then search for another job.

- The government offers unemployment insurance of $300 per week.

Suppose a worker gets a job offer of $1,000 per week. Then the decision is easy: accept that job. The more difficult case is when the worker gets an offer of $500 per week. By accepting this job, the worker gets $200 more than with unemployment insurance. But accepting the job also has an **opportunity cost**. It means that the worker loses out on the chance of getting the higher paying job next week. So what should the worker do?

If you think about this problem, you will probably realize that the answer depends on how likely the worker is to get the better job by waiting. If most of the available jobs are the ones that pay $1,000 per week, then it is likely to be worth waiting. On the other hand, if most of the jobs pay only $500 a week, then the worker might have to wait a long time for the better job, so it is likely better to accept the one that pays $500 a week. More generally, in a world where there are lots of different jobs paying lots of different wages, the best thing for the worker to do is to pick something we call the **reservation wage**. Workers can follow this rule:

- Accept a job if it offers a wage above the reservation wage.
- Reject a job if it offers a wage below the reservation wage.

Bargaining

If a worker and a firm meet and determine that the match is good, then they proceed to determine wages. There are two ways in which this might happen.If you read Chapter 6, you will see some close parallels between the mechanisms we discuss here and the ways in which a buyer and a seller may agree on a price. One possibility is the one we just discussed. Firms post vacant jobs and at the same time advertise the wage. If a worker qualifies for the job, then that worker will accept the job if the wage exceeds the reservation wage.

There can also be bargaining between a worker and a firm. A firm will make a profit based on the difference between the marginal product of hiring the extra worker and wages paid to the worker. So a firm will choose to hire the worker as long as the wage is below the marginal product of labor. This is a firm's valuation of the job. A worker will be willing to take a job as long as the wage exceeds the reservation wage. This is a worker's valuation of the job. In Figure 9.4.1 we show both.

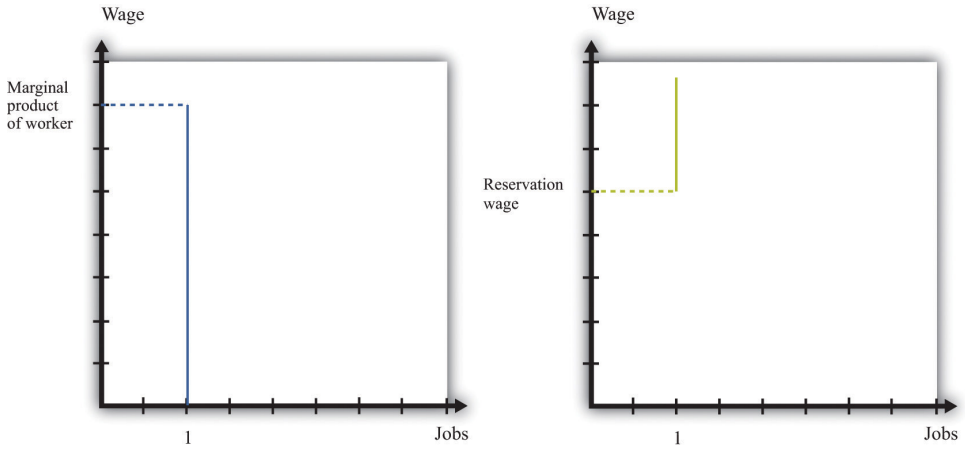


Figure 9.4.1: The Valuation of a Job. A firm follows the decision rule: "Offer the job if a worker's marginal product exceeds the wage." A worker follows the decision rule: "Accept the job if the wage exceeds the reservation wage."

As long as the marginal product of labor for a worker is higher than a worker's reservation wage, there is something to gain by employing a worker. The match is potentially a good one. But how will these gains be split? The answer depends on the relative bargaining power of a worker and a firm, which in turn depends on the information that they each possess.

As an example, suppose that a firm knows a worker's reservation wage and can make a take-it-or-leave-it offer. It would then offer a real wage slightly above the worker's reservation wage. The worker would accept the job, and all the surplus from the employment relationship would flow to the firm. At the other extreme, suppose that a worker knows his marginal product at the firm, so the worker can make a take-it-or-leave-it offer. Then the worker will offer to work at a real wage slightly below his marginal product. The firm would accept the offer, and all the surplus of the employment relationship would flow to the worker.

Another source of bargaining power for a worker is the other options available to him. If a worker comes into a negotiation with a good job offer from another firm in hand, then this will increase his reservation wage.

Workers can also enhance their bargaining power by negotiating together. Firms typically have many workers. Sometimes these workers group together, form a union, and bargain jointly with a firm. When workers organize in this way, they generally have more bargaining power because they can threaten to strike and shut down the firm. In this way, unionized workers get more of the surplus from their jobs.

In reality, workers don't know their exact marginal product and firms don't know the exact reservation wage of their workers. Not surprisingly, bargaining in such situations is more complicated to analyze. Sometimes, gains from trade are not realized. Suppose, for example, that a worker's reservation wage is below his marginal product. But a firm, thinking that a worker's reservation wage is really low, makes the worker a very low wage offer. If this wage offer is below the worker's reservation wage, then the worker will decline the offer and search again—despite the fact that there were gains to trade. Unfortunately, private information prevented the firm and the worker from realizing these gains.

Posting Vacancies

The final element in the search process is the vacancies posted by firms. You can see these vacancies on Monster.com, in the newspaper, and in magazines. Vacancies are costly to post, and it is expensive to evaluate workers. They are the analogy on a firm's side to the costly search on a worker's side.

You can think of a firm's decision of posting vacancies as being very similar to the labor demand for a firm. Firms want to expand output and thus post more vacancies whenever the marginal revenue of selling an extra unit of output increases relative to marginal cost. This could happen, for example, if the demand curve faced by a firm shifts outward. To expand output, the firm needs to hire more workers. It does so by posting vacancies, interviewing workers, and eventually bargaining with the best qualified ones to fill the open positions.

We noted earlier that labor demand also depends on wages: as the real wage decreases, a firm's real marginal cost decreases, so it will want to hire more workers and expand output. When we think of search and bargaining, say through Monster.com, there is no "market wage" that a firm takes as given. Instead, the wage comes from a bargain between a worker and the firm. But the wage that is eventually agreed on will depend on the outside options of workers and firms. As the prevailing wage in the market decreases, a firm will be able to hire workers at a lower wage and will choose to post more vacancies and expand its workforce.

Key Takeaways

- The search process brings together workers and vacancies of firms. This process lies behind the supply and demand curves for labor.
- For many searches, it is best to follow a reservation wage strategy: accept a job if and only if the wage exceeds the reservation wage.
- Wages are determined through a bargaining process. Sometimes this is through a take-it-or-leave-it offer of the firm. Often there is bargaining between a firm and its workers (or their union) to share the surplus of the employment relationship.

Exercises

1. If a worker becomes very impatient, what will happen to his reservation wage?
2. When Walmart comes to town, what will happen to the vacancies posted by competing stores? What will happen to the reservation wages of salespeople?

This page titled 9.4: Search is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

9.5: Government Policies 9.5: Government PoliciesEdit Update titlecancel

Learning Objectives

1. What government policies impact job creation and destruction?
2. What are the effects of trade on job creation and destruction?

Governments are very interested in job creation. A political leader whose economy loses a large number of jobs without creating new ones is unlikely to be reelected. On a local level, state and local governments compete fiercely to have firms locate in their region by offering lucrative tax reductions. This is seen as a way to create jobs in the local economy. We now examine some of these policy interventions and trace out their implications, focusing on three policies: restrictions on closing plants, policies that promote small businesses, and trade policies.

Plant-Closing Restrictions

In the United States, if you want to close a factory, you do not have to have approval of the government, but an act of Congress—the Worker Adjustment and Retraining Notification ActUS Department of Labor, "Other Workplace Standards: Notices for Plant Closings and Mass Layoffs," elaws Employment Law Guide, accessed March 14, 2011, www.dol.gov/compliance/guide/layoffs.htm.—requires you to announce your intentions ahead of time. According to the US Department of Labor, "The Worker Adjustment and Retraining Notification Act (WARN) protects workers, their families, and communities by requiring employers to provide notification 60 calendar days in advance of plant closings and mass layoffs."US Department of Labor, *The Worker Adjustment and Retraining Notification Act (WARN)*, accessed January 22, 2011, www.dol.gov/compliance/laws/comp-warn.htm.

This law was passed in 1988 during a time of higher than average unemployment in the United States. Similar restrictions apply in some European countries, such as Spain and France. You cannot simply close an unproductive plant; employees must be given advance notice, and government approval may be required. Such restrictions on plant closings are intended to reduce job destruction. After all, if you make something more expensive to do, then less of it will be done. If it becomes more expensive to close plants, then fewer jobs will be destroyed by the exit of plants.

Economists point out, however, that the incentives of such policies are complicated and go beyond the effects on job destruction. To see why, think about our earlier discussion of entry and exit. When a firm decides to enter into an industry, it compares the profit flow from operating to the entry cost. When a firm thinks about the profits it will earn if it enters, it recognizes that if the demand for its product disappears, it can exit and thus avoid periods of negative profits. But if you take this option to exit a market away from a firm, then the value of entering an industry will decrease. Fewer firms will enter, and fewer jobs will be created. Thus laws that make it costly to close plants will also reduce job creation. The effect on *net* job creation is unclear.

Small Business Promotion

Started in the 1950s, the Small Business Administration (SBA; http://www.sba.gov) is a US government agency whose goal is to protect and promote small businesses. Small firms obtain preferential treatment in terms of taxes, regulation, and other policies. Part of the argument in favor of promoting and protecting small businesses is the view that job creation is centered on these firms. According to the SBA website, small businessesSee http://www.sba.gov/sites/default/files/files/leg_priorities112th.pdf.

- represent 99.7 percent of all employer firms and
- have generated 60–80 percent of net new jobs annually over the last decade.

So it appears that small businesses are critical to an economy.

We must remember that these are indeed small firms, however. Suppose there were 5 firms in an economy. Four of them have 2 workers, and the fifth has 92 workers. The typical firm then has 2 workers: 80 percent of the firms in this economy have 2 or fewer workers. From the perspective of workers, though, things are rather different. Ninety-two percent of the workers are employed by the single large firm. If you ask workers how many employees are in their firm, the typical worker will say 92. *Most firms have few workers, but most workers are employed by the large firm.*

This is not far from the reality of the US economy, where much economic activity (employment and output) is centered on relatively few firms. A recent study of about 5.4 million businesses found that 182,000 of them operate multiple units. Dividing 182,000 by 5.4 million, we learn that these larger firms are less than 4 percent of the total number of firms. But they account for

about 61 percent of the revenue of the business sector of the economy. So most firms are relatively small, but those that are large are huge compared to the rest.Steven J. Davis et al., “Measuring the Dynamics of Young and Small Business: Integrating the Employer and Non-employer Universes” (NBER Working Paper 13226, 2007), accessed January 30, 2011, http://www.nber.org/papers/w13226.

Davis, Haltiwanger, and Schuh point out that “large firms and plants dominate the creation and destruction of jobs in the manufacturing sector.”See Steven J. Davis, John C. Haltiwanger, and Scott Schuh, *Job Creation and Destruction* (Boston, MA: MIT Press, 1998), chap. 7, sect. 4. Larger firms and plants both destroy and create more jobs. For example, at a job destruction rate of 10 percent a year, a small firm of 50 workers will destroy 5 jobs, while a large plant with 1,000 workers will destroy 100 jobs. So even if the job creation and destruction *rates* are higher for small firms, this does not necessarily mean that these small entities create and destroy more *jobs* than large firms do.

This is not to say that small firms are unimportant. Most of the large firms in the economy started small. Likewise, all the older, more successful firms were once young firms. So any impact the SBA has on either small or young firms will influence these firms as they age and grow. However, the rationale for the SBA is not completely clear. Normally, government interventions are based on the idea of either correcting some problem with the operation of free markets or redistributing resources. It is not clear whether the SBA fulfills either of these roles.

Trade Policy

Job creation and destruction are also affected by things that happen outside US borders. The removal of trade barriers allows countries to benefit more fully from the gains from trade. But in the process, some jobs are destroyed, while others are created.

Job destruction frequently takes center stage during debates on trade policy. In the early 1990s, for example, the United States was contemplating a reduction in trade barriers with its neighbors—Canada and Mexico—through negotiation of the North American Free Trade Agreement (NAFTA). Ross Perot, a third-party candidate for the US presidency in 1992 and 1996, was extremely critical of NAFTA. His focus was on job destruction, and he was famous for forecasting “a giant sucking sound” as employment opportunities moved from the United States to Mexico in response to NAFTA.

The loss of some jobs from a reduction in trade barriers is part of the adjustment one would expect. For countries to reap the gains from trade brought about by the removal of trade barriers, production patterns across countries must change. That process leads to job destruction and creation. Firms that used to produce certain goods in one country exit, as firms in other countries start to produce those goods instead. Workers at the exiting firms will certainly lose their jobs, but other jobs are created in the economy at the same time.

NAFTA was implemented in January 1994. More than 15 years later, it is still difficult to say exactly what the effects were and will be of NAFTA. Economics is not a laboratory science. It is not possible to subject the economies of Canada, Mexico, and the United States to this reduction in trade barriers, holding everything else the same. Instead, we have to look at data from before and after 1994 to try to infer the effects of NAFTA. But of course many other economic factors have also changed over this period. In parts of the United States where manufacturing jobs have been lost over the last 15 years, there is a tendency to hold NAFTA responsible. In fact, there is little evidence that NAFTA led to net job destruction.

What *has* happened over the last decade is that the US manufacturing sector has been exposed to increased competition from other countries, most notably China. It is this trade that has had a bigger impact on US manufacturing. At the same time, this has meant that NAFTA has been less of a success story for the *Mexican* economy than was predicted and hoped, as US consumers have purchased very cheap goods from China rather than Mexico.

Key Takeaways

- In the United States, firms are able to close plants if they choose to do so. This is not the case in all countries.
- The government promotes small businesses, viewing them as a source of job creation.
- The reduction of trade barriers creates new jobs and destroys others.

Exercises

1. If plants, once opened, were never allowed to be closed, what would this do to the incentives of a firm to open a plant?
2. In your college classes, what is the analogue of the statement that *most firms have few workers, but most workers are employed by the large firm*? Is it that most classes are small, but most students are in large classes?

This page titled 9.5: Government Policies is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

9.6: End-of-Chapter Material 9.6: End-of-Chapter MaterialEdit Update titlecancel

In Conclusion

During election season in the United States, the adverse effects of trade on jobs are often talked about extensively. In the 2008 Michigan presidential primary election, for example, candidates offered different ways in which they claimed they would help the automobile industry and bring more jobs to the Michigan economy. In the South Carolina primary in the same year, job losses in the textile industry received lots of attention. One large textile manufacturer in the region, Swift, had been cutting jobs steadily and was closing up, apparently planning to move production to South America. Individuals who lost their jobs due to this closing reported that they experienced a period of unemployment as they searched for a new job. Some found new jobs, either in an automobile assembly plant or working on optic fibers. Others moved from working in manufacturing to services. This is job creation and destruction in action. It happens all the time, all across the world.

Viewed abstractly, job creation and destruction are healthy processes for an economy. Through the process of job creation and destruction, workers are induced to move from less productive to more productive jobs. Such movement enhances the overall productivity of an economy. From the perspective of individual workers, however, the process looks very different. Job destruction means that people lose their jobs, which are a source of income and perhaps also of pride and dignity. They may have to spend some period of time unemployed, and they may lose important benefits, such as health insurance. They may have to relocate in search of jobs that are being created elsewhere; such relocation can be difficult and costly.

In sum, although the productivity of an economy as a whole may increase, this need not translate into improvements for workers who lose their jobs. Some find higher-paying jobs, but others, particularly those with few skills, see their wages decrease. One of the big challenges faced by governments and policymakers is to encourage the efficient reallocations of workers while minimizing the individual hardships that workers confront.

Key Links

- US Census Bureau website on employment dynamics: lehd.did.census.gov/led/index.html
- Bureau of Labor Statistics data on the labor force characteristics: http://www.bls.gov/cps
- Facts about working for Walmart: http://walmartstores.com/pressroom/factsheets
- Information about NAFTA: http://www.export.gov/fta/NAFTA/index.asp?dName=NAFTA
- Department of Labor on WARN: www.dol.gov/compliance/laws/comp-warn.htm

exercises

1. A statistic called “unit labor cost,” which is the cost per unit of output of the labor input, is often calculated. How is this different from marginal cost?
2. Would you think that a firm’s managers would have a different viewpoint than its workers on whether or not a plant should be shut down?
3. All else being the same, which firms would you expect to be more capital intensive—those with labor contracts that pay high wages or those with easy access to funds on capital markets?
4. How is the labor demand curve of a firm influenced by the cost of other inputs, such as energy?
5. If a firm operates with high fixed costs, should it set a higher price for its output to be able to cover those fixed costs?
6. Would a firm ever remain in business even though it is earning negative profits in the current year? How does this decision depend on the interest rate?
7. Besides labor, what other markets can you think of where search is important?
8. All else being the same, who will have a higher reservation wage—someone who can receive unemployment insurance for 13 weeks or someone who can receive unemployment insurance for 26 weeks?
9. What do you think is the role of “friends” in helping you find your first job? What about subsequent jobs?
10. In economic downturns, what happens to the ratio of unemployed workers to vacancies?
11. Why do people quit jobs? Do you think that the number of job quitters is higher or lower during economic downturns??
12. In many European countries, it is very difficult for a firm to close one of its plants. What might be the effects of an increase in the cost of closing a plant on job creation, destruction, and reallocation?

13. Suppose that the establishment of a Walmart in a nearby town led to the creation of 100 jobs in that store and the destruction of 150 retail jobs in the town. Is the net loss of 50 jobs enough reason to oppose the opening of the Walmart? What other benefits might a Walmart bring? What other costs might it impose?
14. What might be the effects of a reduction in child-care costs on an unemployed worker's reservation wage?

Spreadsheet Exercise

1. Revisit Table 9.3.1 "Monthly Costs of Production" through Table 9.3.3 "Demand and Profit after Walmart Comes to Town". Suppose your fixed operating costs were $12,000 instead of $14,000. Redo the tables with this change. Should you stay in business after Walmart arrives? Explain.

Economics Detective

1. The Bureau of Labor Statistics produces data on labor turnover called JOLTS (http://www.bls.gov/jlt). Using that information, create a table and a plot of data to illustrate what has happened to job openings and the quit rate since January 2007. How would you explain these findings?

This page titled 9.6: End-of-Chapter Material is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

10: Making and Losing Money on Wall Street

10.1: Financial Roller Coasters
10.2: A Walk Down Wall Street
10.3: The Value of an Asset
10.4: Asset Markets and Asset Prices
10.5: Efficient Markets
10.6: End-of-Chapter Material
10.7: Appendix- A General Formulation of Discounted Present Value

---

This page titled 10: Making and Losing Money on Wall Street is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

10.1: Financial Roller Coasters 10.1: Financial Roller CoastersEdit Update titlecancel

In 2006 and 2007, the financial district in Shanghai, China, was in a frenzy. Figure 10.1.1 shows the value of stocks in that market since its inception in 2000. Starting in early 2006, the value of stocks traded on this market exploded. The market rose by 130 percent in 2006; by May 2007, it was up over 50 percent for that year.See finance.yahoo.com/q/hp?s=000001.SS&=00&=1&=2006&=11&=31&=2007&=d&=66&=330. The market peaked in late 2007 and is currently at about 50 percent of that value. A lot of money was made by those who invested in the Shanghai market. And unfortunately a lot of money was lost.



Figure 10.1.1: Shanghai Stock Exchange Index. This figure shows the closing prices on the Shanghai stock exchange between January 2000 and April 2010. Stock prices rose rapidly from 2006, peaking in October 2007, but decreased substantially over the next year. In late 2009, stocks rebounded again.

These gains attracted many investors. Funds from abroad poured into Shanghai. The savings accounts of Chinese households were another source of investment funds. From a *People's Daily* Internet article posted on May 13, 2007, we learn the following: "More than 70 billion yuan (9.1 billion U.S. dollars) was transferred from savings accounts in Shanghai to stock trading accounts in the first four months of this year, the Shanghai branch of the People's Bank of China estimated on Saturday. In April alone, [savings deposits denominated in Chinese currency] with Chinese banking institutions decreased by 8.5 billion yuan (1.1 billion U.S. dollars).""Chinese Pour Savings Deposits into Stock Market," *People's Daily*, May 13, 2007, accessed March 14, 2011, http://english.peopledaily.com.cn/200705/13/eng20070513_374113.html.

During May 2007, stories circulated about households spending many hours carefully evaluating individual stocks and market returns. At the same time, it appeared that many relatively uninformed individuals were simply betting on the market, gambling on a quick return.

We said that some investors made money in the Shanghai market. Does that mean there is a lot of money *to be made* by investing in that market? These phrases sound similar but mean very different things. It is one thing to look back at a market and say you could have made money investing in that market. It is quite another to forecast that you will be able to make a high return in a market in the future. Investors who were attracted to the market in late 2007 had a very different experience: they lost a lot of money. Those who came into market in late 2008 were again able to profit as the market value rose over the following year.

In this chapter, we study the markets for different kinds of assets. Assets include stocks—such as are traded in Shanghai, on Wall Street, and in other financial centers around the globe—but, as we will see, there are many other kinds of assets as well. Information on assets is easy to obtain. If you open almost any newspaper, the business section contains an enormous amount of detailed information on stocks sold in a variety of markets. That same section will contain information on bonds, which are another type of frequently traded asset. Part of our interest in this chapter is defining these assets more precisely. The terms *stocks* and *bonds* are used commonly, but we want to understand exactly what these assets are and how they are traded.

As we wrote this chapter, we had no idea whether we, too, should be putting our personal savings in the Shanghai stock exchange or in some other market around the globe. In the middle of 2007, it looked as if the surge in the Shanghai market was over. Market participants were concerned that the time of high gains had ended. Yet by November 2007, market values had again started to escalate. And then, as we said, the market peaked in late 2007 and decreased rapidly for the next year. This is part of the story of asset markets. They are extremely volatile and unpredictable. When you see these high returns in Shanghai and other markets, you might wonder:

"Can I get rich by trading stocks and bonds?"

Road Map

This chapter begins with a walk down a fictionalized Wall Street, where we describe various kinds of assets. We focus mainly on financial markets, although we will look at other assets as well. Financial markets are familiar to many of us from the financial pages of newspapers or reports on the evening news. Such markets provide a link between borrowers and lenders (Figure 10.1.2). Many of us are borrowers from banks, perhaps because we have a student loan, a car loan, or a mortgage for a house. Much of what we borrow from banks comes from deposits placed in banks by other households. Firms also borrow in the financial markets. They issue stock and sell bonds in financial markets to finance their investment in new factories and machines.



Figure 10.1.2. Financial markets link borrowers and lenders.

We then turn to a discussion of the *pricing* of assets. We begin by thinking about an unusual asset: a fruit tree. A fruit tree gives us a certain amount of fruit each year, and the value of the tree depends on the value of the fruit it produces. We explain how to calculate the *value* of a fruit tree that lives for several years and yields an uncertain crop, and we show how exactly the same principles apply to the valuation of stocks, bonds, houses, and other assets. Finally, we explain why—if financial markets are functioning well—the *price* of an asset will equal its value.

Finally, we ask whether it is easy to make money by trading assets. We explain that the gains and losses from trading assets are based on two factors: (1) luck and (2) the skill of investors who quickly recognize profit opportunities before others notice these opportunities. If financial markets are functioning well, then it is very difficult for the casual investor to make money consistently by trading financial assets. And even if—as many believe—financial markets do not function perfectly, this still does not mean that there is easy money to be made.

This page titled 10.1: Financial Roller Coasters is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

10.2: A Walk Down Wall Street 10.2: A Walk Down Wall StreetEdit Update titlecancel

Learning Objectives

1. What are the different types of assets traded in financial markets?
2. What can you earn by owning an asset?
3. What risks do you face?

Wall Street in New York City is the financial capital of the United States. There are other key financial centers around the globe: Shanghai, London, Paris, Hong Kong, and many other cities. These financial centers are places where traders come together to buy and sell assets. Beyond these physical locations, opportunities for trading assets abound on the Internet as well.

We begin the chapter by describing and explaining some of the most commonly traded assets. Ownership of an **asset** gives you the right to some future benefit or a stream of benefits. Very often, these benefits come in the form of monetary payments; for example, ownership of a stock gives you the right to a share of a firm's profits. Sometimes, these benefits come in the form of a flow of services: ownership of a house gives you the right to enjoy the benefits of living in it.

Stocks

One of the first doors you find on Wall Street is called the *stock exchange*. The stock exchange is a place where—as the name suggests—stocks are bought and sold. A **stock (or share)** is an asset that comes in the form of (partial) ownership of a firm. The owners of a firm's stock are called the shareholders of that firm because the stock gives them the right to a share of the firm's profits. More precisely, shareholders receive payments whenever the board of directors of the firm decides to pay out some of the firm's profits in the form of **dividends**.

Some firms—for example, a small family firm like a corner grocery store—are privately owned. This means that the shares of the firm are not available for others to purchase. Other firms are publicly traded, which means that anyone is free to buy or sell their stocks. In many cases, particularly for large firms such as Microsoft Corporation or Nike, stocks are bought and sold on a minute-by-minute basis. You can find information on the prices of publicly traded stocks in newspapers or on the Internet.

Stock Market Indices

Most often, however, we hear not about individual stock prices but about baskets of stocks. The most famous basket of stocks is called the *Dow Jones Industrial Average* (DJIA). Each night of the week, news reports on the radio and television and newspaper stories tell whether the value of the DJIA increased or decreased that day. The DJIA is more than a century old—it started in 1896—and is a bundle of 30 stocks representing some of the most significant firms in the US economy. Its value reflects the prices of these stocks. Very occasionally, one firm will be dropped from the index and replaced with another, reflecting changes in the economy.You can learn more about the DJIA if you go to NYSE Euronext, "Dow Jones Industrial Average," accessed March 14, 2011, www.nyse.com/marketinfo/indexes/dji.shtml.

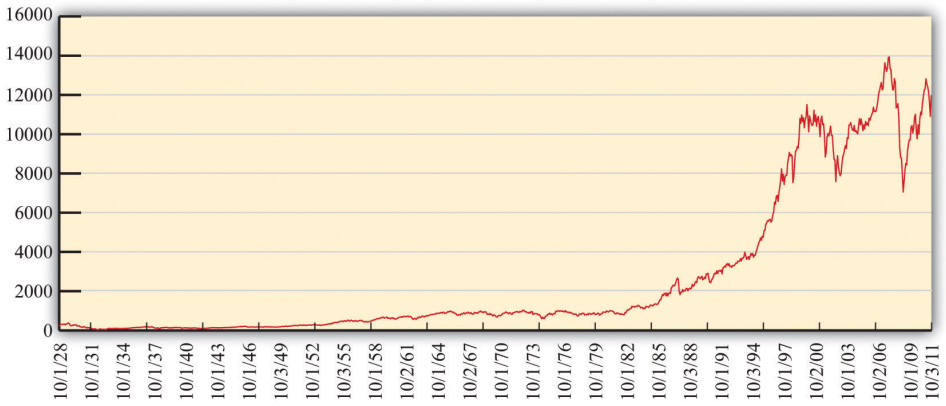


Figure 10.2.1: The DJIA: October 1928 to July 2007. This figure shows the closing prices for the DJIA between 1928 and 2010.

Source: The chart is generated from finance.yahoo.com/q?s=^DJI.

Figure 10.2.1 shows the Dow Jones Industrial Average from 1928 to 2011. Over that period, the index rose from about 300 to about 12,500, which is an average growth rate of about 4.5 percent per year. You can see that this growth was not smooth, however. There was a big decrease at the very beginning, known as the stock market crash of 1929. There was another very significant drop in October 1987. Even though the 1929 crash looks smaller than the 1987 decrease, the 1929 crash was much more severe. In 1929, the stock market lost about half its value and took many years to recover. In 1987, the market lost only about 25 percent of its value and recovered quite quickly.

One striking feature of Figure 10.2.1 is the very rapid growth in the DJIA in the 1990s and the subsequent decrease around the turn of the millennium. The 1990s saw the so-called Internet boom, when there was a lot of excitement about new companies taking advantage of new technologies. Some of these companies, such as Amazon, went on to be successful, but most others failed. As investors came to recognize that most of these new companies would not make money, the market fell in value. There was another rise in the market during the 2000s, followed by a substantial fall during the global financial crisis that began around 2008. Very recently, the market has recovered again.

If these ups and downs in the DJIA were predictable, it would be easy to make money on Wall Street. Suppose you knew the DJIA would increase 10 percent next month. You would buy the stocks in the average now, hold them for a month, and sell them for an easy 10 percent profit. If you knew the DJIA would decrease next month, you could still make money. If you currently owned DJIA stocks, you could sell them and then buy them back after the price decreased. Even if you don't own these stocks right now, there is still a way of selling first and buying later. You can sell (at today's high price) a promise to deliver the stocks in a month's time. Then you buy the stocks after the price has decreased. This is called a *forward sale*. If this sounds as if it is too easy a way to make money, that's because it is. The ups and downs in the DJIA are not perfectly predictable, so there are no easy profit opportunities of the kind we just described. We have more to say about this later in the chapter.

Although the DJIA is the most closely watched stock market index, many others are also commonly reported. The *Standard and Poor's 500* (S&P 500) is another important index. As the name suggests, it includes 500 firms, so it is more representative than the DJIA. If you want to understand what is happening to stock prices in general, you are better off looking at the S&P 500 than at the DJIA. The *Nasdaq* is another index, consisting of the stocks traded in an exchange that specializes in technology-based firms.

We mentioned earlier that the DJIA has increased by almost 5 percent per year on average since 1928. On the face of it, this seems like a fairly respectable level of growth. Yet we must be careful. The DJIA and other indices are averages of stock prices, which are measured in dollar terms. To understand what has happened to the stock market in real terms, we need to adjust for inflation. Between 1928 and 2007, the price level rose by 2.7 percent per year on average. The average growth in the DJIA, adjusted for inflation, was thus 4.8 percent − 2.7 percent = 2.1 percent.

The Price of a Stock

As a shareholder, there are two ways in which you can earn income from your stock. First, as we have explained, firms sometimes choose to pay out some of their income in the form of dividends. If you own some shares and the company declares it will pay a dividend, either you will receive a check in the mail or the company will automatically reinvest your dividend and give you extra shares. But there is no guarantee that a company will pay a dividend in any given year.

The second way you can earn income is through **capital gains**. Suppose you own a stock whose price has gone up. If that happens, you can—if you want—sell your stock and make a profit on the difference between the price you paid for the stock and the higher price you sold it for. Capital gains are the income you obtain from the increase in the price of an asset. (If the asset decreases in value, you instead incur a *capital loss*.)

To see how this works, suppose you buy, for $100, a single share of a company whose stock is trading on an exchange. In exchange for $100, you now have a piece of paper indicating that you own a share of a firm. After a year has gone by, imagine that the firm declares it will pay out dividends of $6.00 per share. Also, at the end of the year, suppose the price of the stock has increased to $105.00. You decide to sell at that price. So with your $100.00, you received $111.00 at the end of the year for an annual return of 11 percent:

(We have used the term *return* a few times. We will give a more precise definition of this term later. At present, you just need to know that it is the amount you obtain, in percentage terms, from holding an asset for a year.)

Suppose that a firm makes some profits but chooses not to pay out a dividend. What does it do with those funds? They are called *retained earnings* and are normally used to finance business operations. For example, a firm may take some of its profits to build a new factory or buy new machines. If a firm is being managed well, then those expenditures should allow a firm to make higher

profits in the future and thus be able to pay out more dividends at a later date. Presuming once again that the firm is well managed, retained earnings should translate into extra dividends that will be paid in the future.

Furthermore, if people expect that a firm will pay higher dividends in the future, then they should be willing to pay more for shares in that firm today. This increase in demand for a firm's shares will cause the share price to increase. So if a firm earns profits but does not pay a dividend, you should expect to get some capital gain instead. We come back to this idea later in the chapter and explain more carefully the connection between a firm's dividend payments and the price of its stock.

The Riskiness of Stocks

Figure 10.2.1 reminds us that stock prices decrease as well as increase. If you choose to buy a stock, it is always possible its price will fall, in which case you suffer a capital loss rather than obtain a capital gain. The riskiness of stocks comes from the fact that the underlying fortunes of a firm are uncertain. Some firms are successful and earn high profits, which means that they are able to pay out large dividends—either now or in the future. Other firms are unsuccessful through either bad luck or bad management, and do not pay dividends. Particularly unsuccessful firms go bankrupt; shares in such a firm become close to worthless. When you buy a share in a firm, you have the chance to make money, but you might lose money as well.

Bonds

Wall Street is also home to many famous financial institutions, such as Morgan Stanley, Merrill Lynch, and many others. These firms act as the financial intermediaries that link borrowers and lenders. If desired, you could use one of these firms to help you buy and sell shares on the stock exchange. You can also go to one of these firms to buy and sell bonds. A **bond** is a promise to make cash payments (the **coupon**) to a bondholder at predetermined dates (such as every year) until the maturity date. At the **maturity date**, a final payment is made to a bondholder. Firms and governments that are raising funds issue bonds. A firm may wish to buy some new machinery or build a new plant, so it needs to borrow to finance this investment. Or a government might issue bonds to finance the construction of a road or a school.

The easiest way to think of a bond is that it is the asset associated with a loan. Here is a simple example. Suppose you loan a friend $100 for a year at a 6 percent interest rate. This means that the friend has agreed to pay you $106 a year from now. Another way to think of this agreement is that you have *bought*, for a price of $100, an asset that entitles you to $106 in a year's time. More generally (as the definition makes clear), a bond may entitle you to an entire schedule of repayments.

The Riskiness of Bonds

Bonds, like stocks, are risky.

- The coupon payments of a bond are almost always specified in dollar terms. This means that the real value of these payments depends on the **inflation rate** in an economy. Higher inflation means that the value of a bond has less worth in real terms.
- Bonds, like stocks, are also risky because of the possibility of bankruptcy. If a firm borrows money but then goes bankrupt, bondholders may end up not being repaid. The extent of this risk depends on who issues the bond. Government bonds usually carry a low risk of bankruptcy. It is unlikely that a government will default on its debt obligations, although it is not impossible: Iceland, Ireland, Greece, and Portugal, for example, have recently been at risk of default. In the case of bonds issued by firms, the riskiness obviously depends on the firm. An Internet start-up firm operated from your neighbor's garage is more likely to default on its loans than a company like the Microsoft Corporation. There are companies that evaluate the riskiness of firms; the ratings provided by these companies have a tremendous impact on the cost that firms incur when they borrow.

Inflation does not have the same effect on stocks as it does on bonds. If prices increase, then the fixed nominal payments of a bond unambiguously become less valuable. But if prices increase, firms will typically set higher nominal prices for their products, earn higher nominal profits, and pay higher nominal dividends. So inflation does not, in and of itself, make stocks less valuable.

> **Toolkit: Section 31.8**
>
> You can review the meaning and calculation of the inflation rate in the toolkit.

One way to see the differences in the riskiness of bonds is to look at the cost of issuing bonds for different groups of borrowers. Generally, the rate at which the US federal government can borrow is much lower than the rate at which corporations borrow. As the riskiness of corporations increases, so does the return they must offer to compensate investors for this risk.

Real Estate and Cars

As you continue to walk down the street, you are somewhat surprised to see a real estate office and a car dealership on Wall Street. (But this is a fictionalized Wall Street, so why not?) Real estate is another kind of asset. Suppose, for example, that you purchase a home and then rent it out. The rental payments you receive are analogous to the dividends from a stock or the coupon payments on a bond: they are a flow of money you receive from ownership of the asset.

Real estate, like other assets, is risky. The rent you can obtain may increase or decrease, and the price of the home can also change over time. The fact that housing is a significant—and risky—financial asset became apparent in the global financial crisis that began in 2007. There were many aspects of that crisis, but an early trigger of the crisis was the fact that housing prices decreased in the United States and around the world.

If you buy a home and live in it yourself, then you still receive a flow of services from your asset. You don't receive money directly, but you receive money indirectly because you don't have to pay rent to live elsewhere. You can think about measuring the value of the flow of services as rent you are paying to yourself.

Our fictional Wall Street also has a car dealership—not only because all the financial traders need somewhere convenient to buy their BMWs but also because cars, like houses, are an asset. They yield a flow of services, and their value is linked to that service flow.

The Foreign Exchange Market

Further down the street, you see a small store listing a large number of different three-letter symbols: BOB, JPY, CND, EUR, NZD, SEK, RUB, SOS, ADF, and many others. Stepping inside to inquire, you learn that that, in this store, they buy and sell foreign currencies. (These three-letter symbols are the currency codes established by the International Organization for Standardization (http://www.iso.org/iso/home.htm). Most of the time, the first two letters refer to the country, and the third letter is the initial letter of the currency unit. Thus, in international dealings, the US dollar is referenced by the symbol *USD*.)

Foreign currencies are another asset—a simple one to understand. The return on foreign currency depends on how the exchange rate changes over the course of a year. The (nominal) **exchange rate** is the price of one currency in terms of another. For example, if it costs US$2 to purchase €1, then the exchange rate for these two currencies is 2. An exchange rate can be looked at in two directions. If the dollar-price of a euro is 2, then the euro price of a dollar is 0.5: with €0.5, you can buy US$1.

Suppose that the exchange rate this year is US$2 to the euro, and suppose you have US$100. You buy €50 and wait a year. Now suppose that next year the exchange rate is US$2.15 to the euro. With your €50, you can purchase US$107.50 (because US$(50 × 2.15) = US$107.50). Your return on this asset is 7.5 percent. Holding euros was a good investment because the dollar became less valuable relative to the euro. Of course, the dollar might increase in value instead. Holding foreign currency is risky, just like holding all the other assets we have considered.The currency market is also discussed in Chapter 8.

The **foreign exchange market** brings together suppliers and demanders of different currencies in the world. In these markets, one currency is bought using another. The law of demand holds: as the price of a foreign currency increases, the quantity demanded of that currency decreases. Likewise, as the price of a foreign currency increases, the quantity supplied of that currency increases. Exchange rates are determined just like other prices, by the interaction of supply and demand. At the equilibrium exchange rate, the quantity of the currency supplied equals the quantity demanded. Shifts in the supply or demand for a currency lead to changes in the exchange rate.

> **Toolkit: Section 31.20**
>
> You can review the foreign exchange market and the exchange rate in the toolkit.

Foreign Assets

Having recently read about the large returns on the Shanghai stock exchange and having seen that you can buy Chinese currency (the yuan, which has the international code CNY), you might wonder whether you can buy shares on the Shanghai stock exchange. In general, you are not restricted to buying assets in your home country. After all, there are companies and governments around the world who need to finance projects of various forms. Financial markets span the globe, so the bonds issued by these companies and governments can be purchased almost anywhere. You can buy shares in Australian firms, Japanese government bonds, or real estate in Italy.Some countries have restrictions on asset purchases by noncitizens—for example, it is not always possible for foreigners to

buy real estate. But such restrictions notwithstanding, the menu of assets from which you can choose is immense. Indeed, television, newspapers, and the Internet report on the behavior of both US stock markets and those worldwide, such as the FTSE 100 on the London stock exchange, the Hang Seng index on the Hong Kong stock exchange, the Nikkei 225 index on the Tokyo stock exchange, and many others.

You could buy foreign assets from one of the big financial firms that you visited earlier. It will be happy to buy foreign stocks or bonds on your behalf. Of course, if you choose to buy stocks or bonds associated with foreign companies or governments, you face all the risks associated with buying domestic stocks and bonds. The dividends are uncertain, there might be inflation in the foreign country, the price of the asset might change, and so on. In addition, you face exchange rate risk. If you purchase a bond issued in Mexico, you don't know what exchange rate you will face in the future for converting pesos to your home currency.

You may feel hesitant about investing in other countries. You are not alone in this. Economists have detected something they call *home bias*. All else being equal, investors are more likely to buy assets issued by corporations and governments in their own country rather than abroad.

A Casino

Toward the end of your walk, you are particularly surprised to see a casino. Stepping inside, you see a casino floor, such as you might find in Las Vegas, Monaco, or Macau near Hong Kong. You are confronted with a vast array of betting opportunities.

The first one you come across is a roulette wheel. The rules are simple enough. You place your chip on a number. After the wheel is spun, you win if—and only if—you guessed the number that is called. There is no skill—only luck. Nearby are the blackjack tables where a version of 21 is played. In contrast to roulette, blackjack requires some skill. As a gambler in blackjack, you have to make choices about taking cards or not. The objective is to get cards whose sum is as high as possible without going over 21. If you do go over 21, you lose. If the dealer goes over 21 and you don't, you win. If neither of you goes over 21, then the winner is the one with the highest total. There is skill involved in deciding whether or not to take a card. There is also a lot of luck involved through the draw of the cards.

You always thought of stocks and bonds as serious business. Yet, as you watch the players on the casino floor, you come to realize that it might not be so peculiar to see a casino on Wall Street. Perhaps there are some similarities between risking money at a gambling table and investing in stocks, bonds, or other assets. As this chapter progresses, you will see that there are some similarities between trading in financial assets and gambling in a casino. But you will learn that there are important differences as well.

Key Takeaways

- Many different types of assets, such as stocks, bonds, real estate, and foreign currency, are traded in financial markets.
- Your earnings from owning an asset depend on the type of asset. If you own a stock, then you are paid dividends and also receive a capital gain or incur a capital loss from selling the asset. If you own real estate, then you have a flow of rental payments from the property and also receive a capital gain or incur a capital loss from selling the asset.
- Risks also depend on the type of asset. If you own a bond issued by a company, then you bear the risk of that company going bankrupt and being unable to pay off its debt.

exercises

1. If you live in a house rather than rent it, do you still get some benefits from ownership? How would these benefits compare with the income you could receive if you rented out the house?
2. What assets are subject to the risk of bankruptcy?

This page titled 10.2: A Walk Down Wall Street is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

10.3: The Value of an Asset 10.3: The Value of an AssetEdit Update titlecancel

Learning Objectives

1. What factors determine the value of an asset?
2. How do you use discounted present value to calculate the value of an asset?
3. How is risk taken into account when valuing an asset?

Our basic explanation of assets reveals that there are two ways in which you can earn money from holding an asset: (1) You may receive some kind of payment that we call a *flow benefit*—a dividend payment from a stock, a coupon payment from a bond, a rental check from an apartment, and so on. (2) The price of the asset may increase, in which case you get a capital gain. You might guess that the price of an asset should be linked in some way to the payments you get from the asset, and you would be right. In this section, we explain how to determine the price of an asset. To do so, we use two tools: **discounted present value** and **expected value**.These tools are discussed at length in Chapter 5.

Toolkit: Section 31.4 and Section 31.7

You can review the meaning and calculation of discounted present value and expected value in the toolkit.

The Value of an Orange Tree

Imagine that you own a very simple asset: an orange tree. The orange tree pays a "dividend" in the form of fruit that you can sell. What is the value to you of owning such a tree? You can think of this value as representing the most you would be willing to pay for the orange tree—that is, your valuation of the tree. As we proceed, we will link this value to the price of the orange tree.

We begin by supposing your orange tree is *very* simple indeed. Next year, it will yield a crop of precisely one orange. That orange can be sold next year for $1. Then the tree will die. We suppose that you know all these things with certainty.

The value to you of the orange tree today depends on the value of having $1 next year. A dollar next year is not worth the same as a dollar this year. If you have a dollar this year, you can put it in the bank and earn interest on it. The technique of discounted present value tells us that you must divide next year's dollar by the **nominal interest factor** to find its value today:

Here and for the rest of this chapter we use the nominal interest factor rather than the **nominal interest rate** to make the equations easier to read. The interest factor is 1 plus the interest rate, so whenever the interest rate is positive, the interest factor is greater than 1. We use the *nominal* interest factor because the flow benefit we are discounting has not been corrected for inflation. If this flow were already corrected for inflation, then we would instead discount by the **real interest factor**.

Toolkit: Section 31.6

You can review nominal and real interest rates and nominal and real interest factors in the toolkit.

To see why this formula makes sense, begin with the special case of a nominal interest rate that is zero. Then using this formula, the discounted present value of a dollar next year is exactly $1. You would be willing to pay at most $1 today for the right to receive $1 next year. Similarly, if you put $1 in a bank paying zero interest today, you would have exactly $1 in the bank tomorrow. When the nominal interest rate is zero, $1 today and $1 next year are equally valuable. As another example, suppose the nominal interest rate is 10 percent. Using the formula, the discounted present value is = $0.909. If you put $0.909 in a bank account paying a 10 percent annual rate of interest (an interest factor of 1.1), then you would have $1 in the bank at the end of the year.

A Tree That Lives for Many Years

Our orange tree was a very special tree in many ways. Now we make our tree more closely resemble real assets in the economy. Suppose first that the tree lives for several years, yielding its flow benefit of fruit for many years to come. Finding the value of the tree now seems much harder, but there are some tricks that help us determine the answer. Orange trees—like stocks, bonds, and other assets—can be bought and sold. So suppose that next year, you harvest the crop of one orange, sell it, and then also sell the tree. Using this strategy, the value of the tree is as follows:

The first term is the same as before: it is the discounted present value to you of the crop next year ($1.00 in our example). The second term is the price that you can sell the tree for next year. After all, if the tree lives for 10 years, then next year it will still have 9 crops remaining and will still be a valuable asset.

This expression tells us something very important. The value of an asset depends on

- the value of the flow benefit (here, the crop of oranges) that you obtain while owning the asset,
- the price of the asset in the market when you sell it.

The insight that the value of the tree equals the value of the crop plus next year's price greatly simplifies the analysis. If you know the price next year, then you know the value of the tree to you this year. Of course, we do not yet know how the price next year is determined; we come back to that question later.

We can now give a more precise definition of the **return on an asset**: it is the amount you obtain, in percentage terms, from holding the asset for a year. The return has two components: a flow of money (such as a dividend in the case of a stock) and the price of the asset. In the case of the orange tree, the return is calculated as

Because we know that

it follows that

In this simple case, the return on the asset is equal to the nominal interest rate. If we wanted the real return, we would use the real interest factor (1 + the real interest rate) instead.

A Tree with a Random Crop

So far we have assumed that you know the orange crop with certainty. This is a good starting point but is not realistic if we want to use our story to understand the value of actual assets. We do not know for sure the future dividends that will be paid by a company whose stock we might own. Nor do we know the future price of a stock or a bond.

Looking back at the tree that lives for one year only, imagine you do not know how many oranges it will yield. Start by assuming that you can buy a tree that lasts for one period and whose crop is not known with certainty. The value of the tree depends on the following.

- The expected value of the crop. You must list all the possible outcomes and the probability of each outcome. For example, Table 10.3.1 "Expected Crop from an Orange Tree" shows the case of a tree where there are three possible outcomes: 0, 1, or 2 oranges. The probability of 0 oranges is 10 percent—that is, 1 in 10 times on average, the tree yields no fruit. The probability of 1 orange is 50 percent: half the time, on average, the tree yields 1 fruit. And the probability of 2 oranges is 40 percent. The expected crop is obtained by adding together the numbers in the final column: 1.3 oranges.
- A **risk premium** is an addition to the return on an asset that is demanded by investors to compensate for the riskiness of the asset. This adjustment reflects the riskiness of the crop and how **risk-averse** the owner of the tree is. If the owner is **risk-neutral**, there is no need for a risk premium. Obviously enough, if the crop is known with certainty, there is also no need for a risk premium.

Toolkit: Section 31.7

You can review the concepts of risk aversion and risk-neutrality in the toolkit.

Table 10.3.1: Expected Crop from an Orange Tree

Outcome (Number of Oranges)	Probability	Probability × Outcome
0	0.1	0
1	0.5	0.5
2	0.4	0.8

The easiest way to see how the risk premium works is to recognize that someone who is risk-averse will demand a higher return to hold a risky asset. Earlier, we said that the return on an asset without risk equals the nominal interest rate. In the case of a risky asset, however,

From this we can see that there is a relationship between risk and return. If the crop is not risky, then the risk premium is zero, so the return equals the nominal interest rate. As the crop becomes riskier, the risk premium increases, causing an increase in the return per dollar invested.

We can see how the risk premium affects the value of the tree by rearranging the equation:

For a given expected crop, the higher is the risk premium, the lower is the value of the tree.

The Value of an Asset in General

We have been talking about orange trees because they nicely illustrate the key features of more complex assets. We can combine the insights from our analysis of the orange tree to obtain a fundamental equation that we can use to value all kinds of assets:

We apply this equation throughout the remainder of the chapter. To keep things simple, however, we will suppose most of the time that there is no risk premium—that is, we will discount using the nominal interest factor alone, except when we explicitly want to talk about the riskiness of different assets. We can now use this formula to value assets that are more familiar, such as bonds, stocks, cars, and houses.

The Value of a Bond

Suppose that you want to value a bond that lasts only one year. You will receive a payment from the borrower next year and then—because the bond has reached its maturity date—there will be no further payments. Naturally enough, the bond is worthless once it matures, so its price next year will be zero. This bond is like the first orange tree we considered: it delivers a crop next year and then dies. Hence we can value the bond using the formula

For example, if the coupon on the bond called for a payment of $100 next year and the nominal interest rate was zero, then the value of the bond today would be $100. But if the nominal interest rate was 10 percent, then the value of the bond today would be = $90.91.

If the bond has several years until maturity,

This expression for the value of a bond is very powerful. It shows that a bond is more valuable this year if

- the coupon payment next year is higher,
- the bond will sell for a higher price next year, or
- interest rates are lower.

We explained earlier that bonds are subject to inflation risk. There are two ways of seeing this in our example. Imagine that inflation increases by 10 percentage points.

This inflation means that the coupon payment next year will be worth less in real terms—that is, in terms of the amount of goods and services that it will buy. Also, from the **Fisher equation**, we know that increases in the inflation rate translate into changes in the nominal interest rate. If inflation increases by 10 percentage points and the real rate of interest is unchanged, then the nominal rate increases by 10 percentage points. So the discounted present value of the bond decreases. Inflation risk might cause a bondholder to include a risk premium when valuing the bond.

> **Toolkit: Section 31.8**
>
> You can review the Fisher equation in the toolkit.

The Value of a Stock

Now let us use our general equation to evaluate the dividend flow from stock ownership. Imagine you are holding a share of a stock this year. You can hold it for a year, receive the dividend payment if there is one, and then sell the stock. For now we treat both the dividend and the price next year as if they are known for sure. What is the value of a share under that plan?

This equation is similar to the one we used for the fruit tree and the bond. The flow benefit in this case is the dividend paid on the stock. Because the dividend is received next year, we have to discount it back to the current year using the nominal interest factor. The other part of the value of the share comes from the fact that it can be sold next year. Again, that share price must be discounted to put it in today's terms. If the share does not pay a dividend next year, then its value is even simpler: the value of the share this year equals its price next year discounted by the nominal interest factor.

The return to owning the share comes in two forms: the dividend and the gain from selling the share next year. To calculate the return per dollar invested, we divide the dividend and future price by the value of a share this year:

Table 10.3.2 "Discounted Present Value of Dividends in Dollars" shows an example where we calculate the value of a stock using two different interest rates: 5 percent and 10 percent.

Table 10.3.2: Discounted Present Value of Dividends in Dollars

Dividend	Price Next Year	Discounted Present Value (5%)	Discounted Present Value (10%)
1	2	2.86	2.73
1	4	4.76	4.55
2	4	5.71	5.45

The Value of a House

There are other familiar assets that can also be valued in the same way. A house is an asset that delivers a benefit each year in the form of providing shelter. The value of a house is the flow of services that it provides over the coming year plus the price it could be sold for next year. Of course, instead of living in your house and enjoying the service flow, you could rent it out instead. Then

For a house and similar assets, the value today reflects

- the flow of services of the asset over a year,
- the resale value next year, and
- the interest rate that is used to discount the future flows.

This completely parallels what we have already found for both bonds and stocks.

Key Takeaways

- The value of an asset is the most you would pay to own that asset. The value today is the discounted value of the sum of the dividend (or service flow) plus the future price of the asset.
- Because the return of owning an asset comes in the future, you use discounted present value to calculate the current value of the asset. If the dividend and future price are not corrected for inflation, then you discount using the nominal interest rate. If the dividend and future price have already been corrected for inflation, then you discount using the real interest rate.
- The value of an asset is reduced by a risk premium that takes into account the riskiness of the asset and your risk aversion.

Checking your understanding

1. plain why an increase in the price of an asset in the future will increase its value today. Is this a violation of the law of demand?
2. In the section "The Value of a House", we talked about houses. Can you think of other assets that could be valued using a similar formula?
3. Revise Table 10.3.1 "Expected Crop from an Orange Tree" so that the probability of getting 0 oranges is 0 and the probability of getting 3 oranges is 0.1. What is the expected crop from this tree? Is it more or less valuable to you?

This page titled 10.3: The Value of an Asset is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

10.4: Asset Markets and Asset Prices 10.4: Asset Markets and Asset PricesEdit Update titlecancel

Learning Objectives

1. What is arbitrage?
2. How are asset prices determined?

So far we have focused on the value of an asset to an individual: “What is the value to you of the asset (fruit tree, bond, stock, car, house, etc.) you are holding?” Now we want to go a step further and see what the market price is for the asset. We already know that the two are connected. For example, when we valued a bond, we wrote

Part of the value of a bond to you is the price you can sell it for on the market next year. Now we explain that the *current* price of a bond is closely connected to its current value.

Arbitrage

Assets are traded in markets around the world. Typically, there are a large number of (potential) buyers and sellers for any given asset: thousands of people might be willing to buy Microsoft Corporation stock or sell government bonds if they felt the price was right. Also, assets are homogeneous: one US government 10-year bond is the same as another. This means that asset markets are a good example of **competitive markets**, which means that we can look at asset markets using **supply and demand**.

Toolkit: Section 31.9

You can review supply and demand and competitive markets in the toolkit.

To derive the supply and demand curves for assets, we use the idea of **arbitrage**. This is the act of buying and then reselling an asset to take advantage of profit opportunities. The idea of arbitrage is to “buy low” and “sell high.” Arbitrage is usually carried out across two markets to profit from any difference in prices. The strict definition of arbitrage refers to buying and selling where there is no risk, meaning that profits can be made with certainty. A weaker meaning of arbitrage allows risk to be associated with the process.

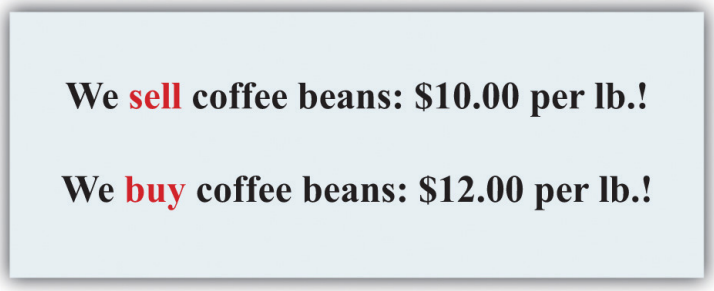


Figure 10.4.1: Arbitrage at a Coffee Shop

Imagine you passed a coffee shop and saw the sign shown in Figure 10.4.1. This would make an economist salivate, not because of the prospect of good coffee but because it presents an opportunity for arbitrage. Facing an offer like this, you could immediately go and buy a pound of coffee beans for $10. Then you could turn around and sell the coffee at $12 per pound. You would have made $2 easy profit. Forget about drinking the coffee: just buy and sell, buy and sell, pound after pound—and become a billionaire. This is an example of arbitrage.

Sadly, you will never see a coffee shop making you an offer like this. We are confident of this because any coffee shop that made such an offer would very quickly go out of business. After all, if you can make a profit by buying at a low price and selling at a high price, then whoever is on the other side of these transactions is making a loss.

Arbitrage in the Supply-and-Demand Framework

We can think about arbitrage using the supply-and-demand framework. There are two markets: in one the coffee shop sells coffee, while in the other the coffee shop buys coffee. The demand of potential buyers would be extremely large, and the supply of coffee (from people selling it back) would likewise be very large. With the prices for buying and selling coffee as stated in the sign, demand could never equal supply in these two markets. An arbitrage possibility like this is not consistent with market equilibrium.

Using similar logic, we can argue that *the price of an asset will equal its value*. To see why, we begin again with an orange tree that will yield an orange worth $1 next year. Owners of this asset value it at

They will be willing to sell the asset at this price but not if the price is any lower. They would definitely want to sell if the price were higher. But buyers can perform exactly the same calculation. They would be willing to buy the asset at this price but not if the price were any higher. They would definitely want to buy if the price were lower. Figure 10.4.2 shows the supply of and demand for trees in this case.

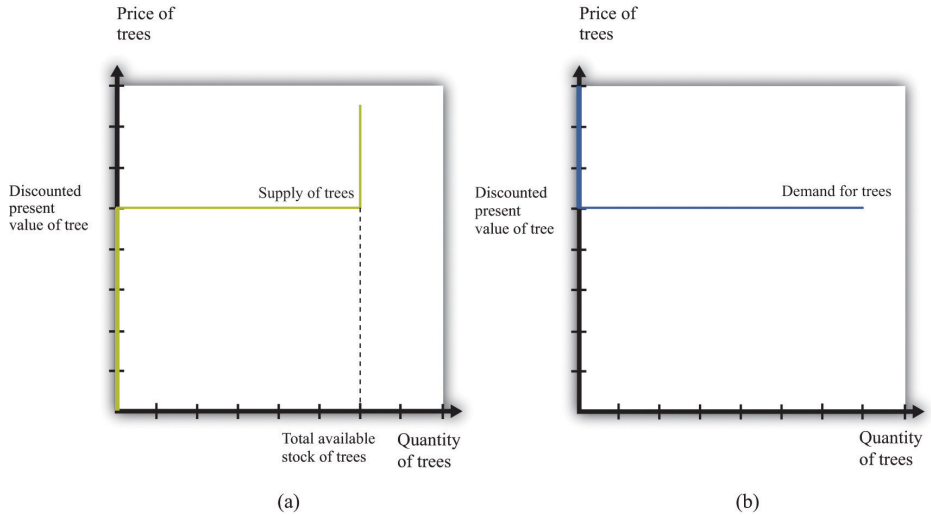


Figure 10.4.2: Asset Demand and Supply. (a) Owners of trees are willing to sell at a price equal to the discounted present value of the tree, and the supply curve is flat (perfectly elastic) up to the total available stock of trees. (b) Potential buyers of trees are willing to buy at a price equal to the discounted present value of the tree, and the demand curve is flat (perfectly elastic).

The supply side is shown in part (a) of Figure 10.4.2. There is a given stock of trees available. For prices below the asset value, no one wants to sell the asset. At prices above the value, everyone wants to sell the asset. So the supply curve is horizontal at a price equal to the asset value, all the way up to the point where every tree is on the market. At that point, the supply curve becomes vertical. The demand function is in part (b) of Figure 10.4.2. At a price above the discounted present value of the tree, the quantity demanded is zero: no one will pay more than the discounted present value for the asset. If the price equals the value, the demand is flat (horizontal). At a price below the value, the asset looks like a great deal because there are arbitrage opportunities. So demand is very large.

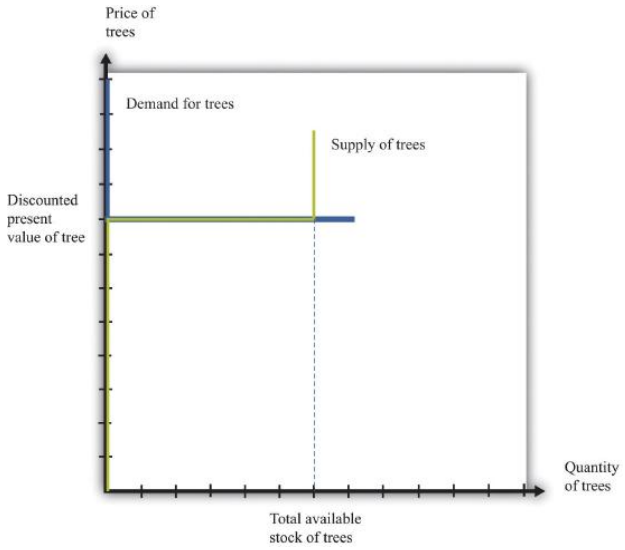


Figure 10.4.3: Asset Market Equilibrium. Because buyers and sellers place the same value on the tree, the demand and supply curves lie on top of each other at this value, so the price will equal the discounted present value of the tree.

We put these curves together in Figure 10.4.3. Both supply and demand are horizontal at a price equal to the discounted present value of the asset. Thus at this price, and at this price only, supply equals demand. We obtain a powerful prediction: *assets will be priced at their discounted present value*. If we see the prices of assets (such as stocks or bonds) increase or decrease, this model of the asset market tells us to attribute these variations to changes in the discounted present value of dividends.

The supply and demand curves in these figures look rather untraditional. We are used to seeing upward sloping supply curves and downward sloping demand curves. But in the market for the tree, everyone values the asset in exactly the same way. That valuation is given by the discounted present value of the dividend stream. As a result, Figure 10.4.3 does not tell us how much people will trade or if they will trade at all. When the price equals the discounted present value, buyers are indifferent about buying, and sellers are indifferent about selling. Everyone is happy to trade, but no one particularly wants to trade. In reality, though, the market for an asset may look much more like a “normal” supply-and-demand diagram, as in Figure 10.4.4, with an upward-sloping supply curve and downward-sloping demand curve. The reason is that different individuals may have differing views about the discounted present values of the asset, because Chapter 6 "eBay and craigslist" discusses in some detail the reasons why people trade. We explain that important motives for trade are that people have different tastes and skills. To these we can add the two motives just mentioned here.

- different buyers and sellers have different information that causes them to make different forecasts of future dividends, or
- different buyers and sellers differ in their attitudes toward risk.

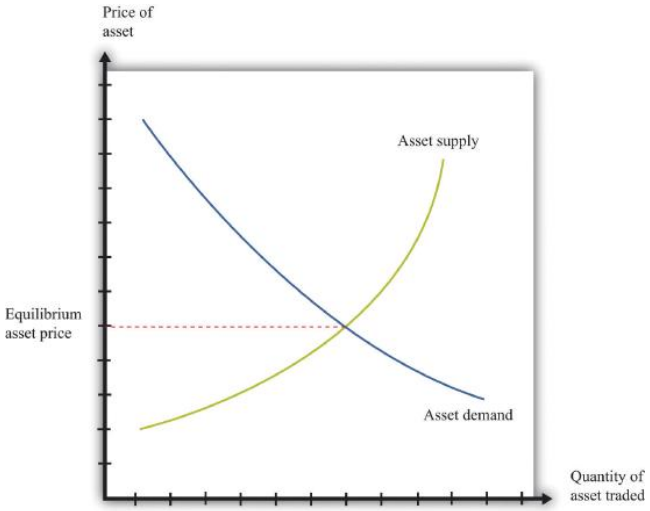


Figure 10.4.4: Asset Market Equilibrium: A More Familiar View. If potential buyers and sellers of an asset differ in their beliefs about the dividend from that asset or differ in terms of their degree of risk aversion, then we obtain demand and supply curves of the familiar form.

For example, suppose some buyers are optimistic about future dividends from a stock, while others are pessimistic. Optimistic buyers will calculate a high discounted present value and have a high willingness to pay. Pessimistic buyers will calculate a lower discounted present value and be willing to pay less for the asset. Such differences in views can hold for sellers as well. Alternatively, suppose some buyers and sellers are more risk-averse than others. The less risk-averse the buyer, the higher the price he is willing to pay because he uses a lower risk premium when calculating his discounted present value. The less risk-averse the seller, the higher the price she is willing to accept.

There is one last, more subtle point. We have been imprecise—intentionally—about what exactly it means for an asset to be risky. Buyers and sellers care not only about assets in isolation but also about how those assets fit into their entire portfolio—that is, the entire collection of assets that they own. An asset that seems very risky to one person may appear less risky to another because he holds other assets that balance out the risks. The riskiness of an individual asset depends on the **diversification** of the portfolio as a whole.

In Figure 10.4.3, all traders in the market valued the asset in exactly the same way, so arbitrage guaranteed that the price equals the discounted value of the flow benefit. In Figure 10.4.4, there is no immediate guarantee that this will still be true. Even with differences in valuation, however, we expect that the price of an asset is still likely to be (at least approximately) equal to its true discounted present value. In particular, if some traders in the market do not care about risk and are accurately informed about the flow benefit, arbitrage will still keep the market price close to the discounted present value of the stock.

We have not yet explained how supply and demand actually come together in financial markets—that is, who actually makes the market? If you study the workings of a market such as the New York Stock Exchange, you will learn that it works through specialized traders. If you want to buy a stock, you typically contact a stockbroker who communicates your demand to his firm on Wall Street. Another broker then takes that order onto the floor of the stock exchange and looks for a seller. If a seller is found, then a deal can be made. Otherwise, the broker can place your order with another specialist who essentially “makes the market” by buying and selling securities at posted prices. So in the end, the market has some elements of posted prices (take-it-or-leave-it offers) and some elements of a double-oral auction.Both of these are discussed in more detail in Chapter 6.

The Price and Value of a Long-Lived Asset

Previously, we explained how to value an asset assuming you hold it for one year, receive the flow benefit (the fruit, the dividend, or the coupon payment), and then sell it at the current market price. We said that (assuming no risk premium)

We have also discovered that, in general,

$$value\ of\ asset\ =\ price\ of\ asset.$$

We combine those two pieces of information to complete our study of the valuation of assets. Imagine an orange tree that lives for two years and yields a crop valued at $1 each year. We already know that we can value the tree as follows:

But now we know that the *price* of the tree next year will equal the *value* of the tree next year:

Next year, we can apply exactly the same formula:

Why is this true? The year after next, this particular tree will be worthless because it will be dead. So the value of the tree today is

This is a more complicated formula. It tells us that the value of the tree today is the discounted present value of the flow benefit tomorrow plus the discounted present value of the value of the tree tomorrow—*which is itself the discounted present value of the flow benefit the year after*. In other words, the value of the tree today is the discounted present value of the flow benefits over the entire lifetime of the tree. What is more, we could use exactly the same logic if the tree were to yield a crop for 3 years, 10 years, or 100 years.

There is one last step. If we again use the idea that the price of the tree should equal its value, then we can conclude the following:

$$price\ of\ tree\ this\ year\ =\ discounted\ present\ value\ of\ the\ flow\ of\ benefits\ from\ the\ tree.$$

This logic applies to all assets, not only trees, so we can now apply it to bonds and stocks:

$$price\ of\ bond\ today\ =\ discounted\ present\ value\ of\ the\ flow\ of\ payments\ from\ the\ bond$$

and

$$price\ of\ share\ today\ =\ discounted\ present\ value\ of\ the\ flow\ of\ dividend\ payments.$$

A final note on uncertainty. We have been assuming that dividends are known with certainty. If they are not, then we need to modify the valuation of the stock by (1) replacing "dividend" with "expected dividend" and (2) adding a risk premium to the interest rate. As discussed previously, the adjustment for risk will reflect both the riskiness of the stock and the aversion to risk of investors. Riskier stocks generally have a lower value and a higher expected return.

Key Takeaways

- Arbitrage entails the buying and selling of assets to make a profit. In equilibrium, there are no profits to be made through arbitrage.
- The price of an asset is (approximately) equal to the discounted present value of the flow of benefits (dividends, service flow, etc.) from the asset.

Exercises

1. Suppose an orange tree lives for two years, with a crop of five oranges in the first year and three in the second year. The price is $2 in the first year and $5 in the second year. If the interest rate is 20 percent, what is the price of the orange tree?
2. Explain why an increase in interest rates will reduce the price of a bond.

This page titled 10.4: Asset Markets and Asset Prices is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

10.5: Efficient Markets 10.5: Efficient MarketsEdit Update titlecancel

Learning Objectives

1. Is it possible to make large profits in asset markets?
2. Is it easy to make large profits in asset markets?
3. What are some factors that cause asset prices to increase and decrease?
4. Why do asset prices respond to new events but not forecasted ones?
5. Are markets efficient?

The title of this chapter speaks of making and losing money on Wall Street. We have gone into considerable detail about what determines the price of assets, but we have not yet discussed how easy or hard it is to make money by buying and selling these assets.

Can Easy Profits Be Made on Wall Street?

Our fictional Wall Street contained places where you could buy many different kinds of assets, such as real estate and automobiles as well as stocks and bonds. But it also contained a building that wasn't selling assets at all: the casino.

Is Wall Street Like a Roulette Wheel?

Is buying and selling shares like gambling on a roulette wheel, where gains and losses are purely a matter of luck? To answer this question, think more about the uncertainty associated with buying stocks and bonds. Suppose we are buying a stock that will pay dividends over four years, as in Table 10.5.1 "Discounted Present Value of Dividends in Dollars", and suppose that the interest rate is 5 percent. From Table 10.5.1 "Discounted Present Value of Dividends in Dollars", we know that the discounted present value of the stock is $609.61. We then expect this will also be the price of the stock.

Table 10.5.1: Discounted Present Value of Dividends in Dollars

Year	Dividend ($)	Discounted Present Value ($, Interest Rate = 5%)
1	100	95.2381
2	90	81.63265
3	120	103.6605
4	400	329.081
Discounted present value (all years)		609.61

Can you make money buying and selling this stock? It seems unlikely. If the price of a stock is equal to the present discounted value of the flow of dividends, then you get what you pay for. If you sell the stock, then instead of an asset that would have paid you the equivalent of $609.61, you receive $609.61 in cash. If you buy the stock, the reverse is true. Either way, you are no richer or poorer after the transaction; you are just holding your wealth in a different form.

Economists often use the metaphor of $100 bills lying on the ground to describe a situation where easy money can be made. Our example of the arbitrage opportunity in the coffee shop, where you could buy beans at $10 a pound and resell them at $12 a pound, is an example: getting rich in that coffee shop is as easy as picking up money on the floor. But if the value of a stock is the discounted present value of the dividends that it will pay, then there is no easy money to be made. There are no $100 bills lying around. You should not anticipate spectacular earnings from owning assets. You can earn a reasonable rate of return, equal to the nominal interest rate, but no more. A market with the characteristic that you cannot expect to earn abnormal profits is called an **efficient market**. In such a market, the price of an asset accurately reflects the best forecast of the value of that asset. The value of an asset is the discounted present value of the flow benefits.

Yet this may strike you as odd. There are many people who certainly make it rich by buying and selling stocks, bonds, and other assets. Some of these individuals—like Warren Buffett or George Soros—are household names. Does the fact that some people get very rich on Wall Street mean that markets are not efficient?

There are a couple of possible answers to this question. The first is that even when markets are efficient, some people may get rich. Go back to the stock we were considering earlier but with one small change. Imagine that the numbers in Table 10.5.1 "Discounted Present Value of Dividends in Dollars" are now *expected* dividends. They tell you what this stock can be expected to pay out on average, but they are not guaranteed. For example, it might be the case that in year four, the firm will pay a dividend of $800 with 50 percent probability but will also pay $0 with 50 percent probability. The expected dividend is $400. If people care only about the expected value of the dividends (if they are risk-neutral), the price of the stock still equals the (expected) discounted present value.

Now, there are still no $100 bills on the ground. You cannot expect to make unusual profits by buying this stock or others like it. However, some people will get lucky and thus get rich. If you bought this stock, and it ended up paying the high dividend, you would end up with a return nicely above the market interest rate. If it failed to pay the dividend, however, you would get a lower return than the market interest rate. As we looked around the economy, we would see some lucky investors earning high returns and other unlucky investors earning low returns. In this world, buying and selling assets in the stock market is really not that different from betting on a roulette wheel. Buying an asset is like placing your chip on a certain number. If the number comes up, you get rich. If it does not, you lose.

When you go to a casino, you should not *expect* to win at roulette. But this does not mean that you can never earn spectacular amounts of money. It happens frequently: casinos thrive on advertising these success stories. The same goes for buying an asset, such as a stock. Suppose you buy a share in a pharmaceutical company. The price of the share when you purchase it might indeed equal the expected discounted present value of dividends. Yet the following week the company could have a major discovery that will allow it to be much more profitable in the future. Expected dividends will increase, and so will the stock price. This is certainly good news for you. But it is also no different from getting lucky on the roulette wheel.

Spectacular successes tend to be more visible than losses. In the 1990s, some people earned large amounts (sometimes spectacularly large amounts) from certain successful Internet companies. But many other people lost money on Internet companies that were ultimately unsuccessful, and you are less likely to hear about them.

Is Wall Street Like a Blackjack Table?

More than luck is required when investing on Wall Street, however. Just as there are skilled players of blackjack in the casino, there are people who are skilled in assessing the prospects of different firms in an economy.

If the price of a share equals the discounted present value of the dividends that a company will pay, then the total value of all the shares in a company should equal the discounted present value of the total profits the firm will pay out—both now and in the future. The total value of all the outstanding shares in a firm is called its **market capitalization**.

The price of a share increases whenever a firm's market capitalization increases, and a firm's market capitalization should increase whenever there is reason to think that the firm has become more profitable—either now or in the future. If markets are efficient, therefore, we expect share prices to respond to *new information* about a firm. Traders in large financial firms make their money in part by gathering new information about firms and then acting very quickly on that new information. News about a firm—good or bad—is likely to be incorporated quickly into a firm's market price, but a trader who can move fast can make money from these movements.

When economists use the metaphor of $100 bills lying on the ground, they are pointing out that if opportunities for easy profits arise, they will disappear very quickly. It is not impossible that someone will drop a $100 bill. But it is highly unlikely that that bill will lie unretrieved for more than a few minutes. If there is easy money to make in the market by buying a stock, professional traders will jump on the opportunity quickly. This has an important implication for the rest of us. If you read in the newspaper today that Merck Pharmaceuticals has just announced a new pharmaceutical compound that is highly effective in treating lung cancer, there is no point in calling your broker and instructing him to buy Merck stock on the basis of this news. Somebody—a very smart trader with her ear to the ground and lots of knowledge about the pharmaceutical industry—might well have made money buying Merck stock at the first hint of this news. But by the time there is an announcement in the paper, the increase in Merck's expected profits has long been factored into the price.

Are Markets Efficient?

The theory of efficient financial markets is very powerful because it gives us a key to understanding the prices of assets. Go back to our equation for the value of an asset:

In the case of a share, the flow benefit is the value of the dividend. If the price next year is the discounted present value of further dividends from that point in time onward, then this equation is another way of saying that the value of the share today equals the discounted present value of dividends.

But what if the price that everyone expects an asset to sell for next year is—for some reason—much higher than the discounted present value of the flow benefit from the asset starting next year? As an example, consider a house. You buy a house in part to enjoy living in it—to enjoy "housing services." You also buy a house as a possible source of capital gains if the price of the house increases. Imagine you live somewhere where everyone seems to think housing prices will increase a lot over the next few years. Then you should be willing to pay more for a house. After all, if you anticipate a large capital gain in five years from selling the house, you can pay more for it now and still expect a good return. So if everyone expects the price of houses to increase in the future, then the current demand for houses increases, so the current price increases. The price increase today reflects the expectation of higher prices in the future.

Now fast forward to next year. We can say the same thing applied to next year: "If everyone expects the price of houses to increase in the future, then the current demand for houses increases, so the current price increases." This can go on from year to year: housing prices are high because everyone expects higher prices in the future. Higher prices have an element of self-fulfilling prophecy: prices increase because everyone thinks prices will continually increase.

This is sometimes called a housing *bubble*. In a bubble, the increase in the price of housing does *not* reflect an increase in the value people place in housing services. In the language of economics, the price is not changing because of any change in the *fundamentals*. Furthermore, it is possible that prices will not actually keep increasing. If everyone suddenly becomes more pessimistic about the future of housing prices, then the capital gains that everyone anticipated are gone, and housing prices collapse. In this case, the bubble bursts, and prices fall—sometimes very rapidly.

Many economists think that something like this happened in the early stages of the global financial crisis that started in around 2007. The price of housing in many markets had been increasing substantially, and people expected this to continue. At some point, people stopped being so confident that house prices would keep increasing—and, sure enough, the price of houses then decreased rapidly.

The same idea applies to stocks. If everyone believes the value of a particular stock will be higher in the future, then the price will be bid up today. If these beliefs persist, they can sustain a bubble in the stock. If everyone believes that stocks will generally increase in price, then this can lead to a bubble in the entire stock market.

Data on the prices of stocks can perhaps help us see if the efficient market view is accurate, or if we instead see lots of bubbles. This sounds like an easy exercise but is actually very hard. It is difficult to calculate the discounted present value of expected dividends because it requires forecasts of the future. The usual interpretation of this evidence is that the efficient market hypothesis is not capable of explaining all the variations in asset prices, particularly over short periods of time.

An extreme example illustrates this point well. Here is a quotation from an article that appeared in the *Wall Street Journal* after the collapse in the US stock market in 1987: "Calmly appraised, the intrinsic value of American industry didn't fall 23 percent in a day."Roger Lowenstein, "After the Fall: Some Lessons Are Not So Obvious," *Wall Street Journal*, August 25, 1997. "The intrinsic value of American industry" refers to the total market capitalization of all the firms quoted on the stock exchange. It is hard to explain such short-run variations in asset prices from the perspective of discounted present value of dividends.

Economist Robert Shiller claims that stock markets can exhibit "irrational exuberance." He argues that asset prices move around too much to be explained by theories that rely on the discounted present value of dividends to the price of assets. Instead, the fluctuation of prices, at least over short periods of time, might also be influenced by expectations and bubbles. More generally, Shiller is one of many financial economists who believe that economic theory needs to be supplemented with some ideas from social psychology to do a better job of explaining the performance of financial assets. Such **behavioral finance** has identified several *anomalies*—that is, occasions on which asset prices apparently diverge from the values predicted by efficient market theory.

Despite the insights of behavioral finance, most economists take the view that, at the very least, efficient market theory is the best starting point for thinking about asset prices. Efficient market theory provides us with two key insights: (1) the price of a stock should reflect expectations about future profits, which means that (2) the price of a stock should change when—and only when—there is new information that changes those expectations. Many economists nonetheless think this approach is incomplete and that behavioral finance can also help us understand financial markets. A word of caution: even if markets are not always efficient, this still does not mean that there are easy ways to make money on Wall Street.

Changes in Asset Quantities and Asset Prices

Each weekday in the United States, around 5 p.m. Eastern Standard Time, there are reports on the performance of the markets that day. At other times of the day, you can learn about other markets around the world. Newscasters report on the volume of trade (the number of shares exchanged) and some index of the price of stocks, such as the DJIA. In economic terms, these are reports about the price and quantity in a market. Therefore we can use the supply-and-demand framework, and more specifically the tool of **comparative statics**, to consider what makes asset prices increase and decrease.

> Toolkit: Section 31.16
>
> You can review how to carry out a comparative statics exercise in the toolkit.

If we take the efficient-market view that the price of an asset equals its discounted present value, then any change in the price of an asset must be due to some change in its expected discounted present value. If, for example, the price of General Motors stock increases, this should reflect some information about the prospects and hence future dividends of this company. Moreover, any information that makes the price increase or decrease must be *new*. If it were old information that everyone in the market already knew, then it would have already been factored into the stock price. Hundred-dollar bills do not stay on the ground for long. So what are some of the big events that can change asset prices?

Product Development News

Part of the profitability of a firm comes from its innovative activities in developing and marketing new products. Open a computer magazine, for example, and you will see hundreds of advertisements for a wide range of new products. How does news about a new product affect the price of a firm's stock?

Consider the following story:

Pfizer stock tumbled Monday after the world's biggest drugmaker abruptly pulled the plug on its most important experimental medicine—a drug meant to treat heart disease that instead caused an increase in deaths and heart problems in people taking it in a clinical trial.

Shares of Pfizer sank about 11 percent in afternoon trading as investors worried what the New York-based company would do to replace the product, torcetrapib, in its pipeline.

Pfizer CEO Jeffrey Kindler unveiled his company's pipeline at an analyst meeting last week, before the bad news on torcetrapib.

Trading was heavy with more than 235 million shares changing hands by mid-afternoon—nearly seven times the stock's average daily volume.

On Saturday, Pfizer and the Food and Drug Administration announced that the drug company would halt a clinical trial of torcetrapib due to an increased rate of death and heart problems in patients who took it.

Just two days earlier, Pfizer's new CEO Jeffrey Kindler had told hundreds of investors and analysts at a research meeting that the drugmaker could seek approval for the medicine as early as next year if clinical data supported it.Aaron Smith, "Heart Drug Pulled, Pfizer Tumbles," *CNNMoney.com*, December 4, 2006, accessed January 29, 2011, http://money.cnn.com/2006/12/04/news/companies/pfizer_stock/index.htm.

Pfizer's announcement is the **exogenous variable** in this comparative statics exercise. The story tells us that the announcement led to a decrease in price and a large amount of trading. Let us try to make sense of this. First the announcement clearly contained new information that the markets had not anticipated. Indeed, the article tells us that, a few days previously, there had been positive information about this product.

Because market participants did not previously know the results of the clinical trials, and because the announcement is bad news, analysts and market professionals immediately revise downward their estimates of the future profitability of Pfizer. They now expect that dividends in future years will be lower than they had previously thought. This reduction in the discounted present value of dividends causes the stock price to decrease.

The way this works in the supply-and-demand framework is shown in Figure 10.5.1. The bad news influences both the demand and supply curves. Prospective buyers of the stock are not willing to pay as much for it, given the bad news. So at a given stock price,

the quantity demanded decreases. In Figure 10.5.1, the demand curve shifts to the left. Owners of the stock are now less interested in holding their shares. So at a given price, the quantity supplied increases. This means that the supply curve shifts to the right.

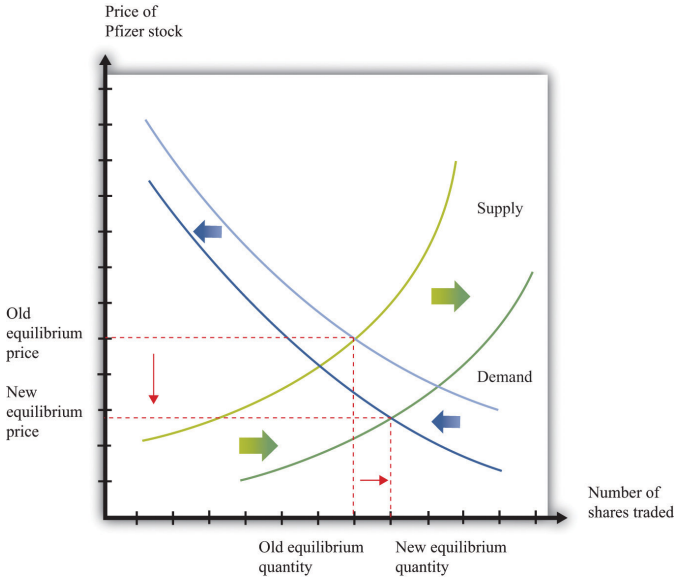


Figure 10.5.1: Bad News about a Firm's Product Reduces the Value of Its Stock.. Bad news about a clinical trial causes many current holders of Pfizer stock to want to sell their shares and makes people less likely to want to buy Pfizer stock.

The bad news has an unambiguous effect on the price of the stock: it decreases. The effect of the announcement on the quantity traded is not clear. It depends on the steepness of the curves and the relative shifts in supply and demand. In this particular example, we know that there was an unusually large amount of trading, so the equilibrium quantity increased.

Had the results of the clinical trials already been leaked to the market, there would have been no new information from the announcement. The test results would already have been incorporated in the asset price. The supply and demand curves would not have moved at all.

Monetary Policy

In most modern economies, interest rates are set (or more precisely heavily influenced) by a central bank through its conduct of monetary policy. This process and its effects on an economy are studied in detail in macroeconomics courses. Here, we look at the effect of monetary policy on asset prices.

Because monetary policy influences interest rates, the link to asset prices is immediate. We know that the price of an asset equals its discounted present value, and that interest rates are used for the discounting process. A change in interest rates therefore directly affects asset prices. Using the example in Table 10.5.1 "Discounted Present Value of Dividends in Dollars", if a monetary authority were to increase the interest rate from 5 percent to 10 percent, then the asset price would decrease by about $81.

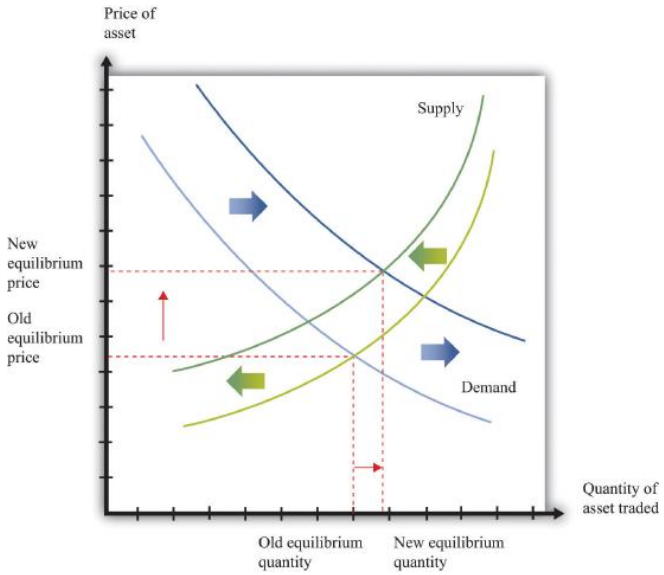


Figure 10.5.2: Equilibrium Asset Prices Respond to a Decrease in Interest Rates. A reduction in interest rates shifts the demand curve for assets (stocks, bonds, and other assets) to the right and shifts the supply curve to the left.

Figure 10.5.2 shows the impact of monetary policy on the supply-and-demand framework. As the interest rate decreases, the discounted present value of an asset increases. So at a given price, the demand for an asset increases. This is shown as an outward shift in demand. Also, at a given price, the supply of the asset decreases. As the interest rate decreases, more holders of the asset want to hold onto it, so the quantity supplied to the market is lower at each price. A rightward shift of the demand curve, combined with a leftward shift in the supply curve, tell us that the price of an asset increases when the interest rate decreases. The effect of the interest rate change on the quantity traded is ambiguous: it depends on the relative slopes of the curves and how much the curves shift.

Multiple Asset Markets

On the walk down Wall Street, we noticed that there were many markets, all trading simultaneously. Yet we have looked at the market for each particular asset in isolation. This is a fine tactic for understanding how to do comparative statics. But the real world is more complex: a single event can impact multiple markets.

Take, for example, the bad news on the test results for Pfizer. We saw that this news forced the firm's stock price to decrease. But Pfizer may have also borrowed in years past by issuing bonds. What will happen to them when the bad news hits the bond market? The price of a bond is the discounted present value of the interest payments on a bond over its lifetime, taking into account the possibility of bankruptcy. So the bad news from the pharmaceutical company ought to depress the price of its bonds. This happens largely because the chance of bankruptcy—while surely small—increases with the bad news.

We observed earlier that the global financial crisis had its origins, at least in part, in the real estate market. The effects quickly spread beyond the housing market. Many people had borrowed to buy houses, and these loans—known as mortgages—were financial assets held by banks. When housing prices decreased, the value of these mortgages decreased as well. Moreover, these mortgages were often combined in various ways to make new assets. When housing prices decreased, the price of these related assets decreased as well. And because financial institutions sometimes had difficulty working out the value of their assets, there was a risk that they would go bankrupt. This in turn meant that anyone who had lent to such an institution now had an asset that was less valuable. Thus a collapse in the price of one asset—houses—led to a decrease in the price of all sorts of other assets in the economy as well.

Key Takeaways

- If markets are efficient, then, on average, there are no excessive profits to be made in asset markets. Some people will be lucky and do better than average, while others will be unlucky and do worse than average.
- Efficient markets provide a benchmark for asset valuation, though asset prices may sometimes deviate from these values.

- We use comparative statics to study the effects of changes in asset supply and demand on prices. Shifts in asset demand may come from new information about a new product or a new technique established by a firm. Monetary policy may also influence the demand for an asset.
- Asset markets respond to events, like the surprise announcement of a new product. Asset markets do not respond to changes today that were announced in the past, such as a change in interest rates by a central bank that was announced (or forecasted) weeks earlier.

check your understanding

1. Explain how the expected return from playing slot machines is negative even though lots of people have great stories about winning money playing slot machines. Can you tell a similar story about stock markets?
2. If you see housing prices in a city decrease by 50 percent over six months, how would you explain this using the efficient markets viewpoint?
3. Suppose there is good news about a company's future product. What happens to the value of its stock?
4. Using the supply-and-demand framework, show how an increase in interest rates can increase the quantity of an asset sold. Now show how a decrease in interest rates can increase the quantity of an asset sold. What are the key differences between the two figures you just created?

This page titled 10.5: Efficient Markets is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

10.6: End-of-Chapter Material 10.6: End-of-Chapter MaterialEdit Update titlecancel

In Conclusion

The performance of the stock market is one of the most closely watched of all economic statistics. This chapter provided some clues as to why people care so much about the value of stocks and other assets.

One reason is that people save by purchasing stocks and other assets. Thus savers want to know what determines the value of assets in the economy. Having read this chapter, you should now understand that the value of any asset is closely linked to the flow of benefits that the asset provides. Indeed, if markets are efficient, then the value of any asset should equal the discounted present value of the flow of benefits.

There are two other reasons why we pay so much attention to the stock market. (1) If the value of a stock reflects the discounted present value of expected dividends, then the market capitalization of a firm represents the best guess as to the value of that firm—which depends ultimately on the profits that it will generate in the future. In that case, a stock market index represents our best guess of the overall value of all firms. It truly is a measure of an economy as a whole. (2) The stock market plays a key role in allocating an economy's saving to those firms that can make the most profitable use of those funds.

Key Links

- NYSE: www.nyse.com/about/education/1095056911801.html
- S&P 500: http://www.standardandpoors.com/home/en/us
- Robert Shiller
  - Home page: http://www.econ.yale.edu/~shiller
  - Irrational exuberance: http://www.irrationalexuberance.com

exercises

1. List the factors you think would make stock prices increase and decrease in the Shanghai stock exchange.
2. An October 2007 article in the *Economist* magazine discusses land prices and office rents. According to the article, rents have recently increased, and land prices have been increasing in the past few years as well. Why would land prices and rents move together?
3. Following from Question 2, what do you think has happened to the price of office buildings as rental rates have increased?
4. Suppose an orange tree yields a crop of one orange after the first year and then two oranges in the second year. As before, let the price of an orange be $1 in both years. What is the value to you of buying the tree today and then selling it next year, after you have harvested the first orange? (Hint: first find the value of the tree tomorrow and then use that as the price for selling the tree.)
5. Suppose an orange tree lives for two years, with a crop of five oranges in the first year and three in the second year. The price is $2 in the first year and $5 in the second year. If the interest rate is 20 percent, what is the price of the orange tree?
6. (Advanced) The table titled Discounted Present Value Exercise provides information about the crop from an orange tree as well as the interest rate for a tree that lives four years. Assume that the price of oranges is $1 in the first year and then increases at 10 percent per year. What is the discounted present value of this tree?
7. Suppose prospective buyers of houses become very optimistic about the future prices of houses. Existing owners, on the other hand, become very pessimistic about the future value of houses. What happens to the price of houses today?
8. Suppose housing markets are efficient. If you see rapidly increasing prices in a market, do you think that rental rates are increasing as well?
9. Explain how contractionary monetary policy can reduce housing prices.
10. (Advanced) In the first row of Table 10.3.2 "Discounted Present Value of Dividends in Dollars", we considered a stock that pays a dividend of $1 this year and that will have a price of $2 next year. Suppose the inflation rate from this year to next year is 5 percent. There are two ways that you can correct for this inflation.
    1. You can leave next year's price in nominal terms and deflate by the nominal interest rate, as we did in the table.
    2. You can adjust next year's price and put it in terms of today's dollars, so next year's price is a "real price." Then you can discount using the real interest rate, which you can get from the Fisher equation.

Show that you get the same answer for the discounted present value using the second method as using the first method. (Note: when the interest rate is 10 percent, you should get exactly the same answer; when the interest rate is 5 percent, there will be a very small difference because the Fisher equation is an approximation.)

11. Can you think of an exogenous event that would cause the demand curve but not the supply curve for an asset to shift?
12. (Advanced) Explain why the DJIA and other stock market indices are more useful after they have been adjusted for inflation.

Economics Detective

1. Find data from a stock exchange in another country. Create a version of Figure 10.6.1 for that stock exchange.
2. What is the current annual return on US government bonds? What is the current annual return on government bonds issued by Argentina? How would you explain the differences in returns?
3. Find recent data on the yields on the debt of Ireland, Spain, and Portugal. What happened to these yields, relative to the yield on German debt, in both October 2010 and November 2010? How might you explain the patterns you find? (Hint: think about our discussion of the riskiness of bonds.)
4. The chapter opened with a discussion of the stock market in Shanghai. Suppose you wanted to buy shares of a company trading on that exchange. How would you go about doing that?
5. Look at data on housing prices in your area. Do they fluctuate as much as stock prices?

Spreadsheet Exercise

1. Suppose an orange tree lives for three years, with a crop of 5 oranges in the first year, 3 in the second year, and 10 in the third year. The price is $2 in the first year and $5 in the second and third years. If the interest rate is 20 percent the first year and then 10 percent the next two years, what is the price of the orange tree?

Year	Number of Oranges	Price of Orange	Revenue	Interest Rate
1	5	1.00		0.05
2	6			0.10
3	4			0.075
4	10			0.20

TABLE 10.6.1: DISCOUNTED PRESENT VALUE EXERCISE

This page titled 10.6: End-of-Chapter Material is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

10.7: Appendix- A General Formulation of Discounted Present Value 10.7: Appendix- A General Formulation of Discounted Present ValueEdit Update titlecancel

This section presents a more general way of thinking about discounted present value. The economic idea is the same as the one we encountered when discussing the pricing of orange trees. Here the idea is to isolate the central ideas of discounted present value. We then use this more general formulation to talk about the pricing of stocks in an asset market.

We begin by defining the *t*-period real interest factor between the present date and some future date *t* years from now. The *t*-period real interest factor is simply the amount by which you must discount when calculating a discounted present value of a flow benefit (already adjusted for inflation) that will be received *t* years from now.

Suppose we have an asset that will provide real dividend payments every year for *t* years. Suppose that D_t is the real dividend in period *t*, and R_t is the real interest factor from the current period to period *t*. Then the price of the asset is given by

or

All we did was to divide the dividends (D) due in period *t* by the interest factor R_t and then add them together.

If interest rates are constant over time, then the interest factors are easy to determine. Suppose that the annual real interest rate for one year is *r*. Then $R_1 = (1 + r)$ because this is the factor we would use to discount from next year to the present. What about discounting dividends two periods from now? To discount D_2 to period 1, we would divide by (1 + *r*). To discount that back again to the current period we would again divide by (1 + *r*). So to discount D_2 to the present we divide D_2 by $(1 + r) \times (1 + r) = (1 + r)^2$. That is, $R_2 = (1 + r)^2$. In general, $R_t = (1 + r)^t$ when interest rates are constant.

If real interest rates are not constant over time, the calculation of R_t is more tedious. If $R_1 = (1 + r_1)$, then $R_2 = (1 + r_1) \times (1 + r_2)$, where r_2 is the real interest rate between period 1 and period 2. In the calculation of R_2, you can think of (1 + r_2) as discounting the flow from period 2 to period 1 and then (1 + r_1) as discounting the flow from period 1 to period 0.

This page titled 10.7: Appendix- A General Formulation of Discounted Present Value is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

11: Raising the Wage Floor

11.1: Working at Minimum Wage
11.2: Nominal Wages and Real Wages
11.3: The Effects of a Minimum Wage
11.4: Minimum Wage Changes
11.5: The Minimum Wage and the Distribution of Income
11.6: Empirical Evidence on Minimum Wages
11.7: End-of-Chapter Material

This page titled 11: Raising the Wage Floor is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

11.1: Working at Minimum Wage 11.1: Working at Minimum WageEdit Update titlecancel

Even in rich economies like those of the United States or Western Europe, there are numerous jobs where the level of pay is very low. Perhaps you have experienced this yourself—for example, waiting tables, bagging groceries, or working at a fast-food restaurant. Strikingly, many of these jobs pay exactly the same hourly wage. In 2010 in the United States, for example, the wage for jobs in fast-food restaurants was often $7.25. If you worked for Burger King in Georgia or Arby's in Iowa, you were likely to receive exactly the same wage. In Washington State, you would have earned more—$8.67 an hour—but you would have again found that many different employers were offering exactly the same wage. Had you looked for a job in a fast-food restaurant in 1995 in the United States, you would probably have been offered $4.25. The story is similar in many other countries. In New Zealand, the wage at fast-food restaurants in 2010 was typically NZ$12.75 (about $9.50); in France it was €9.00 (about $12.50).

The fact that different US employers from Wisconsin to Pennsylvania offer the same hourly wage is not a coincidence. It is the result of legislation by the federal government that sets a lower limit on the wage that firms can pay. Such regulations are called *minimum wage laws*, and they are found in many different countries. Figure 11.1.1 is a poster you might have seen where you have worked. This is from the US Department of Labor and outlines your rights as an employee. This chapter is about the origins and consequences of the government intervention summarized in this poster.

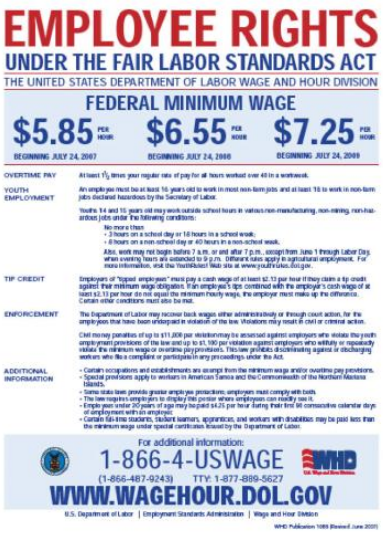


Figure 11.1.1: US Department of Labor Poster

Governments enact such laws because they want to ensure that those who work earn a "living wage." Were you to work in the United States at the current federal minimum wage for 40 hours a week, 50 weeks a year (a total of 2,000 hours), you would earn $14,500. This is slightly above the current poverty level for an individual (which is $11,369) but is well below the average income in the United States. Without minimum wage legislation, the wage earner in a family could have a full-time job, work hard every day, and still not be able to keep the family out of poverty.

Minimum wage laws have been in existence in some parts of the world for a long time. The Industrial Conciliation and Arbitration Act of New Zealand set a minimum wage more than a century ago, in 1894. The first minimum wage in the United States was established in Massachusetts in 1912. Working conditions at that time were terrible in comparison to those in modern developed economies. Women, men, and children worked long hours in very dangerous working conditions for extremely low pay. It was quite natural, confronted with these sweatshops, to feel that the government could do more to actively protect the rights of workers and secure a fair standard of living for them. Those of us fortunate to live in rich economies are now largely spared from such working conditions, but in much of the world, people continue to work in unsafe and unhealthy conditions for very low pay.

The US federal government first established a minimum wage in 1938, as part of the Fair Labor Standards Act. Not all workers were covered by this act, however. The US Constitution charges the federal government with the duty of regulating interstate trade, so the act originally covered only those workers who were involved with trade that crossed state lines. Over time, however,

amendments to the legislation have increased its coverage, and it now applies to all workers.See the following discussion for more details: US Department of Labor, Wage and Hour Division (WHD), "Fact Sheet #14: Coverage Under the Fair Labor Standards Act (FLSA)," revised July 2009, accessed March 14, 2011, http://www.dol.gov/whd/regs/compliance/whdfs14.htm.

Prior to the Fair Labor Standards Act, many states instituted their own minimum wage laws, and minimum wages still differ from state to state. For example, Oregon's current minimum wage is $8.50, about 17 percent higher than the federal minimum. These state-by-state differences complicate the life of economic historians who wish to study minimum wages. But for the economic analyst, these differences are extremely valuable because they are like an experiment: we can compare the experiences of different states with different laws.

In the United States, the minimum wage was raised in 2007, 2008, and 2009. Prior to that, the minimum wage had been constant for a decade; it had last been raised in 1997 following an act of Congress passed in 1995. President Clinton's 1995 message to the Congress, accompanying his minimum wage proposal, laid out arguments for the minimum wage increase. The following quote is from the Congressional Record:

To the Congress of the United States:

I am pleased to transmit for your immediate consideration and enactment the 'Working Wage Increase Act of 1995.' This draft bill would amend the Fair Labor Standards Act to increase the minimum wage in two 45 cents steps—from the current rate of $4.25 an hour to $4.70 an hour on July 4, 1995, and to $5.15 an hour after July 3, 1996.

…

To reform the Nation's welfare system, we should make work pay, and this legislation would help achieve that result. It would offer a raise to families that are working hard, but struggling to make ends meet. Most individuals earning the minimum wage are adults, and the average worker affected by this proposal brings home half of the family's earnings. Numerous empirical studies indicate that an increase in the minimum wage of the magnitude proposed would not have a significant impact on employment. The legislation would ensure that those who work hard and play by the rules can live with the dignity they have earned.

I urge the Congress to take prompt and favorable action on this legislation.*Congressional Record*, February 13, 1995 (House), page H1677-H1678, accessed March 3, 2011, frwebgate2.access.gpo.gov/cgi-bin/TEXTgate.cgi?WAISdocID=C6M4yE/0/1/0&WAISaction=retrieve.

President Clinton's words were forceful, and legislators in many different countries have been convinced by arguments such as these. Yet despite their widespread existence, minimum wage laws are highly contentious. Some commentators and analysts think that minimum wage laws are badly misguided and do much more harm than good.

The superficial appeal of minimum wage legislation is clear: it allows us to know that people who hold down jobs will at least earn a basic living wage—hard work will be rewarded by a minimum standard of living. Sometimes this is expressed differently: those who work hard should receive a "fair" wage for their efforts. (Fair is in quotation marks because not everyone agrees on what is fair and what is unfair, so it is hard to define exactly what the word means here.) Chapter 13 contains more discussion. In other words, minimum wage legislation has a redistributive goal: the aim is to put more of society's resources in the hands of the working poor.

Economics, however, teaches us that many policy actions can have unintended consequences. To assess whether it is a good idea for the government to intervene in this manner, we need to develop a framework for understanding the effects of minimum wage laws. In particular, we want to answer the following questions:

1. What are the consequences of a statutory minimum wage?
2. Why is there so much disagreement about whether the minimum wage is a good idea?

We will not tell you whether or not the minimum wage is a good thing. By the end of this chapter, you should be in a position to make your own informed opinion about this controversial public policy question.

Road Map

Because the minimum wage says that firms are not allowed to pay below a certain price for the labor they hire, it is natural that our analysis focuses on the labor market. Of course, there is no single labor market—rather, we might think about there being many different markets for different types of skilled individuals. Lawyers, plumbers, engineers, web designers, and airline pilots earn much more than the minimum wage. In this chapter, though, we are focusing on people who earn relatively low

wages, which means that we should look at the unskilled labor market. Sellers of labor in this market are not bringing any specialized skills; buyers of labor are not looking for any particular qualifications. The unskilled labor market is largely a market for time.

The minimum wage is set in terms of money—dollars in the United States, euros in France, and so on. Over time, increases in prices can erode the value of the minimum wage. We therefore begin this chapter by explaining how to adjust for the effects of higher prices. We then turn to the unskilled labor market. We look at what happens when we impose a minimum wage in that market, and then we look at what happens when the minimum wage changes.

Who is affected by minimum wage changes? Recognizing that people move in and out of jobs, we go beyond a supply-and-demand framework and consider these dynamic changes in the labor market. When we take into account these movements, we obtain a more sophisticated answer to this question. Then, once we are done with theory, we turn to evidence. We look at what different studies have found about the effects of changes in the real wage, and we assess how well these studies match up with our theory.

This page titled 11.1: Working at Minimum Wage is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

11.2: Nominal Wages and Real Wages 11.2: Nominal Wages and Real WagesEdit Update titlecancel

Learning Objectives

1. What is the difference between the real minimum wage and the nominal minimum wage?
2. What determines the equilibrium real wage and the level of employment?

When the federal minimum wage was first introduced in the United States in 1938, it was set at $0.25 per hour. Since then, Congress has raised the minimum wage several times. Figure 11.2.1 shows the minimum wage since 1938. You can see that the wage increases in steps whenever Congress enacts an increase in the wage.

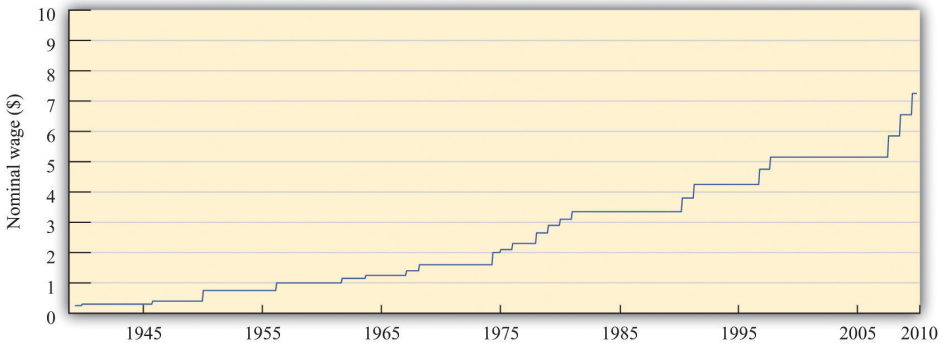


Figure 11.2.1: Nominal Federal Minimum Wage in the United States. The figure shows the federal minimum wage in the United States. When introduced in 1938 the minimum wage was set at $0.25 per hour. The minimum wage, since 2009, is $7.25 per hour.

Source: US Department of Labor.

The repeated increases in the minimum wage are not primarily due to the increased generosity of the US Congress. As you probably know, prices and wages have also tended to increase over time—a process we call **inflation**. The price level in 2007 was, on average, 14.8 times higher than in 1938, so $0.25 per hour is equivalent in modern terms to $3.70 ($0.25 × 14.8 = $3.70). Most of the increase in the minimum wage has simply been about keeping up with inflation. That said, the current minimum wage is $7.25, so the federal minimum wage has increased faster than the rate of inflation since its inception.

We call the wage in dollars the *nominal wage*. It is not the most useful measure of the amount that workers are receiving. Inflation means that a given nominal wage becomes worth less and less over time in terms of the goods and services that it buys. Between 1938 and 1957, for example, the general price level in the United States doubled. Had there been no change in the minimum wage, the $0.25 per hour minimum wage would have been worth only half as much in 1957 as it was when it was established.

From Nominal to Real Wages

The nominal wage is the wage measured in money (dollars in the United States). The **real wage** is the nominal wage in an economy adjusted for changes in purchasing power. It is defined as the nominal wage divided by the general price level:

Workers care about the real wage, not the nominal wage, because the real wage captures the trade-off between leisure time and goods and services. Firms care about the real wage, not the nominal wage, because it measures the true cost of hiring labor. Figure 11.2.2 shows the *real minimum wage*—that is, the minimum wage adjusted for inflation. The real minimum wage is defined as

Toolkit: Section 31.8

The conversion from nominal wages to real wages is an example of the more general idea of correcting for inflation. If you have data expressed in nominal terms (for example, in dollars) and want to covert them to real terms, you should follow the following four steps:

1. Select your deflator. In most cases, the Consumer Price Index (CPI) is the best deflator to use. You can find data on the CPI (for the United States) at the Bureau of Labor Statistics website (www.bls.gov).
2. Select your base year. Find the value of the index in that base year.

3. For all years (including the base year), divide the value of the index in that year by the value in the base year. Notice that the value for the base year is 1.
4. For each year, divide the value in the nominal data series by the number you calculated in step 3. This gives you the value in "base year dollars."

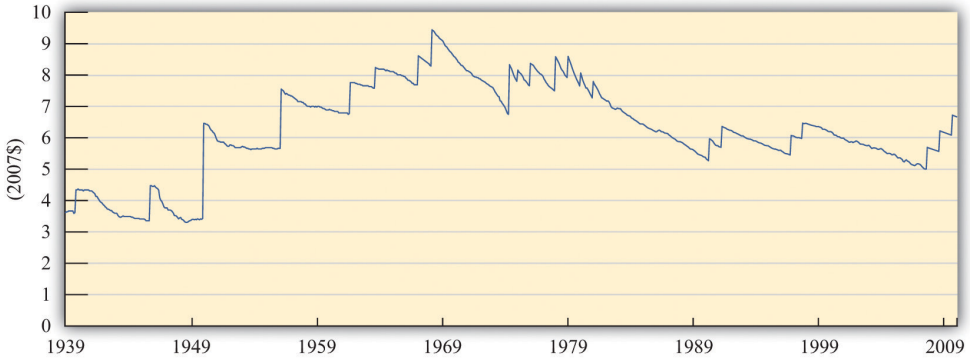


Figure 11.2.2: Real Minimum Wage in the United States. The minimum wage was at its highest in real terms in the 1960s, and the current minimum wage it is still well below that level.

Source: US Department of Labor and Bureau of Labor Statistics. Estimates for 2008 and 2009 are based on 2.5 percent annual inflation (equivalent to the average of the previous two years).

The real minimum wage increases in jumps whenever the nominal wage is increased, but it declines over time as it is eroded by inflation. The erosion of the minimum wage by inflation was recognized by President Clinton in the address that we quoted in our introduction. In that request to Congress he also said:

The first increment of the proposal simply restores the minimum wage to its real value following the change enacted in 1989.

If the Congress does not act now, the minimum wage will fall to its lowest real level in 40 years. That would dishonor one of the great promises of American life—that everyone who works hard can earn a living wage. More than 11 million workers would benefit under this proposal, and a full-time, year-round worker at the minimum wage would get a $1,800 raise—the equivalent of 7 months of groceries for the average family.

When President Clinton referred to "7 months of groceries," he was converting the increase in the minimum wage into real terms, just as our technique for converting to inflation does. Instead of using the bundle of goods that goes into the CPI, however, he was using a bundle of goods representing groceries for the average family.

Nominal and Real Wages in the Labor Market

The challenge when analyzing the minimum wage is that it is set in nominal terms, but workers and firms care about the real minimum wage. To help us understand the difference, we begin with a specific numerical example of the **labor market**. Suppose we have the following labor supply-and-demand equations, where labor supply and labor demand are measured in hours:

$$labor\ supply\ =\ 10,000\ \times\ real\ wage$$

and

$$labor\ demand\ =\ 72,000 - 8,000\ \times\ real\ wage.$$

Think of this example as referring to the weekly demand for and supply of unskilled labor in a small city. It is reasonable to think of this as a **competitive market**, in which market participants will typically agree on a price at or close to the point where supply equals demand. In the supply-and-demand framework, the intersection of the supply and demand curves tells us the equilibrium price in the market and the equilibrium quantity traded. In the labor market, the place where supply and demand meet tells us the equilibrium wage and the equilibrium number of hours worked.

Toolkit: Section 31.3, and Section 31.9

You can find more detail about the underpinnings of labor market supply and demand and the workings of the competitive market in the toolkit.

First we solve for the equilibrium in this market. In equilibrium, the quantity of labor supplied equals the quantity of labor demanded, so

$$10,000 \times \textit{real wage} = 72,000–8,000 \times \textit{real wage}.$$

Add $(8,000 \times \textit{real wage})$ to each side:

$$18,000 \times \textit{real wage} = 72,000.$$

Then divide both sides by 18,000 to obtain

$$\textit{real wage} = 4.$$

If we plug this value of the real wage back into either the supply or the demand equation, we find that the equilibrium quantity of hours worked is 40,000 hours. For example, we might have 1,000 workers, each of whom works a 40-hour week. The equilibrium is illustrated in Figure 11.2.3.

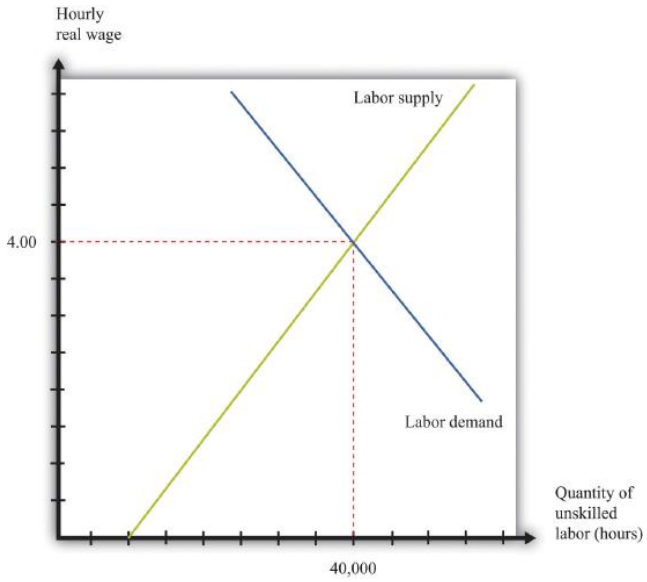


Figure 11.2.3: Labor Market Equilibrium. The market for unskilled labor is in equilibrium at an hourly wage of $4 and a total of 40,000 hours. In this diagram, we assume that the price level is 1, so the real wage equals the nominal wage.

Suppose that this example pertains to the base year. From our discussion of correcting for inflation, we know that in the base year we set the price level equal to 1. When the price level is 1, the real wage equals the nominal wage. In the initial year, therefore,

and

In the base year, the nominal wage is $4 per hour.

Now imagine we have 10 percent inflation, which means that the price of all goods and services in the economy increases by 10 percent over the course of a year. If a household paid $100 a week for groceries last year, it must pay $110 this year; if a household used to pay $500 a month in rent, it must now pay $550; and so on. Turning this around, a dollar is worth less than it used to be; you need $1.10 to purchase what you could have bought for $1 this year. The price level has increased from 1 to 1.1.

To see what this means in terms of the labor market diagram, think about the situation at a given nominal wage, such as $2.20 per hour. Last year, when the price level was 1, households were willing to supply 22,000 hours (= 10,000 × 2.2). But $2.20 now is worth the equivalent of only $2, so households are willing to supply only 20,000 (10,000 × = 10,000 × 2) hours of labor instead. The same idea applies at every wage; households will supply only the amount of labor that they would previously have supplied when the wage was 10 percent higher.

A similar logic applies to the demand for labor. The increase in the price level means that firms get 10 percent more dollars for the goods that they sell. As a consequence, the labor performed by workers generates more dollars than it used to. If it was worth paying $7 for an hour of work before, it is now worth paying $7.70 for that same hour of work.

In terms of real wages, however, nothing has changed. The equilibrium real wage is still $4, as it was before. But because

the nominal wage must increase by 10 percent to match the increase in the price level. The equilibrium in the labor market is shown in Figure 11.2.4. It is no coincidence that this diagram looks exactly the same as Figure 11.2.3; that is the point. An increase in the price level is matched by an increase in the nominal wage, and nothing changes in terms of the real wage or the real equilibrium quantity of labor.

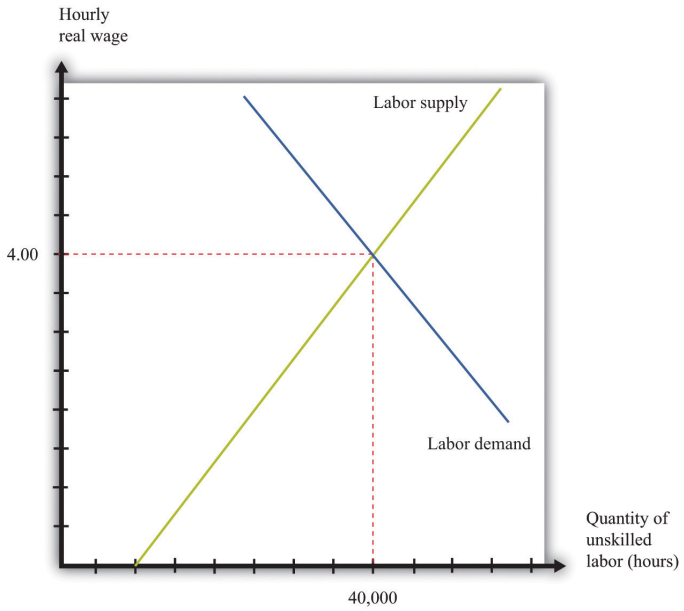


Figure 11.2.4: Labor Market Equilibrium after 10 Percent Inflation. If there is 10 percent inflation, the price level increases from 1 to 1.1, the real wage is unchanged, and the nominal wage increases by 10 percent.

Key Takeaways

- The nominal minimum wage is set by governments. The real minimum wage is the real value of the nominal minimum wage. It is determined by dividing the nominal minimum wage by the price level.
- The levels of the real wage and employment are determined by labor market equilibrium.

Exercises

1. Looking at Figure 11.2.2, explain why the real minimum wage increases very quickly but never decreases very quickly.
2. Why do labor demand and supply depend on the real and not the nominal wage?

This page titled 11.2: Nominal Wages and Real Wages is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

11.3: The Effects of a Minimum Wage 11.3: The Effects of a Minimum WageEdit Update titlecancel

Learning Objectives

1. What happens when a government imposes a minimum wage?
2. If there is inflation under a minimum wage system, what happens to the level of employment?
3. What are the efficiency costs of a minimum wage?

Adam Smith, the 18th-century economist who founded modern economics, had a vivid metaphor for the idea that supply would equal demand in a competitive market: he referred to an "invisible hand" guiding markets to equilibrium. Joan Robinson, a famous economist at Cambridge University in the first half of the 20th century, wrote that "the hidden hand will always do its work but it may work by strangulation."Joan Robinson, "The Pure Theory of International Trade," in *Collected Economic Papers* (Oxford: Basil Blackwell, 1966), 189. What she meant by this was there is no guarantee that the *equilibrium* wage in the market would in fact be a *living* wage.

The Imposition of a Minimum Wage

When the government imposes a minimum wage, firms are not permitted to pay less than the amount that the government mandates. Suppose we are again in the base year, so the price level is 1. Imagine that the market equilibrium wage is $4 per hour, but the government now passes legislation stating that all firms must pay at least $5 per hour. At this wage, supply does not equal demand. Figure 11.3.1 illustrates what happens.

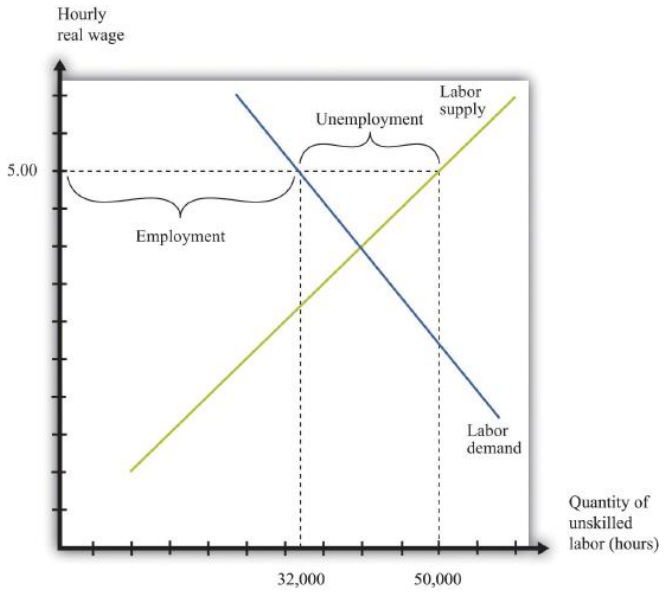


Figure 11.3.1: Labor Market with a Minimum Wage. With a minimum wage of $5, the supply of labor is 50,000 hours, but firms demand only 32,000 hours of labor, so the labor market is not in equilibrium.

Markets are based on voluntary trades. In Figure 11.3.1, we see that sellers (the workers who supply labor) would like to sell 50,000 hours of labor to the market at the set minimum wage—that is, 250 more people would like to have a 40-hour-a-week job when the wage increases from $4 to $5. But firms wish to purchase only 32,000 hours of labor—firms want to hire 200 *fewer* workers (8,000 fewer hours). In a market with voluntary trade, no one can force firms to hire workers. As a result, the equilibrium quantity of labor traded in the market will be determined by how much the firms wish to buy, not how much workers want to sell.

We can now answer our first motivating question of the chapter: what is the consequence of imposing a minimum wage? Two things happen when the government imposes a minimum wage:

1. The amount of labor hired in the market decreases. In our example, the number of unskilled workers employed decreases from 1,000 to 800. Thus while those who have jobs earn a higher wage, there are now some individuals who no longer have jobs. Employment has decreased.

2. At the government-imposed wage, there are more people who want to work than are able to find jobs. Thus the minimum wage has created unemployment. Because 1,250 people would like jobs at a wage of $5 but only 800 jobs are available, 450 people are unemployed; they would like a job at the prevailing wage, but they are unable to find one.

The number of unemployed workers is 450, even though employment decreased by only 200 workers. The difference comes from the fact that the higher wage also means that more people want to work than before. In this case, the higher wage means 250 more people would like a job.

We have assumed in this discussion that everyone works for 40 hours, in which case the number of people employed must decrease by 200. Another possibility is that everyone who wants a job is able to get one, but the number of hours worked by each individual decreases. Because there are 32,000 hours of work demanded and 1,250 people who want jobs, each worker would work 25.6 hours a week. In this situation, we say that there is *underemployment* rather than unemployment. Yet another possibility is that, after the introduction of the minimum wage, the number of people employed stays the same as before (1,000), but those individuals are allowed to work only 32 hours per week. In this case, we have both underemployment (of the previously employed) and unemployment (of the extra workers who want a job at the higher wage). In real-life situations, there may be both unemployment and underemployment.

Inflation and the Minimum Wage

Although inflation made no difference to our basic analysis of the labor market, it does change our analysis of the minimum wage. Minimum wages are fixed in nominal terms and do not automatically change when there is inflation. So if the minimum wage is set at $5 and the price level increases from 1 to 1.1, the real minimum wage declines. Looking back at the definition of the real minimum wage, we find that

The effect of a reduction in the real minimum wage is shown in Figure 11.3.2. At the lower real wage, firms are willing to hire more workers. Employment increases from 32,000 hours to 35,600 hours: 90 more people can find jobs.

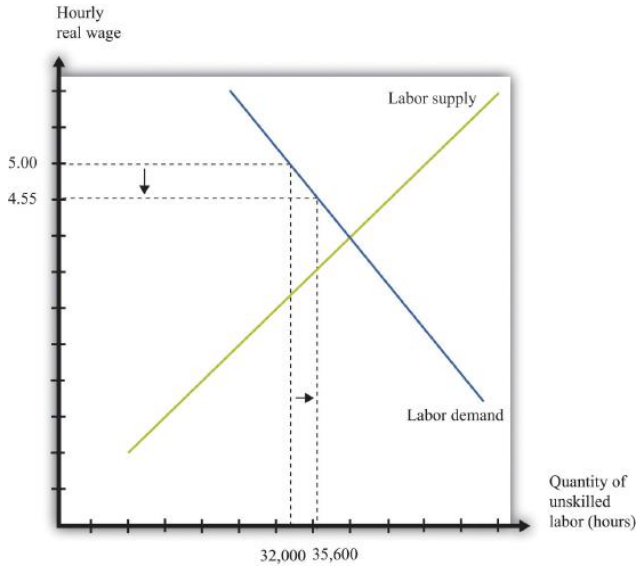


Figure 11.3.2: A Reduction in the Real Minimum Wage. A 10-percent increase in the price level leads to a reduction in the real minimum wage to $4.55 and an increase in employment from 32,000 to 35,600.

In Figure 11.3.2, the real minimum wage of $4.55 is still higher than the equilibrium wage of $4.00. To put the same point another way, the equilibrium nominal wage has increased to $4.40, but this is still below the nominal minimum wage of $5.00. However, if the price level were to increase by 25 percent or more from the base year, the minimum wage would become completely irrelevant. The minimum wage would be below the market wage. Economists say that the minimum wage would no longer be “binding” in this case.

It is exactly this process of increasing prices that lies behind Figure 11.3.2. As the price level increases, the minimum wage becomes worth less in real terms (and has less of an effect on employment). Eventually, Congress acts to increase the minimum wage to bring it back in line with inflation—although, as Figure 11.3.2 shows, Congress has allowed the real minimum wage to decline substantially from its high point in the late 1950s. The reduction in the real minimum wage also leads to a reduction in unemployment, as shown in Figure 11.3.3.

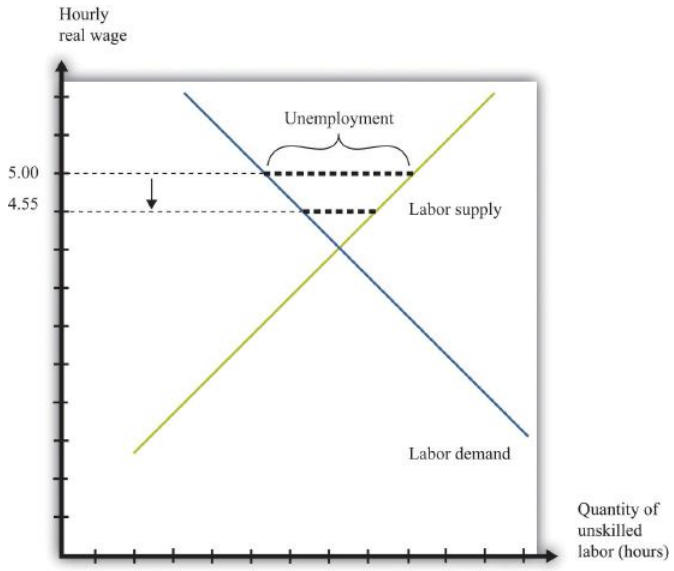


Figure 11.3.3: Effects on Unemployment of a Reduction in the Real Minimum Wage. The reduction in the real minimum wage to $4.55 leads to a decrease in unemployment.

Efficiency Implications of a Minimum Wage

Markets are a mechanism that allow individuals to take advantage of gains from trade. Whenever a buyer has a higher valuation than a seller for a good or service, they can both benefit from carrying out a trade. This is how economies create value—by finding opportunities for mutually beneficial trades.

The minimum wage interferes with this process in the unskilled labor market. It reduces employment, which is the same as saying that fewer transactions take place. Because each voluntary transaction by definition generates a surplus, anything that reduces the number of transactions causes a loss of surplus. In economists' terminology, the minimum wage leads to a departure from **efficiency**. We can represent that inefficiency graphically. Figure 11.3.4 shows the effect of the minimum wage, using the ideas of **buyer surplus** and **seller surplus**.

Toolkit: Section 31.10 and Section 31.11

You can review the different kinds of surplus, as well as the concepts of efficiency and deadweight loss, in the toolkit.

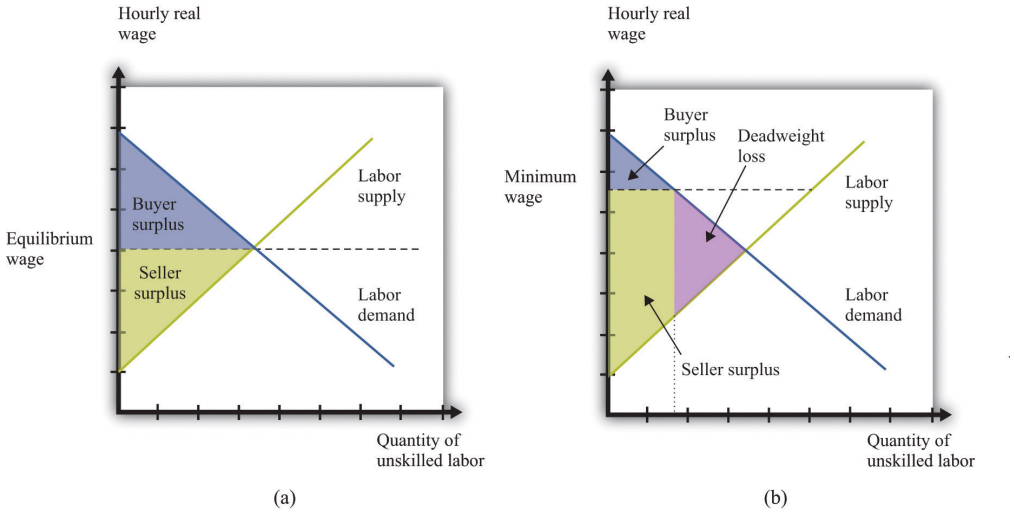


Figure 11.3.4: Deadweight Loss from Minimum Wage. With no minimum wage (a), all the possible gains from trade in the market are realized, but with a minimum wage (b), some gains from trade are lost because there are fewer transactions.

In part (a) of Figure 11.3.4, we see the market without the minimum wage. In the labor market, it is the firm who is the buyer. The total buyer surplus is the **profit** that firms obtain by hiring these workers; it is the difference between the cost of hiring these

workers and the revenues that they generate. Graphically, it is the area below the labor demand curve and above the market wage. The total sellersellers' surplus#8217;s surplus is the benefit that accrues to workers from selling labor time. Sellers of labor (workers) receive surplus equal to the area below the market wage and above the supply curve.

In part (b) of Figure 11.3.4, we show the effect of the minimum wage. As we already know, the higher wage leads to a reduction in employment. Fewer transactions occur, so the total surplus in the market is reduced. Economists call the lost surplus the **deadweight loss** from the minimum wage policy.

The most obvious cost of the minimum wage is this loss of surplus. But there may be other hidden costs as well. Whenever people are prevented from carrying out mutually beneficial trades, they have an incentive to try to get around these restrictions. Firms and workers may try to "cheat," conducting hidden transactions below the minimum wage. For example, a firm might pay a worker for fewer hours than he or she actually worked. Alternatively, a firm might reduce other benefits of the job. Such cheating not only subverts the minimum wage law but also uses up resources because the firm and the worker must devote effort to devising ways around the law and ensuring that they do not get caught.

The loss in surplus could also be greater than is shown in Figure 11.3.4. The figure is drawn under the presumption that the trades taking place in the labor market are the ones that generate the most surplus. But suppose that the minimum wage is $5.00. It is possible that someone who would be willing to work for, say, $2.00 an hour loses her job, whereas someone willing to work for, say, $4.50 is employed.

Remember that if there were no restrictions in the labor market, the wage would adjust so that anyone wanting to work could find a job. This means that both the person willing to work for $2.00 and the person willing to work for $4.50 could both find a job as long as the wage is above $4.50. If the equilibrium wage were $4.00, then the person willing to work for only $4.50 would not be employed. In either case, this is the efficient outcome, consistent with obtaining all the gains from trade.

Key Takeaways

- When the government imposes a minimum wage, the real wage is determined by the minimum wage divided by the price level, not by the interaction between labor supply and demand.
- If there is inflation and a fixed nominal minimum wage, then the level of employment will increase and the real minimum wage will decrease.
- The minimum wage creates deadweight loss because some trades of labor services do not take place.

check your understanding

1. Draw a diagram for a labor market where the minimum wage is not binding.
2. What happens to the real minimum wage and the level of employment if there is *deflation*—that is, if the price level decreases?

This page titled 11.3: The Effects of a Minimum Wage is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

11.4: Minimum Wage Changes 11.4: Minimum Wage ChangesEdit Update titlecancel

Learning Objectives

1. What happens to the levels of employment and unemployment if the real minimum wage increases?
2. What determines the size of the change in employment when the real minimum wage increases?
3. What determines the size of the change in unemployment when the real minimum wage increases?

Suppose that the government is considering an increase in the minimum wage. What should we expect to happen? How will firms and workers respond? One might be tempted simply to ask firms what they would do in the face of an increase in the minimum wage. Unfortunately, this is likely to be both infeasible (or at least prohibitively expensive) and inaccurate. It would be an immense amount of work to interview all the firms in an economy. What is more, there is no guarantee that managers of firms would give accurate answers if they were asked hypothetical questions about a change in the minimum wage. Instead, government statisticians use statistical sampling techniques to interview a random sample of firms in an economy, and they ask them about their actual behavior—they ask questions such as the following: "How many workers do you employ at present?" and "How much do you pay them?" The data from such surveys are useful but do not directly help us determine the effects of a change in the minimum wage. For this we need more theory.

The Effect of a Minimum Wage Increase on Employment and Unemployment

Figure 11.4.1 amends our view of the labor market to show an increase in the minimum wage from $5 to $6. (We suppose that the price level is constant, so an increase in the nominal minimum wage implies an increase in the real minimum wage.) The increase in the minimum wage leads to a reduction in the level of employment: employment decreases from 32,000 to 24,000. Labor is now more expensive to firms, so they will want to use fewer hours. At the same time, the higher minimum wage means that more people would like jobs. The increase in the amount of labor that people would like to supply, and the decrease in the amount of labor that firms demand, both serve to increase unemployment.

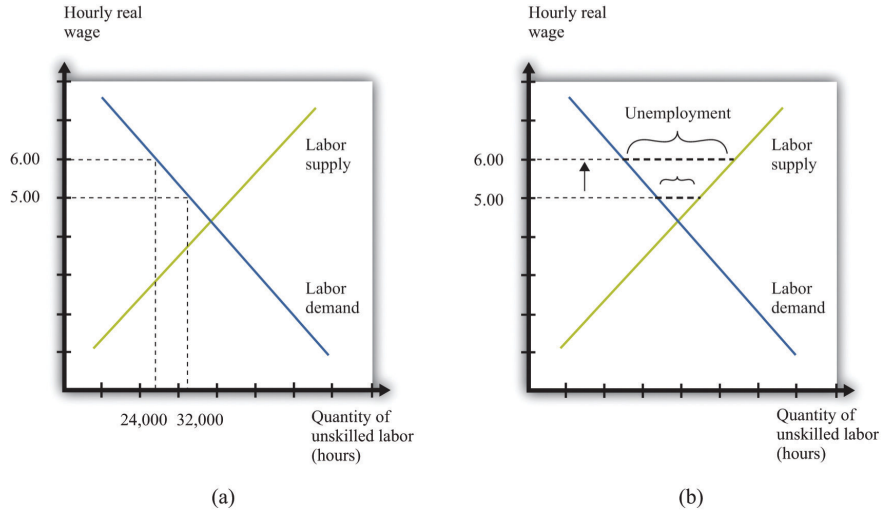


Figure 11.4.1: Effects of Increasing the Real Minimum Wage. An increase in the value of the hourly real minimum wage from $5 to $6 leads to a decrease in employment from 32,000 hours to 24,000 hours (a) and an increase in unemployment (b).

Our model generates a *qualitative* prediction: an increase in the minimum wage will decrease employment and increase unemployment. At the same time, the wage increase will ensure that those with jobs will earn a higher wage. So we can see that there may be both advantages and disadvantages of increasing the minimum wage. To go further, we have to know how big an effect such a change would have on employment and unemployment—that is, we need the *quantitative* effects of a higher minimum wage.

To understand the quantitative effects, we want to know when to expect big or small changes in employment or unemployment—which depends on the **wage elasticity of labor demand** and the labor supply. Remembering that the wage is simply the price in the labor market, the wage elasticity of demand is an example of the **price elasticity of demand** in a market:

From Figure 11.4.1, we can see that the wage elasticity of labor demand tells us everything we need to know about the effects of a change in the wage on employment. If the demand curve is relatively elastic, then a change in the minimum wage will lead to a relatively large change in employment. If the demand curve is relatively inelastic, then a change in the minimum wage will lead to a relatively small change in employment. This is intuitive because the elasticity of labor demand tells us how sensitive firms' hiring decisions are to changes in the wage. An elastic demand for labor means that firms will respond to a small change in the wage by laying off a large number of workers, so the employment effect will be large. The elasticity of labor supply is not relevant if we are concerned only with employment effects. This is illustrated in Figure 11.4.2 and summarized in Table 11.4.1 "Employment Effects of a Change in the Real Minimum Wage".

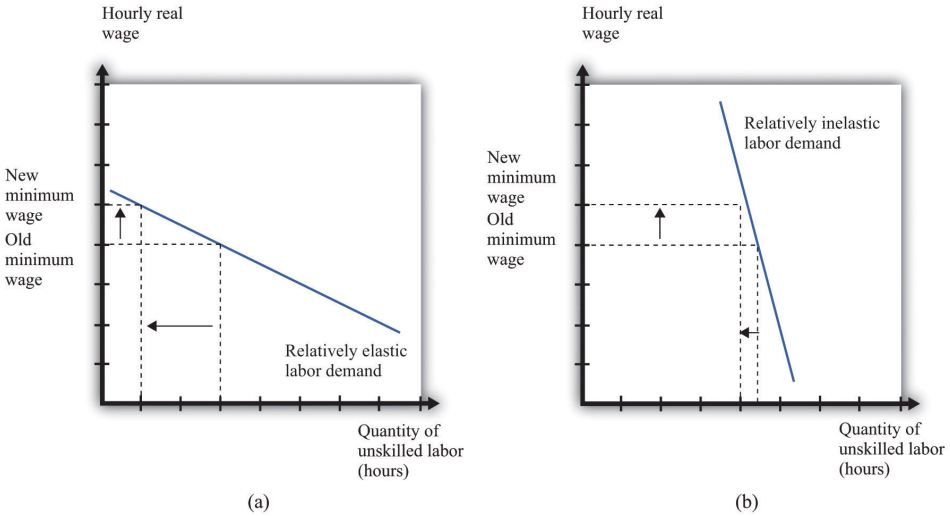


Figure 11.4.2: The Employment Effect of a Change in the Minimum Wage. If labor demand is relatively elastic (a), a change in the minimum wage has a big effect on employment, while if labor demand is relatively inelastic (b), the same change in the minimum wage has a much smaller effect on employment.

Table 11.4.1: Employment Effects of a Change in the Real Minimum Wage

	Effect on Employment
Elastic demand	Large change
Inelastic demand	Small change

If we are interested in the effect on unemployment, however, we must look at both demand and supply. A worker is counted as unemployed if he or she is looking for a job but does not currently have a job. The labor supply curve tells us how many workers are willing to work at a given wage; those who are not employed are looking for a job. To understand the effects of the minimum wage on unemployment, we need to look at the mismatch between supply and demand at the minimum wage, so we must look at the supply of labor as well as the demand for labor. The **price elasticity of supply** measures the responsiveness of the quantity supplied to a change in the price: in the case of the labor market, we obtain the **wage elasticity of labor supply**:

Toolkit: Section 31.2

You can review the general definition and calculation of elasticities in the toolkit.

The more elastic the labor supply curve, the bigger the change in labor supply for a given change in the real wage (Figure 11.4.3). A bigger change in labor supply means a bigger change in unemployment. Combining this with Table 11.4.1 "Employment Effects of a Change in the Real Minimum Wage", we get the results summarized in Table 11.4.2 "Unemployment Effects of a Change in the Real Minimum Wage". If demand and supply are both inelastic, the change in the minimum wage has little effect on unemployment. The higher wage does not make much difference to firms' hiring decisions (inelastic demand), and it does not induce many additional workers to look for a job (inelastic supply). The overall effect on unemployment is small. By contrast, if both curves are elastic, then an increase in the wage will lead to a big decrease in the number of jobs available and a big increase in

the number of job seekers. If we can find good estimates of the elasticities of labor demand and supply, we will be able to make good predictions about the likely effect of an increase in the minimum wage.

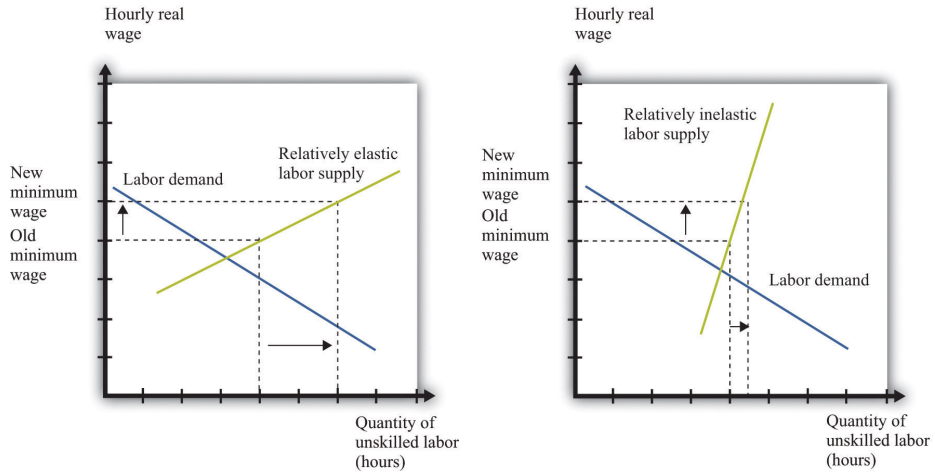


Figure 11.4.3: The Unemployment Effect of a Change in the Minimum Wage. If the labor supply is relatively elastic (a), a change in the minimum wage has a big effect on unemployment, while if the labor supply is relatively inelastic (b), the same change in the minimum wage has a much smaller effect on unemployment.

Table 11.4.2: Unemployment Effects of a Change in the Real Minimum Wage

	Effects on Unemployment	
	Elastic Supply	**Inelastic Supply**
Elastic demand	Very large change	Large change
Inelastic demand	Large change	Small change

Key Takeaways

- All else being the same, an increase in the real minimum wage will reduce employment and increase unemployment.
- The size of the change in employment when the minimum wage increases is determined by the elasticity of the labor demand curve.
- The size of the change in unemployment when the minimum wage increases is determined by the elasticities of the labor demand and supply curves.

check your understanding

1. Why doesn't the elasticity of labor supply matter for the effects of changes in the real minimum wage on employment?
2. If prices increase, what will happen to the level of unemployment when there is a binding minimum wage?

This page titled 11.4: Minimum Wage Changes is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

11.5: The Minimum Wage and the Distribution of Income 11.5: The Minimum Wage and the Distribution of IncomeEdit Update titlecancel

Learning Objectives

1. Which parts of the economy are affected by the minimum wage?
2. When the minimum wage increases, who gains and who loses?
3. What is an equity-efficiency trade-off?

We said earlier that governments impose minimum wages because they care about ensuring that the working poor earn a fair wage. Another way of saying this is that the minimum wage is an intervention by the government that is meant to change the distribution of society's resources. If unskilled workers are going to earn more, then this means they are obtaining more of the total resources available in an economy. And if they are getting more, then somebody else must be getting less. We would like to have some way of thinking about the effects of the minimum wage on the distribution of income.

To talk about distribution, we need to divide society into groups and then examine how much each group gets. One group is obviously those who receive the minimum wage—the working poor. Another group we need to consider is the unemployed. And then there is everybody else: all of those individuals who are sufficiently skilled to have jobs that pay more than the minimum wage. This is a large group, encompassing electricians and CEOs, but—for our present purposes—it makes sense to group them all together. Let's call them the "relatively rich," by which we mean that they are richer than unemployed or employed unskilled workers. So we have three groups: the unemployed, the working poor, and the relatively rich. How are these groups affected by an increase in the minimum wage?

Winners and Losers from the Minimum Wage

We know that the working poor are made better off by an increase in the minimum wage—after all, the whole point of the increase is to ensure that these individuals earn more. We can see this gain in Figure 11.5.4: it is the extra surplus that sellers obtain at the expense of buyers. Remember, though, that some of the working poor will lose their jobs as a result of the higher minimum wage. So our first conclusion is that those members of the working poor who keep their jobs are unambiguously made better off.

The increase in the minimum wage means that there are more people who are unemployed. Those who are *already* unemployed are not directly affected by the increase in the minimum wage. These unemployed individuals may be indirectly affected, however, because it becomes harder for them to find jobs.

Are there any effects on the relatively rich? The answer is yes. The increase in the minimum wage means that firms will earn lower profits. We can see this because buyer surplus is reduced in Figure 11.5.4. Although firms are just legal entities, they are owned by individuals. When a firm earns lower profits, the shareholders of that firm receive lower income. The working poor and the unemployed are not, for the most part, individuals with portfolios of stocks; the shareholders of firms are the relatively rich. Thus the relatively rich are made worse off by the increase in the minimum wage. Our broad conclusion is therefore that the working poor benefit from an increase in the minimum wage, but everybody else in society is made worse off.

There is another concern when we think about the distribution of income. One consequence of the minimum wage is that jobs become a scarce resource: more people want jobs than there are jobs available. We must consider how this scarce resource is allocated. Do workers line up outside factory gates? In this case, the time that they spend waiting in line is an additional cost of the minimum wage. Does some individual control who gets hired? Then there is the potential for corruption, whereby jobs are sold, meaning that the gains from the minimum wage flow not to workers but to this individual instead.

Does the Minimum Wage Benefit Unskilled Labor as a Whole?

We have concluded so far that the minimum wage benefits the working poor but at the cost of creating unemployment: some people who used to have jobs will lose them as a result of the minimum wage. Because of the flows between unemployment and employment in a dynamic labor market, it does not really make sense to think of the unemployed and the employed as different people. If we instead think about unskilled labor as a whole—a group that includes both those with jobs and those unable to find jobs—what can we conclude about the effects of an increase in the minimum wage? There are a few different ways of looking at this question.

The Wage Bill

First, we can look at total wages paid in the labor market, sometime called the *wage bill*. By looking at the wage bill, we can find out if the additional wages earned by the working poor exceed the wages lost by those who find themselves unemployed.

Total wages are equal to the total hours worked multiplied by the hourly wage:

$$total\ wages\ =\ real\ minimum\ wage\ \times\ hours\ worked.$$

Because we are measuring the wage in real terms, total wages are likewise measured in real terms. Figure 11.5.1 provides a graphical interpretation of the wage bill: total wages paid are given by the shaded rectangle.

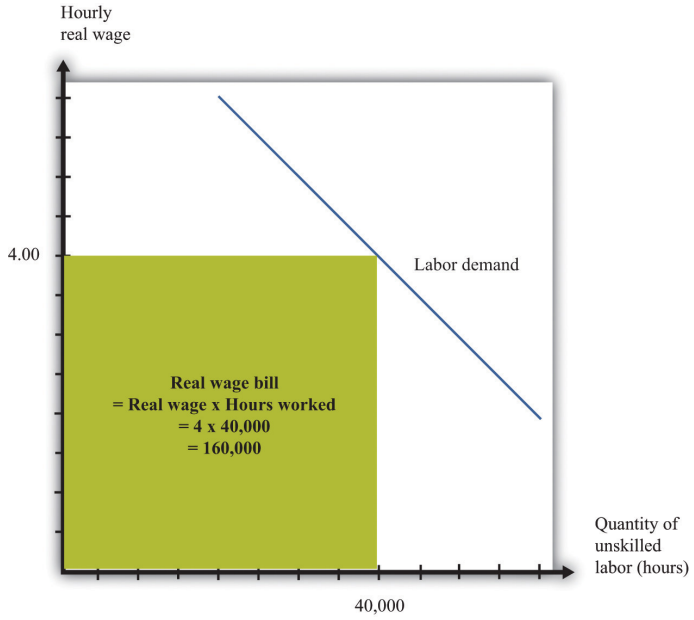


Figure 11.5.1: The Wage Bill. The wage bill is equal to the rectangle under the demand curve. For example, if the real wage is $4 per hour and employment is 40,000 hours, then the wage bill is 160,000.

From this equation it can be shown that To derive this equation, we first apply the rules of growth rates to obtain

$$percentage\ change\ in\ total\ wages\ =\ percentage\ change\ in\ real\ minimum\ wage\ +\ percentage\ change\ in\ hours\ worked.$$

Then divide both terms by the percentage change in the minimum wage and use the definition of the elasticity of demand:

$$percentage\ change\ in\ total\ wages\ =\ percent\ change\ in\ real\ minimum\ wage\ \times\ (1–[–(elasticity\ of\ demand)]).$$

The elasticity of demand is a negative number: if wages increase (that is, the change in the wage is positive), then hours worked decreases (that is, the change in hours worked is negative). It is therefore easier if we use –(elasticity of demand) because this is a positive number. The equation tells us that the change in the total wage is positive if the percentage increase in wages is greater than the percentage decrease in hours worked—in other words, if –(elasticity of demand) is less than 1.

If the demand for labor is relatively sensitive to changes in the wage, employment will decrease significantly following an increase in the minimum wage. Total wages paid will decrease. This is shown in part (a) of Figure 11.5.2. Before the increase, total wages are given by the sum of areas A and B. After the increase, total wages are given by the sum of areas A and C. We get the opposite conclusion if labor demand is inelastic. In this case, an increase in the wage increases the total wages paid. The conclusion is intuitive: if employers do not change their hiring very much when wages increase, then total wages will increase. But if an increase in the minimum wage leads to a big decrease in the demand for labor, total wages paid will decrease.

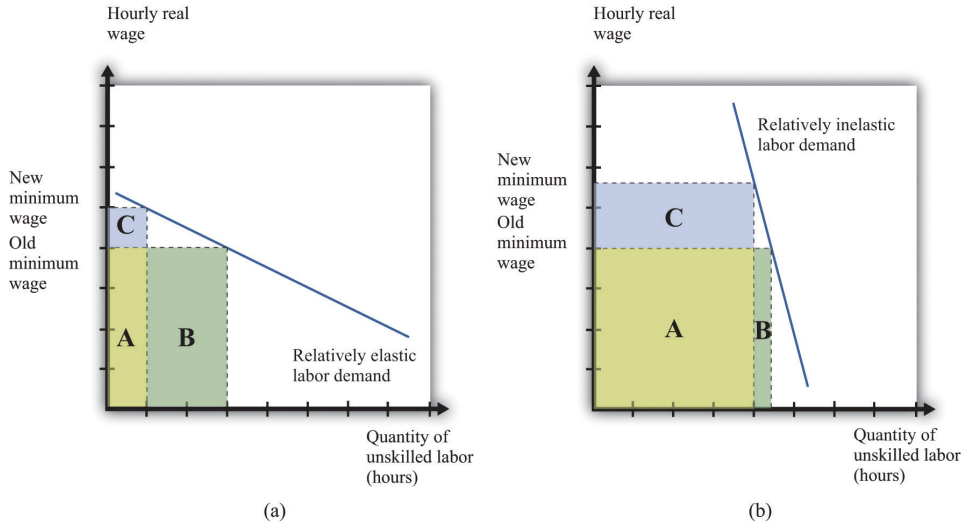


Figure 11.5.2: Effects of an Increase in the Minimum Wage on the Wage Bill. If labor demand is relatively elastic (a), a change in the minimum wage leads to a reduction in the wage bill: the original wage bill is A + B, and the new wage bill is A + C. If labor demand is relatively inelastic (b), the same change in the minimum wage leads to an increase in the wage bill.

Seller Surplus

The wage bill tells us how much workers are paid in total. A better measure of the benefits obtained by workers is the total sellersellers' surplus#8217;s surplus in the market. We cannot measure this exactly unless we know exactly what the labor supply curve looks like, but we can conclude that just looking at the wage bill understates the benefits to workers of the increased wage. The reason is simple and does not even need any diagrams. Following an increase in the minimum wage, workers work fewer hours in total. Everything else being the same, people prefer leisure time to working. For example, suppose that total wages increase following an increase in the minimum wage. Then workers gain twice: they are being paid more, and they are working less.

Even if total wages decrease, workers might still be better off. They might be more than compensated for the lower wages by the fact that they don't have to work as many hours. We are not saying that having a job is a bad thing; those who are working prefer having a job to being unemployed. But those who are working also prefer working fewer hours.

Expected Wage

So far, we have looked at the minimum wage through the lens of a competitive labor market. This is not a bad approach: as we have argued, the unskilled labor market is probably a reasonably good example of a competitive market. It is, however, a static way of thinking about the labor market, when the labor market is in fact highly dynamic. Chapter 9 contains more discussion. People move in and out of jobs: they quit or are laid off from old jobs, and they search for new jobs. A worker who is employed this month may find herself unemployed next month; a worker with no job this month may be hired next month.

Earlier we claimed that an increase in the minimum wage has no direct effect on the unemployed. This is true in the static labor market picture, but once we take a more dynamic view of the labor market, it no longer makes very much sense to draw a hard-and-fast distinction between the employed and the unemployed. Over time, they will include many of the same people. So when we look at the distributional effects of the minimum wage, it is better to draw the distinction between unskilled workers (that is, the employed and unemployed together) and the relatively rich. With this in mind, let us now consider whether the unskilled as a group are likely to benefit from the minimum wage.

We might expect that unskilled workers will spend some of their time employed and some unemployed. When employed, they earn the minimum wage, but when unemployed, they receive much less. To keep things simple, suppose these workers earn nothing when unemployed. On average, the fraction of time that workers spend employed rather than unemployed is given by

We can think of this as the **probability** that a typical unskilled worker will be employed at any given time. Combining this with the idea of **expected value**, we can calculate the expected wage of such a worker. If a worker earns nothing when unemployed, then the expected wage is as follows:

Toolkit: Section 31.7

You can review probability and expected values in the toolkit.

How does this expected wage change when there is an increase in the minimum wage? The answer, as you might expect by now, depends on the elasticities of demand and supply. Specifically, it turns out that the expected wage will increase if

$$-(elasticity\ of\ labor\ demand)\ +\ elasticity\ of\ labor\ supply\ <1.$$

If both demand and supply are sufficiently inelastic, the average wage will increase. Conversely, if they are both relatively elastic, then expected wages will decrease.

There are some things missing from this story. In a more careful analysis, we would take into account the fact that workers are probably **risk-averse** and dislike the randomness of their earnings. We would likewise take into account that unemployed workers obtain some income—perhaps from unemployment insurance. This actually makes it more likely that an increase in the minimum wage will increase the expected wage. These are details, however. We can draw two big conclusions from our discussion so far:

1. Under some circumstances at least, an increase in the minimum wage will have the effect of redistributing income from the relatively rich to unskilled workers. Policymakers may wish to reduce inequality in society, and the minimum wage is one possible tool that they can use. At the same time, the minimum wage comes at a cost to society: it distorts decisions in the labor market and leads to deadweight loss. This is an **equity-efficiency trade-off**.
2. The distributional impact of a change in the minimum wage cannot be deduced from economic theory. It depends on the elasticities of labor demand and supply in the market for unskilled labor and is an empirical question. In the next section, therefore, we turn to the evidence on minimum wages.

Key Takeaways

- The minimum wage affects buyers and sellers of labor services in the markets where the minimum wage is binding. It also affects the owners of these firms.
- Buyers and sellers of labor services in low wage (unskilled) labor markets are directly affected by changes in the real wage. Workers in this market who keep their jobs are better off. Those that lose their jobs are made worse off, at least in the short run. Owners of the firms are hurt because of reduced profits from the minimum wage. The magnitude of these effects depends on the elasticities of labor demand and supply.
- The equity-efficiency trade-off means that policies that increase equity can create inefficiencies. The minimum wage provides an example of that trade-off.

check your understanding

1. Why does an increase in the minimum wage reduce the profits of a firm?
2. If the real minimum wage is reduced by inflation, can the real wage bill paid to workers increase?

This page titled 11.5: The Minimum Wage and the Distribution of Income is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

11.6: Empirical Evidence on Minimum Wages 11.6: Empirical Evidence on Minimum WagesEdit Update titlecancel

Learning Objectives

1. How do economists determine the elasticities of labor supply and labor demand?
2. What are the estimates of these elasticities?
3. What are the estimated effects of an increase in the minimum wage?

For the most part, economists cannot carry out experiments to test their theories and must use much more indirect methods. They must rely on observations that are generated by the everyday experience of individuals in an economy. In a textbook like this, we constantly draw demand and supply curves, and we get so used to seeing these diagrams that we might be fooled into thinking that we can just go out and observe them in the real world. In fact, all we observe are the market outcomes—the equilibrium price and quantity that are traded. Our conception of the labor market might look like part (a) of Figure 11.6.1, but the data that we actually gather look like part (b) of Figure 11.6.1.

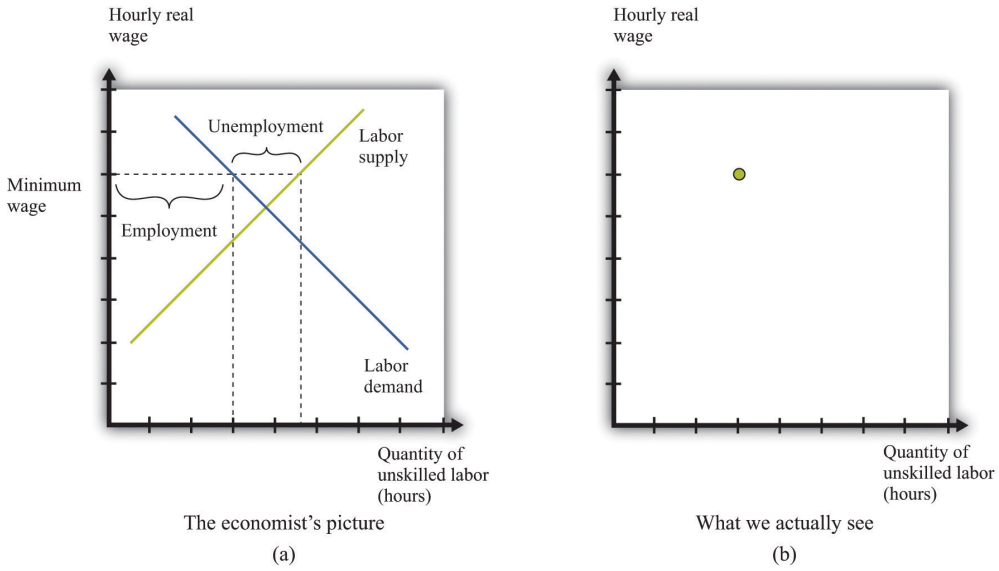


Figure $PageIndex1$: Models and Data. We construct an entire framework based on the supply and demand curves for labor (a), but at any time we observe only a single data point: the wage and the level of employment (b).

If we want to *estimate* a demand curve, we need much more than part (b) of Figure 11.6.1. We need more data points. We need different observations. In the case of the labor market, we might be able to use the fact that the minimum wage changes over time. Figure 11.6.2 shows an example. Minimum wage changes allow us to observe different points on the labor demand curve. Given enough observations, we might be able to get a good idea of what the demand curve looks like and come up with an estimate of labor elasticity.

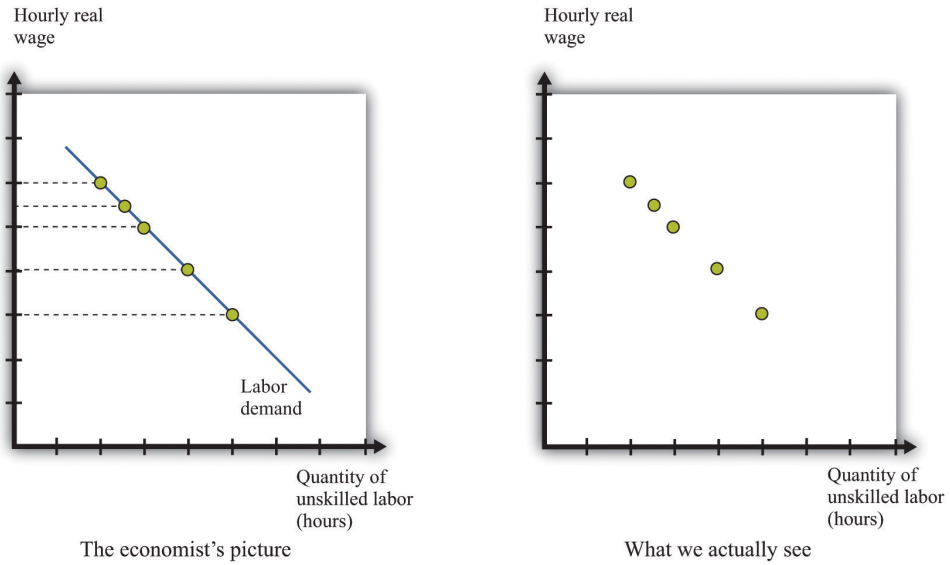


Figure 11.6.2: Inferring Labor Demand from Data. We may be able to infer a demand curve for labor by looking at what happens to the quantity of hours worked as the minimum wage changes.

Reality is messier. In actual labor markets, many things are going on at once. At the same time that the wage is changing, firms might be facing changes in the demand for their products, changes in the costs of other inputs, changes in their technology, or changes in their competitive environment. All of these changes would cause the labor demand curve to shift. As an example, look at Figure 11.6.3. The story here is as follows: The minimum wage is increasing over time, but—perhaps because of increased product demand—the demand for labor is also increasing over time. What we observe is shown in Figure 11.6.3. The unwary analyst, looking at these data, might conclude that minimum wages have little or no effect on the demand for labor.

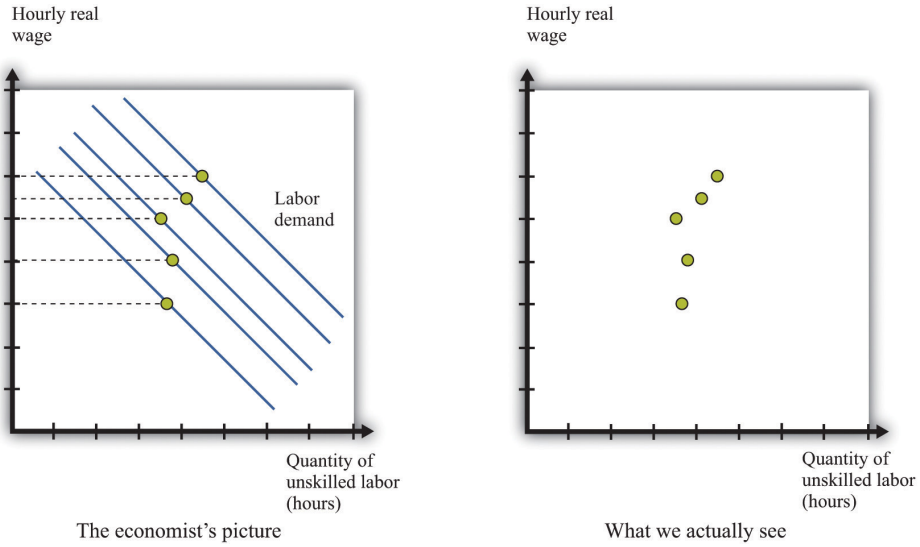


Figure 11.6.3: Difficulties Inferring Labor Demand from Data. When the demand curve for labor is shifting at the same time as the minimum wage is changing, it is more difficult to see the effects of the minimum wage.

Economists who analyze data are forever trying to distinguish effects of interest (for example, minimum wage changes) from those caused by changes in other variables (for example, product demand). The key to successful empirical work is obtaining informative sources of variation and excluding irrelevant sources of variation. In the case of the minimum wage, we might look at differences in the minimum wage at different times, or we might look at differences in the minimum wage in different places (such as specific states in the United States).

Estimates of Labor Demand Elasticity

Labor economist Daniel Hamermesh summarizes the findings from numerous studies in his book, *Labor Demand*.Daniel S. Hamermesh, *Labor Demand* (Princeton, NJ: Princeton University Press, 1996). Based on his review of all these studies,

Hamermesh argues that a good estimate of −(elasticity of labor demand) is about 0.3. This means that if we increase the minimum wage by 10 percent, employment will decrease by about 3 percent. With these, we would conclude that labor demand is relatively inelastic, and the employment and welfare implications of the minimum wage are not that large. Labor supply also tends to be relatively inelastic, partly because the **income effects** and **substitution effects** of changes in the wage tend to be offsetting. It certainly seems possible, then, that an increase in the minimum wage will raise expected wages.

We must be careful, though. Hamermesh notes that labor demand becomes more elastic as the skill level of workers decreases. It is difficult to find substitutes for workers with specialized skills. If you need an electrician, then you must hire an electrician, even if plumbers are much cheaper. If you need an airline pilot, you can't hire a computer programmer. If you are running a store, however, and you find that labor is becoming more expensive, you might be able to upgrade your cash registers and inventory control systems and get by with fewer employees. We might expect the elasticity of demand at the minimum wage to be significantly higher.

It is therefore useful to look at studies that focus directly on minimum wage changes. One approach is to look at the relationship over time between the minimum wage and the employment experience of groups that are most directly affected by the minimum wage. Alternatively, one can look across groups that are differentially affected by the minimum wage to gauge its effect.

A reminder before we proceed. The point of going through this material is not only to help you understand the effects of a change in the minimum wage but also to provide you with a glimpse of how economic research proceeds so that you can do a better job of evaluating evidence that economists compile in all sorts of areas.

Economists have focused particular attention on teenage workers because that group is typically unskilled and is likely to be subject to the minimum wage. (This is not to say that the *effects* of the minimum wage are primarily on teenagers. Indeed, about two-thirds of minimum wage earners are adults.) Figure 11.6.4 graphs the teenage unemployment rate over the period from 1956 to 2007 along with the real minimum wage for this period. We would like to know if the minimum wage has an impact on the teenage unemployment rate. This figure is suggestive of a relationship, particularly in the last couple of decades. For example, the real minimum wage fell during the 1980s, and the teenage unemployment rate also fell during that time. The teenage unemployment rate also often seems to increase around the time that the minimum wage increases. On the other hand, teenage unemployment also fell substantially during the 1960s at a time when there were several increases in the minimum wage.

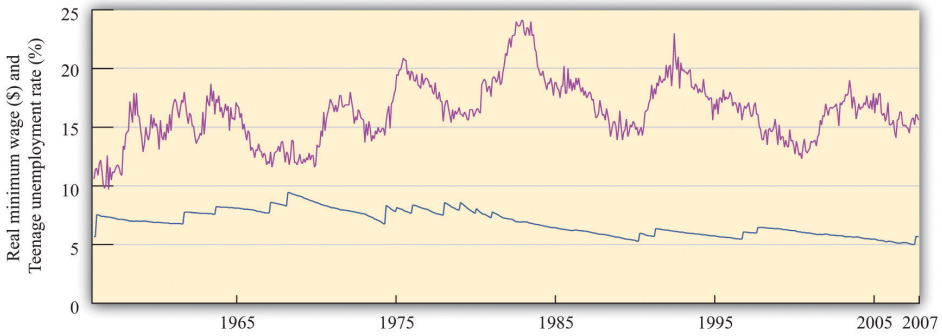


Figure 11.6.4: Teenage Unemployment in the United States. The figure shows the federal minimum wage in the United States, adjusted for inflation, together with the teenage unemployment rate.

Source: US Department of Labor and Bureau of Labor Statistics.

If the only cause of changes in teenage employment were minimum wage changes, Figure 11.6.4 might give us lots of answers. But this is a very big "if." In terms of our analysis of supply and demand, it would amount to saying that the supply and demand for labor didn't change over the entire period. In fact, you can make an enormously long list of things that might have shifted the supply curve, the demand curve, or both. Examples include whether the economy was in a boom or a recession, changes in tax rates, technological advances, and population growth. As we saw in Figure 11.6.3, it is difficult to disentangle the effects of a changing minimum wage from the effects of other changes. Thus we cannot isolate the effects of the minimum wage just by looking at diagrams like Figure 11.6.4.

We need some way to take into account these other factors so that we can focus on the effects of the minimum wage. This is a complicated statistical problem that arises time and again in economics, and the ways of dealing with it go far beyond this textbook. Indeed, this problem is one of the hardest things about studying economic data. But even though the statistical details are complex, there are three simple ideas that are important to understand:

1. The basic idea of these statistical techniques is that economists attempt to include in their analysis—as best they can—all the factors that shift the demand and supply curves. Then they can predict *what would have happened if there had been no change in the minimum wage*. Once they have done this, they can then determine the effect of minimum wage changes.
2. When different economists study the same problem, they do not always reach the same conclusion. The reason that they disagree is usually not because they have different ideas about economic theory. Almost all economists have the same general framework for understanding labor markets, for example. Economists disagree because they use different approaches to take into account all the other factors.
3. If economists could conduct direct experiments, such as those performed in the physical sciences, they would not have this problem. They would be able to completely control the *stimulus* (minimum wage changes), and thus get an accurate estimate of the *response* (changes in employment). Instead, the best we can do is to look at past minimum wage changes and try to determine what happened.

Studies from the 1970s and early 1980s found some evidence that increases in the minimum wage reduced the employment of 16- to 19-year-olds. According to economists David Card and Alan Krueger, the average estimate is that a 10 percent increase in the minimum wage would reduce employment by about 1.5 percent.David Card and Alan Krueger, *Myth and Measurement: The New Economics of the Minimum Wage* (Princeton, NJ: Princeton University Press, 1995). So the implied value of −(elasticity of labor demand) is about 0.15. With an average employment rate of about 50 percent, this means that a 10 percent increase in the minimum wage would reduce the rate of teenage employment by about 0.75 percentage points.

It is striking that this estimate is *lower* than the estimate from looking at labor demand as a whole. What is more, Card and Krueger note that more recent studies produce even smaller estimates of these effects.David Card and Alan Krueger, *Myth and Measurement: The New Economics of the Minimum Wage* (Princeton, NJ: Princeton University Press, 1995), Table 7.3.2 "Calculating Revenues". In their own work, they find that the apparent negative effects of the minimum wage on employment are statistically insignificant. This result holds not only for all 16- to 19-year-olds but also when the sample is split by race and sex.

The evidence from teenage employment studies suggests that, as an empirical matter, the effects of the minimum wage on employment may be very small. This result has surprised many economists. Although economic theory did not suggest actual magnitudes for the labor demand elasticity, the existence of substitutes for unskilled labor did suggest that labor demand would be at least somewhat elastic.

Not surprisingly, there have been many other studies of the minimum wage. Some researchers have found larger employment elasticities for teenage employment than those reported by Card and Krueger. Sometimes we are simply unable to give a definitive answer to empirical questions in economics. This can be frustrating for both students of economics and practitioners—but we are not going to pretend that the world is simpler than it actually is.

Cross-Section Studies

The studies we have just discussed analyzed the minimum wage by looking at minimum wage changes over time. Another approach to analyzing the effects of the minimum wage is to take advantage of differences over individuals rather than variations over time. Economists call these *cross-section studies* because they look at a cross section of different individuals or firms at a point in time. Many of these studies look at the effects of minimum wage changes at the level of an individual worker. Others exploit differences in minimum wage laws across states. Such differences across states give rise to a **natural experiment** because they can substitute, at least in part, for economists' inability to conduct experiments in which only one thing changes at a time.

Here is an example. Recall that in 1938 a minimum wage of $0.25 per hour was put into effect under the Fair Labor Standards Act. In the United States, this minimum wage was about 40 percent of the average manufacturing wage. However, the law also applied to Puerto Rico, which was much poorer, where this minimum wage was about twice the average factory rates. In a book published more than 25 years ago, researchers John Petersen and Charles Stewart noted that the increase in the minimum wage led to numerous factory closings, and there was a dramatic decrease in output and employment.John M. Peterson, and Charles T. Stewart Jr., *Employment Effects of Minimum Wage Rates* (Washington: American Enterprise Institute, 1969). In the case of Puerto Rico, the introduction of the minimum wage apparently had a large adverse effect on employment.

Peterson and Stewart also provide an extensive account of early studies of minimum wages that looked at employment and wages at individual production sites. These studies compared employment before and after a change in the minimum wage in an attempt to infer the effects of the policy. A study of the seamless hosiery industry from 1938 to 1941 is of particular interest. This period saw the introduction of the $0.25 per hour minimum wage, followed by an increase to $0.325 per hour in September 1939. A researcher named A. F. Hinrichs looked at 76 different plants and divided them into two groups: those that paid high wages and

those that paid low wages. We would expect the minimum wage to have a much bigger effect at the low-wage plants. Between September 1938 and September 1939, the low-wage plants had employment losses of 12 percent. Employment at high-wage plants actually expanded by 23 percent, perhaps in part because workers who lost jobs in the low-wage plants became part of the labor supply for the previously higher-wage plants. A similar pattern was noted for the period from 1938 to 1940.

A much more recent study by Card and Krueger is another example of this approach. They studied employment patterns in fast-food restaurants in New Jersey and Pennsylvania. The key to their research was that, during the period of study, the minimum wage was increased in New Jersey but not in Pennsylvania. (Remember that individual states sometimes set minimum wages above the federal minimum.) From this natural experiment, Card and Krueger found that the increased minimum wage in New Jersey actually seemed to have *increased* employment.

The evidence from other countries (both cross-section studies and studies over time) is likewise mixed. A recent Organisation for Economic Co-operation and Development report has a summary of minimum wage studies from different countries: "Making the Most of the Minimum: Statutory Minimum Wages, Employment and Poverty," accessed March 14, 2011, http://www.oecd.org/dataoecd/8/57/2080222.pdf. One study of Greek labor markets, for example, found a negative effect for men but a positive effect for women. Another study found negative effects for Mexico but not for Colombia. Different researchers in France have come to different conclusions about the effects of the minimum wage there; researchers in New Zealand likewise disagree; and so on.

Beyond Employment Effects

We started with what seemed to be some simple questions about the minimum wage. The answers turned out to be quite complex. Empirical research does not deliver a definitive answer about whether minimum wages have a big effect on employment. This leads to some disagreement among economists, particularly because the minimum wage is a politically charged issue. From the perspective of policymaking, the lack of a consensus creates difficulty in formulating good policy. On the other hand, the lack of a consensus provides a stimulus for continued work on these important issues.

Though we have emphasized the employment effects of minimum wage changes, there are other effects of minimum wages as well. First, remember that the main argument in favor of minimum wages is that they are a vehicle for redistributing income toward the working poor. Card and Krueger present a detailed analysis of the types of individuals most likely to be directly affected by minimum wage changes. Although empirical work often focuses on the employment of teenage workers, young workers are not the only group in the labor market that is paid close to the minimum wage. About 50 percent of the workers affected by the April 1990 increase in the minimum wage were older than 24 years old, for example.

How much income then flows to these workers as a consequence of an increase in the minimum wage? Card and Krueger conclude that the increase in the minimum wage during 1990 and 1991 had only a tiny effect on the distribution of income. They calculate that the minimum wage increase from $3.35 to $4.25 transferred about $5.5 billion of income to low-wage earners. This amounts to about 0.2 percent of family earnings. The host of transfer programs in place in the United States swamps the effects of the minimum wage on the redistribution of income. The evidence from other studies and countries is broadly in line with this conclusion: several studies find some effects of minimum wages on income distribution, but these effects are typically small.

Second, we can think about the effect of the minimum wage on firms. An increase in the minimum wage increases firms' **marginal cost** of production. As a consequence, firms will increase their prices and sell less output. Because of this, increases in the minimum wage reduce profits, so we might expect to see this reflected in the share prices of firms that employ minimum wage workers. Relative to the large empirical literature on employment effects, the implications for employers have been largely neglected. Card and Krueger survey the evidence and find relatively small effects on the stock market value of firms.

Key Takeaways

- Economists use data from labor market outcomes (wages and employment) to infer the shapes of labor supply and demand curves. A key part of this inference is to isolate economic variation to trace out one of the curves.
- Based on many studies, −(elasticity of labor demand) is about 0.3. So if we increase the minimum wage by 10 percent, employment will decrease by 3 percent.
- Studies that look directly at the effects of minimum wage changes find minimal effects of minimum wage changes on employment and unemployment.

check your understanding

1. Why is it so difficult to determine the effects of minimum wage changes on unemployment?
2. What is a natural experiment?
3. Why do changes in the real minimum wage allow researchers to trace out the labor demand curve?

This page titled 11.6: Empirical Evidence on Minimum Wages is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

11.7: End-of-Chapter Material 11.7: End-of-Chapter MaterialEdit Update titlecancel

In Conclusion

The minimum wage is a public policy that is debated the world over. It has widespread public support because there is something very appealing about the notion that those who work hard will be rewarded with a reasonable standard of living. Economists, living up to their reputation as dismal scientists, point out that this is all very well, but there may be unintended consequences of such a policy:

- A minimum wage leads to a reduction in employment and to unemployment.
- A minimum wage leads to fewer trades than in a competitive market and therefore to inefficiency.

The theory of the minimum wage is straightforward and convinces most economists that, even if the minimum wage has some benefits in terms of the distribution of income, it carries costs with it as well. There is less consensus among economists about whether the redistribution brought about by the minimum wage is desirable. Although the working poor benefit, others are made worse off—including those who are unemployed, who are perhaps even more in need of help than the working poor. Individuals differ in their beliefs about how society's resources should be distributed, and there is no right answer to the question "what is fair?"

When we looked at the *evidence* on minimum wages, however, we found that the picture is much less clear. Although some studies are in line with the predictions of theory, many studies suggest that, in practice, the effect of the minimum wage on employment is minimal. At the same time, the effect of the minimum wage on the distribution of income is small as well. In the end, it is difficult to resist concluding that the minimum wage is much less important—in terms of both benefits and costs—than one would think from the rhetoric of the debate.

When you read the newspapers or watch television, you will frequently hear economists offer different viewpoints on economic policies. These disagreements are typically not because economists differ on the theory. The disagreements often come down to different opinions about how to analyze and interpret economic data. Remember as well that television and print journalists go out of their way to find differing points of view because that makes a better story, so the disagreement you see in the media is usually not representative of economists as a whole. That said, economists also have different political viewpoints, and they are sometimes guilty of letting their political preferences cloud their economic analysis.

Having gone through the arguments in this chapter, you should be better able to assess debates and discussion on the minimum wage the next time it comes to the forefront of public debate. This chapter has a much broader purpose, however. We have been studying the effect of government intervention in a market, and we have shown how we can use our tools of supply and demand to understand the likely effects of that intervention. There are many other examples of government interventions that can be investigated using very similar reasoning.

Finally, we have learned something about how empirical work is conducted in economics. Because economists cannot conduct experiments, they are forced to trawl through messy data in an attempt to test their theories. It is difficult to be sure that the variables in which we are interested are indeed changing enough to be useful, and it is even more difficult to disentangle those changes from all the other irrelevant changes that affect the data that we observe.

Key Links

- Bureau of Labor Statistics: www.bls.gov
- Department of Labor: http://www.dol.gov
- US minimum wage laws: www.dol.gov/esa/minwage/america.htm

exercises

1. List three jobs you think probably pay minimum wage and three jobs that you think do not.
2. Illustrate an increase in the minimum wage when both demand and supply are (a) relatively inelastic and (b) relatively elastic. Explain why the change in unemployment is smaller when the curves are inelastic.
3. Explain why the deadweight loss from the minimum wage is larger if labor demand is relatively elastic.
4. How does the elasticity of labor supply affect the deadweight loss from the minimum wage? Specifically, if labor supply is more elastic, is the deadweight loss smaller or larger? What is the economic intuition behind your answer?

5. (Advanced) Draw a version of Figure 11.7.4 for the case where a single individual controls access to scarce jobs. Suppose that she is able to charge job searchers a fee (the same fee for all searchers) equal to the difference between the minimum wage and the wage that workers would be willing to accept. What area of the figure does she obtain?
6. In our discussion of the evidence of the effects of minimum wage changes, we said, “If economists could conduct direct experiments, such as those performed in the physical sciences, they would not have this problem.” Exactly what problem were we referring to?
7. What is the difference between cross-sectional and time-series studies? Does one hold “more things fixed” than the other?
8. Suppose the government imposed a maximum wage in the market for some high-paying job. Draw a diagram to illustrate this market. What would be the consequences of this maximum wage?
9. Explain why, when we analyze the minimum wage, the elasticity of labor supply affects the unemployment rate but not the employment rate.
10. Why does the government not set a minimum wage for corporate lawyers and airline pilots?
11. If the rate of inflation is 10 percent higher than expected, and –(elasticity of labor demand) is 5, what will happen to the employment level in jobs that pay the minimum wage?
12. What happens to the rate of unemployment of minimum wage workers if the rate of inflation is lower than expected?
13. Does the elasticity of labor supply have an effect on the change in the wage bill when there is an increase in the minimum wage? Does this elasticity have an effect on the unemployment rate when the minimum wage changes?
14. (Advanced) Using the discussion of estimation of labor demand, if you could conduct an experiment to see the effects of a minimum wage increase, what exactly would you do?
15. (Advanced) Using supply and demand curves in the market for fast food, what are the effects of an increase in the minimum wage in this market? Think about shifts in both the supply curve and the demand curve. Explain your predictions.
16. Why isn’t an increase in the minimum wage just a redistribution from firms to workers?
17. A politician is in favor of getting rid of the minimum wage entirely. How would you argue against that proposal?

Economics Detective

1. Pick three countries and find the minimum wage in each country.
2. Find a country that does not have a minimum wage. Do you think the lack of a minimum wage means that workers are badly treated in that country?
3. Find some recent discussion of minimum wage legislation in either the United States or some other country. What arguments were made to support the minimum wage? What arguments were made against the minimum wage?

Spreadsheet Exercise

1. (Advanced) Find data on the minimum wage and the price level for another country. Construct a real minimum wage series for that country.

This page titled 11.7: End-of-Chapter Material is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

12: Barriers to Trade and the Underground Economy

12.1: An Unusual Shopping List

12.2: How the Government Controls What You Buy and Sell

12.3: Limits on Trade across Borders

12.4: Government and the Labor Market

12.5: End-of-Chapter Material

This page titled 12: Barriers to Trade and the Underground Economy is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

12.1: An Unusual Shopping List 12.1: An Unusual Shopping ListEdit Update titlecancel

You come downstairs one morning and find a note on the table.

Please go to the store today and buy the following:

- A bag of sugar
- Two pints of milk

So far there is nothing unusual about this. You plan to go to the grocery store on your way home that evening. Then you read on.

- One carton of cigarettes
- One bottle of whiskey

These are a bit trickier. If you are like many readers of this book, you may not be allowed to purchase alcohol or possibly even cigarettes. In the United States, you must be 21 or over to buy alcohol and over 18 (or 19 in some states) to purchase cigarettes. Depending on where you live, it may also be quite inconvenient to purchase alcohol. In some places, by law, alcohol is sold only at certain times of day. In some places—certain states in the United States and certain countries in Europe, for example—it is sold only in government-run stores.

Many goods, like alcohol, are restricted in terms of who can buy them, when they can be purchased, and where they can be purchased. Alcohol laws differ from country to country. In most European countries, for example, you can buy alcohol at the age of 18. The laws also change over time. Thirty years ago, 18-year-olds could buy alcohol in the United States as well. Ninety years ago, it was illegal for anyone to buy alcohol in the United States.

Next on the list is the following.

- Two tickets for the sold-out rock concert in town tomorrow night

This may also be difficult. You know that you can probably find someone who has tickets and is willing to sell them, but you know that local laws say that this, too, is illegal. So-called scalping of tickets is forbidden. Still, if you go to eBay, you'll probably be able to find some tickets for sale.

Then the list gets stranger:

- Six Cohiba cigars
- One French raw milk camembert
- Four ounces of marijuana

At this point (at least if you are living in the United States), you begin to seriously worry. You search the Internet for "Cohiba" and discover that these cigars are manufactured in Cuba, but you vaguely remember that it is illegal to import goods from Cuba to the United States. You know that camembert is a French cheese, but "raw milk" sounds strange. More online investigation informs you that it is also illegal to import cheeses into the United States unless they are made from pasteurized milk. Apparently, raw milk cheeses may carry dangerous bacteria. As for marijuana, you already know that it is illegal in the United States.

You read on.

- Also, please hire a cleaning person (an undocumented migrant worker would probably be the cheapest)

This is another transaction that you know is illegal. That said, you know that there are many illegal immigrants working in your town. It would be easy to find someone to hire if you were willing to break the law. With some foreboding, you turn the list over and read the other side.

- Finally, please buy one human kidney (suitable for transplant).

Most of the things that were on the list up to this point were goods or services that you would probably be able to find if you had to. Even though some of them could not be purchased legally, it would not be too hard to find out where to purchase most of them. (Oddly, it would probably be easier to get the marijuana than the cheese.) A human kidney is a different proposition, however. You're pretty sure, even without research, that buying and selling human organs is illegal, and you would have no idea where to go to buy a kidney even if you were willing to break the law.

We know that the market interaction of buyers and sellers creates value in an economy.We discuss this in detail in Chapter 6, and Chapter 8. In a market, sellers supply a good or a service, and buyers demand that good or service. Because each transaction is voluntary, the value that the buyer places on the good is always greater than its value to the seller. This means that each trade creates some value. In addition, if the market is competitive, all value-creating trades occur in the market; there are no disappointed buyers or sellers.

This logic suggests that governments should be doing everything in their power to encourage and facilitate trade. Yet, in practice, there are several ways in which governments do the opposite: they actively intervene to restrict trade. We have just listed a large number of examples, and you can surely think of many more. We would like to understand all the restrictions that are deliberately put in place to impede trade.

Our main aim here is *not* to analyze the rationales behind these restrictions, although we do briefly explain some of them. In other chapters, we provide more insight into precisely why governments impose these and other limitations on our ability to transact with one another.See in particular Chapter 13, and Chapter 15. Our goal in this chapter is to explore *what happens* when governments interfere with trade in different ways.

One message of this chapter is a reiteration of the gains from trade, together with the recognition that they provide a powerful incentive for people to get together and transact with one another. It seems that whenever the government steps in to try to prevent them from trading, people still try to find a way around these restrictions. The gains from trade are a powerful motivator. Indeed, people continue to trade even when this is an illegal act that carries a significant risk of fines or imprisonment. We use the term *underground economy* to describe where these trades occur. The question we want to answer in this chapter is as follows:

What are the consequences of government restrictions on trade?

Road Map

In this chapter, we will see many different ways in which governments intervene. For most of our analysis, we use the supply-and-demand framework. We analyze different kinds of government policy and examine the following questions:

- What happens to prices and quantities?
- What happens to welfare and the distribution of income?
- What happens to incentives? Are there any resulting unintended consequences?

We organize our discussion by looking at different categories of restrictions on trade. First, we look at the sale of goods and services in domestic markets. Then we turn to restrictions in international markets for goods and services. Finally, we turn to restrictions not on goods and services but on labor, both within and across countries.

This page titled 12.1: An Unusual Shopping List is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

12.2: How the Government Controls What You Buy and Sell 12.2: How the Government Controls What You Buy and SellEdit Update titlecancel

Learning Objectives

1. What tools does the government use to control market transactions?
2. Why might the government restrict trades?
3. What are the effects of these restrictions on market outcomes and welfare?

Looking at your shopping list, there are some items that you simply cannot buy. For example, marijuana and the raw milk cheese from France are not available for purchase in stores in the United States. And depending on your age as well as the time and day of the week, you may not be able to buy the cigarettes and whiskey. We begin by discussing these types of market interventions.

Closing Down the Market

The most fundamental intervention in a market occurs when the government closes down trading completely—that is, the government simply says that it is illegal to trade certain goods or services. Examples are numerous and stem from many different motivations.

Health and safety. Most governments ban addictive drugs, such as heroin, cocaine, and marijuana. The primary reason is that these and similar drugs are deemed to be harmful to those who use them. A secondary reason is that governments may think—rightly or wrongly—that the trade of such drugs also has other harmful implications, such as increased crime.

Governments also ban trade in other products for similar reasons of health and safety. One of the functions of government in most countries is to oversee the safety of products, both generally and more specifically in terms of health risks. In the United States, the Food and Drug Administration certifies factories and food processing and also oversees the approval of pharmaceuticals. Meanwhile, the Bureau of Consumer Protection (http://www.ftc.gov/bcp) is charged with ensuring that goods meet certain legislated safety standards. Goods that do not meet these standards cannot be legally traded. For example, it is illegal to sell a new car without seatbelts and airbags in the United States.

Ethics, morality, and religion. The exchange of some goods and services is banned for ethical or moral reasons. Examples include the trading of human organisms, the sale of alcohol, and various forms of prostitution.

The ban on the trading of human organs is rooted primarily in an ethical belief that buying and selling body parts is wrong. Many people argue that the moral case for banning organ selling is very shaky, and the world would be a better place if such trades were allowed. It is true that many find the idea of trading body parts for dollars to be repulsive. We have a sense that people would sell a kidney only if they were truly in desperate financial straits, and there is something terrible about the image of, say, a mother selling a kidney to feed her children. Yet there are people who die every day because doctors are unable to find a suitable organ donor in time; that, too, is a sad image.

In many places, the consumption and sale of alcoholic beverages is forbidden, often for religious reasons. The sale of alcohol is prohibited in some Muslim countries, such as Saudi Arabia and Kuwait. Religious pressure also led to a 13-year ban on alcohol in the United States under the 18th Amendment to the Constitution; this state of affairs was known as *Prohibition*. Indeed, in many counties in the United States, the sale of alcohol is still prohibited. Likewise, many other countries in the world have regions that are “dry.”

Not surprisingly, there is often disagreement about which trades should be ruled immoral or unethical. Different laws in different countries regarding the sale of alcohol are one illustration of this. Another example is prostitution, which is illegal in many places yet legal in others. For example, prostitution is legal (although heavily regulated by the government) in the Netherlands and in parts of Nevada.

Fairness. Sometimes, the government simply takes the view that certain trades are unfair. For example, scalping—the reselling of tickets to concerts and exhibitions—is frequently prohibited for this reason. The following story illustrates that people often see the reselling of tickets as unfair.

Perry Loesberg wanted to surprise his 10-year-old daughter Amy with tickets to “Hannah Montana,” the sizzling-hot concert tour featuring 14-year-old TV star Miley Cyrus.

Instead, he was the one surprised. Though he bought a $30 fan club membership to get access to tickets ahead of the general public, and then logged on to the Ticketmaster Web site before the general public sale began, Loesberg still came up empty-handed.…

Tickets to each of the 54 shows on the "The Best of Both Worlds: Hannah Montana and Miley Cyrus" tour…sold out within minutes of going on sale. Almost immediately, online marketplaces such as StubHub and craigslist were offering dozens of seats, many selling for more than $2,000 each. Tickets were originally priced at $22–$66.

What kind of ignited parents is I think they thought it should be more fair," said Debra Rathwell, senior vice president for AEG Live, the tour's promoter…"We would like fans to sit in these seats. But everything you do, [scalpers] find a way to skirt around it."

Many of the purchasers are parents unfamiliar with the post-Internet ticket market. They were amazed at the availability of tickets —not to mention the high prices—on the re-sale market.…

Many have pointed to computer software programs that allow users to, in essence, cut in line on the Ticketmaster Web site.…The outcry over Hannah Montana is unusual for other reasons. Ray Waddell, senior editor at Billboard magazine, said parents and children are being disappointed, and their complaints have found sympathetic ears, including the attorneys general of Missouri, Arkansas, Connecticut and Pennsylvania…

It has been like the Wild West out there," said Waddell. "Things are going to tighten up, (there will be) more regulation about who's selling, who's buying and how they are getting their tickets."Peggy McClone, "Parents Are Angry about 'Hannah Montana' Ticket Sales," *The Star Ledger*, October 29, 2007, accessed January 29, 2011, http://www.nj.com/news/index.ssf/2007/10/parents_are_upset_about_hannah.html.

Restrictions on Who Can Trade

There are many products that can be legally traded, but the government places substantial conditions on the terms of those trades. For example, several legal goods and services cannot be purchased by minors and can be sold only by licensed sellers. Obvious examples are alcohol and cigarettes, but there are many others. Casino gambling is restricted to adults. Many pharmaceuticals can be sold only by licensed pharmacists and bought only with a doctor's prescription.

These restrictions vary a lot by time and place, which again tells us that there is no simple right or wrong where these laws are concerned. Different states have different laws. Not all stores can sell liquor. In Sweden, for example, alcohol is sold only in state-run stores. The legal drinking age in Europe is different from the legal drinking age in the United States. Some drugs require a prescription in some countries yet are available over-the-counter in others.

Implications

Figure 12.2.1 shows the **buyer surplus** and **seller surplus** in a **competitive market** and reminds us that the **gains from trade** in a competitive market are at a maximum. All mutually beneficial trades have been carried out. Government interventions in markets typically have the effect of eliminating some or all these gains from trade.

Toolkit: Section 31.10

You can review the different kinds of surplus and the gains from trade in the toolkit.

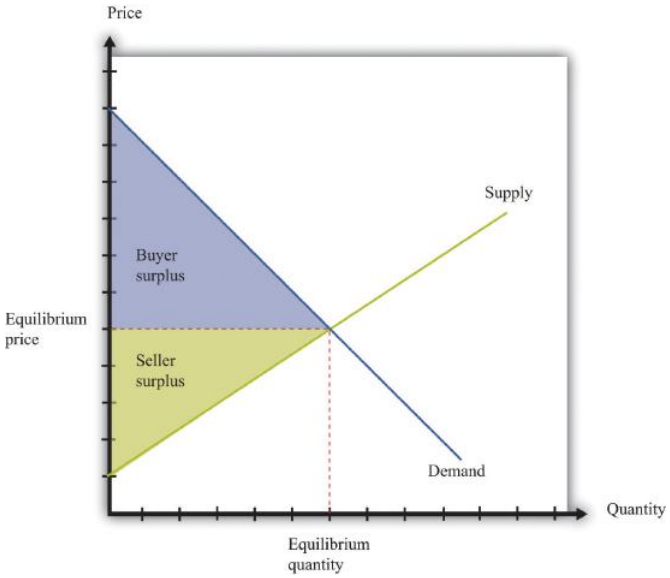


Figure 12.2.1: Supply, Demand, and the Gains from Trade. The area below the demand curve and above the price is the buyer surplus; the area above the supply curve and below the price is the seller surplus.

The economic analysis of the closing of a market is very simple. If the government successfully prevents trade, then the quantity traded is zero. All producer and consumer surpluses in Figure 12.2.1 are lost. An economist's first response to the closing of a market—any market—is that it brings a loss because some potential gains from trade go unrealized. The question then becomes whether any benefits from closing down a market justify the lost gains from trade.

As our examples reveal, there are many reasons for closing a market, so there is no simple answer to the question, "Is it good to shut down a market?" Each argument must be looked at on a case-by-case basis, and the particulars of specific examples are beyond the scope of this book. Entire books have been written, for example, on the market for human organs or the legalization of prostitution.

When you read the shopping list at the beginning of the chapter, you might also have been struck by the fact that the government's success in blocking trade is often limited. You probably would find it difficult to buy a heart for transplant on the open market. But if you know where to go, you could almost certainly buy marijuana. Even if you are underage, you may be able to get a fake identification card and buy alcohol. And buying scalped tickets to a concert or a sports event is usually easy, if you have the money. The economic message is simple and fundamental. When there are gains from trade, people will try to realize those gains. When trades are illegal, economic activity moves into the so-called underground economy but is unlikely to disappear completely.

Rationing

Another way in which governments intervene in markets is not by banning trade outright but by placing a restriction on the quantity traded. In most modern economies, such restrictions are little used in a domestic context but are much more prevalent in international trade. If we look back in history, though, we can find instances of **rationing** in the domestic economy. Rationing means that the quantity available on the market is less than the equilibrium quantity. Some surplus goes unrealized because willing buyers and sellers are prevented from trading. During and after World War II, many basic goods were rationed in the United States, Britain, and elsewhere.

The following excerpt by journalist Joelle Kirch Preksta comes from oral histories of World War II collected by the Carnegie Library.

Ruth showed me several of the ration books she was issued during World War II.…She explained that staples such as sugar, butter, and eggs were rationed in order to help supply our troops overseas and therefore were difficult to obtain in stores.…The following excerpt [is] taken from the "Instructions" section of the books…

1. This book is valuable. Do not lose it.
2. Each stamp authorizes you to purchase rationed goods in the quantities and at the times designated by the Office of Price Administration. Without the stamps, you will be unable to purchase those goods.

Rationing is a vital part of your country's war effort.…Any attempt to violate the rules is an effort to deny someone his share and will create hardship and discontent. Such action, like treason, helps the enemy. Give your whole support to rationing and thereby conserve our vital goods. Be guided by the rule: "If you don't need it, DON'T BUY IT."

The books also contained a warning which indicated that someone who violated the rules for the ration books could be imprisoned for as long as 10 years or fined as much as $10,000.Ruth L. Baxter, interview by Joelle Kirch Preksta, May 21, 2001, Carnegie Library of Pittsburgh, www.carnegielibrary.org/research/pittsburgh/history/ww2/ww27.html.

Despite these strong moral and legal sanctions—comparing black market trading with treason, no less—there was a substantial underground market for all sorts of rationed goods. For example, the Carnegie oral histories describe a young woman in her twenties named Mary: "She somewhat embarrassingly recalled that she was able to dishonestly procure an extra carton of cigarettes every month for herself because her aunt worked at the drug store where they could be purchased. To this day she says she feels somewhat guilty over this unpatriotic indiscretion."Mary Hresko and Mary Vincher Shiner, interview by Mark Kernion, May 21, 2001, Carnegie Library of Pittsburgh, www.carnegielibrary.org/research/pittsburgh/history/ww2/ww29.html.

Figure 12.2.2 shows the implications of quantity rationing. Part (a) of Figure 12.2.2 shows that there is a deadweight loss. We see that a quantity ration does not tell us what the price will be. It could be anywhere between the minimum price that the marginal seller will accept (the price found on the supply curve) and the maximum price that the marginal buyer will pay (the price found on the demand curve). In the absence of any other mechanism, the price is determined by bargaining among buyers and sellers. In the case of World War II rationing, sellers were often in stronger bargaining positions, which pushed the price toward the higher end of the range. For this reason, quantity rations were often supplemented by a maximum price, called a **price ceiling** (part (b) of Figure 12.2.2). Figure 12.2.3 shows a poster from this period.

Toolkit: Section 31.11

You can review the concepts of efficiency and deadweight loss in the toolkit.

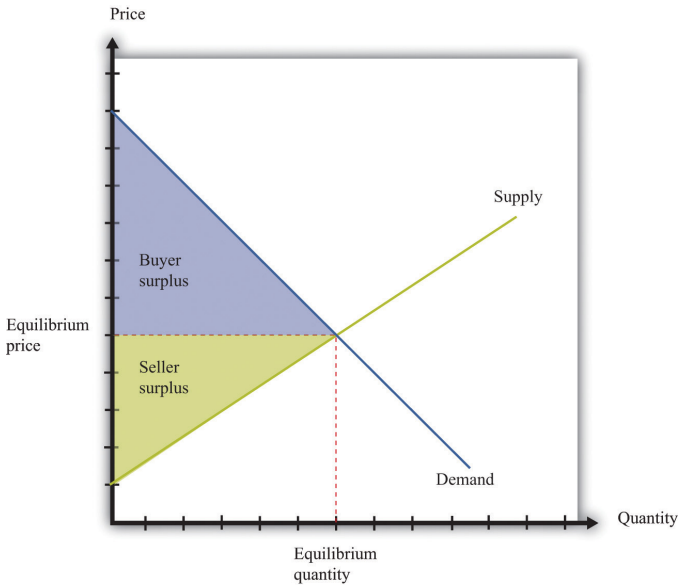


Figure 12.2.2: The Implications of Quantity Rationing. A quantity ration leads to deadweight loss but by itself does not tell us what the price will be.



Figure 12.2.3: A World War II Poster. Price ceilings during World War II led to illegal trading above the fixed price, so the government campaigned to prevent people from trading in these markets.

Hulton Archive/Getty Images

Price Ceiling

It is more common for governments to intervene in an economy by using price tools rather than quantity tools. In particular, governments sometimes intervene using restrictions on how high the price in a market can go. This is called a *price ceiling*. A classic example of a price ceiling is rent control. In New York City and some other places, there are restrictions on how much landlords can increase the rent on apartments.

Figure 12.2.4 illustrates a price ceiling. Notice that (unless there is also a quantity ration in place) the price ceiling must be below the equilibrium price; otherwise the policy is irrelevant. The main economic implications of a price ceiling can be readily seen from this figure.

- Because no one can force you to sell if you don't want to, the quantity traded is determined by the supply curve.
- Because the quantity traded is below the equilibrium quantity, there is an inefficiency (deadweight loss).
- Because the quantity demanders wish to buy exceeds the quantity suppliers wish to sell, there must be some kind of rationing in the market to determine who actually buys the good or the service in question.

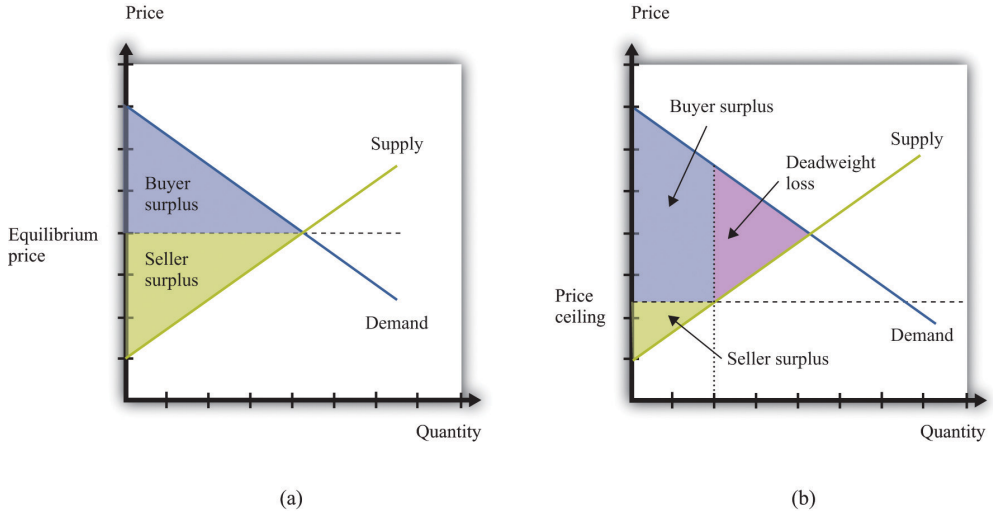


Figure 12.2.4: The Effects of a Price Ceiling. With no price ceiling (a), all the possible gains from trade in the market are realized. With a price ceiling (b), some gains from trade are lost because there are fewer transactions.

Rent controls keep the price of an apartment rental below its equilibrium level. Not surprisingly, lots of people would like to live in rent-controlled apartments. The quantity demanded is greater than the quantity supplied. Because the price is, by law, not allowed to undergo the adjustment that would restore equilibrium in the market, some other kind of rationing must take place instead.

Rent controls are enacted with distributional goals in mind. The aim is to ensure that people with lower incomes are not priced out of the rental market. Put differently, the goal is to redistribute income from sellers to buyers—that is, from landlords to those who are renting apartments. A difficulty with price ceilings is that people have an incentive to try to get around the restrictions in creative ways. There is often more to a transaction than a simple exchange of money for a good or a service. There may be nonmonetary aspects of the transaction that governments find harder to regulate. When apartments are covered by rent controls,

landlords often ask for “key money.” This is an off-the-books, up-front payment that renters must agree to pay before renting the apartment. In other words, it is a polite term for a bribe. In addition, some landlords may not put much money or effort into the upkeep of rent-controlled apartments, thus compensating for the low rent by reducing the quality of the apartment.

In emergency circumstances, temporary price ceilings may be put into effect. These take the form of laws that prevent so-called *price gouging*. For example, in the aftermath of a hurricane, some goods and services are typically very hard to come by. Basic necessities like food and water may be in limited supply. In the weeks and months after such a disaster, building supplies and similar products may be almost completely unavailable.

After Hurricane Katrina, price-gouging laws applied to states affected by the storm.

While there is no federal price gouging law, many states have enacted some type of prohibition or limitation on price increases during declared emergencies. All of the affected states—Louisiana, Mississippi, Alabama, and Florida—have price gouging laws that are triggered by the declaration of an emergency in the state. Generally, the laws prohibit the sale of goods and services in the designated emergency area at prices that exceed the prices ordinarily charged…

However, there exists a general exemption for increased prices that are the result of additional costs incurred for procuring the goods or services in question.

…

In Alabama,…evidence of unconscionable pricing exists “if any person, during a state of emergency declared pursuant to the powers granted to the Governor, charges a price that exceeds, by an amount equal to or in excess of 25% the average price at which the same or similar commodity or rental facility was obtainable in the affected area during the last 30 days immediately prior to the declared state of emergency.”Angie A. Welborn and Aaron M. Flynn, “Price Increases in the Aftermath of Hurricane Katrina: Authority to Limit Price Gouging,” *Congressional Research Service Report for Congress*, September 2, 2005, accessed January 29, 2011, www.fas.org/sgp/crs/misc/RS22236.pdf.

Think about the market for lumber (wood for building purposes) in the first few weeks following a hurricane. Were we to apply supply and demand reasoning to this situation, we would get a diagram like Figure 12.2.5. Because there is a great deal of new construction going on, there is a rightward shift in the demand for lumber. The supply of lumber is likely to be fairly inelastic, at least until it is possible to start bringing supplies in from other states. Thus the shift in the demand will lead to a large increase in the existing price. If the price is allowed to increase to its new equilibrium, existing suppliers will obtain a big gain. Price-gouging laws, however, prevent suppliers from raising their prices in this way.

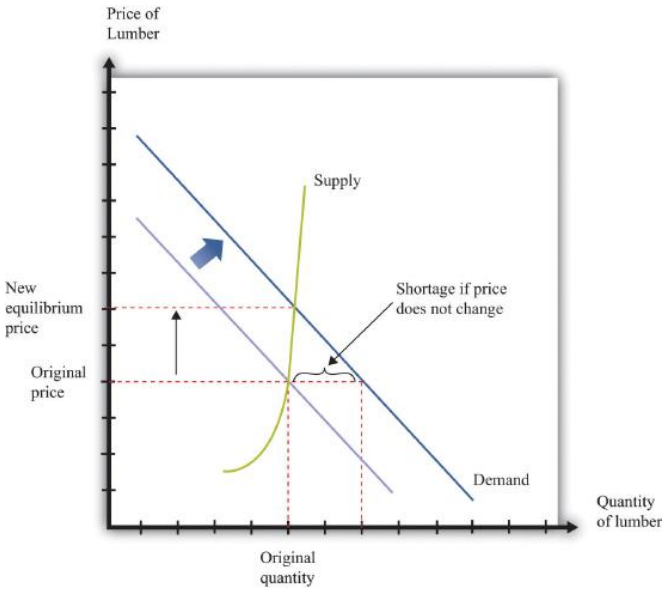


Figure 12.2.5: The Market for Lumber after a Hurricane. If the market were allowed to work, the price of lumber would increase substantially, but there would not be much more wood supplied. If suppliers are not allowed to increase prices, then demand exceeds supply.

This presents two problems. First, suppliers no longer receive the price signal that tells them to bring more wood to market. In the short run, this may not matter so much. After all, Figure 12.2.5 shows that, with inelastic supply, the shift in the demand curve

would not in fact lead to a big increase in the quantity supplied, even if the price were allowed to adjust. In the longer run, though, this is more of a problem because there is less incentive for suppliers from further away to bring in additional lumber.

The second problem with forcing sellers to keep their price fixed is that the increase in demand will lead to a shortage. This is also shown in Figure 12.2.5. Because demand now outstrips supply, the limited supply will have to be rationed in some way. Most likely, what will happen is that demanders will have to queue to get the lumber that they need. The time that they must spend standing in line has an **opportunity cost**; they would rather spend that time doing something else. We can think of the time spent in line as increasing the effective price that they have to pay.

These arguments do not necessarily mean that price-gouging laws have no merit. In the aftermath of a hurricane, many things may be happening. Lumber firms may see a temporary increase in their market power. Such an increase in market power gives them an incentive to increase prices, so price-gouging laws may serve as a way to limit the abuse of monopoly power.

Price Floor

A **price floor** is closely analogous to a price ceiling. The difference, as the name suggests, is that it is a government-imposed *minimum* price rather than a government-imposed maximum price. The government says that all transactions must be at or above this minimum price. The minimum wage is the most important example of a price floor.We devote a whole chapter to the analysis in Chapter 11. With the minimum wage, the aim is to redistribute income from buyers to sellers—that is, from firms to suppliers of unskilled labor.

Figure 12.2.6 illustrates a price floor. The main economic implications of a price floor can be seen from this figure.

- Because no one can force you to sell if you don't want to, the quantity traded is determined by the demand curve.
- Because the quantity traded is below the equilibrium quantity, there is inefficiency (deadweight loss).
- Because the quantity suppliers wish to sell exceeds the quantity demanders wish to buy, there must be some kind of rationing in the market to determine who actually sells the good or the service in question.

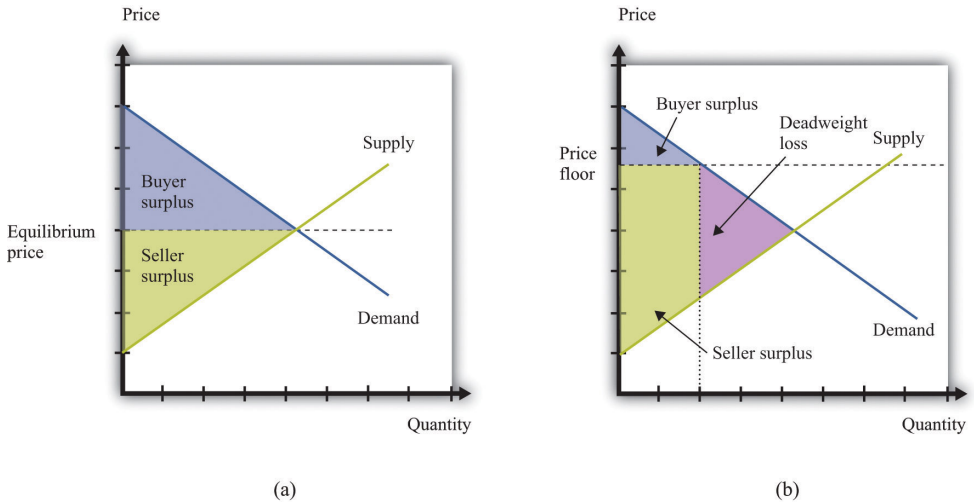


Figure 12.2.6: The Implications of a Price Floor. With no price floor (a), all the possible gains from trade in the market are realized. With a price floor (b), some gains from trade are lost because there are fewer transactions.

Just as renters use key money and other devices to get around rent control, firms (and workers) sometimes devise ways to get around minimum wage requirements. Employers who are forced to pay a minimum wage may provide worse working conditions than those who pay a market wage. Or, if you want to work at a company and are willing to work at less than the minimum wage, you can negotiate a deal with your employer so that you are paid the minimum wage for reported hours but then work additional hours for nothing. The minimum wage regulations in the United States stipulate that this is illegal, punishable with fines of $1,100 per violation.US Department of Labor, "Wages: Minimum Wage," accessed March 14, 2011, http://www.dol.gov/dol/topic/wages/minimumwage.htm.

Sometimes individuals work their way around such restrictions even more blatantly. In the former Soviet Union, price ceilings were put in place in an attempt to keep the prices of basic goods down for households. Martin Walker, a journalist in Moscow, wrote of his experiences with these price ceilings in the food markets outside Moscow.Martin Walker, *Guardian*. A butcher offering to sell

Walker a side of beef assured him that the price per kilogram was fixed. “However,” said the butcher, “the weight is subject to negotiation.”

In the Soviet Union, the limited supply of goods led to long lines for those who wanted to purchase basic commodities, such as bread. You can think of these lines as an additional component of the price: you pay money *plus* the value of the time that you spend standing in line.

Although price ceilings and price floors have different implications for the price in the market, they both imply that *the quantity traded in the market will be less than the equilibrium quantity*. The reason is simple: neither buyers nor sellers can be forced to trade if they do not want to. If the price is above the equilibrium price, the quantity is determined by the amount of the good or the service that people are willing to buy. Some would-be sellers are disappointed because they cannot find someone to buy from them. With a minimum wage, for example, not everyone who wants a job can find one. If the price is below the equilibrium price, the quantity is determined by the amount of the good or the service that people are willing to sell. Some would-be buyers are disappointed because they cannot find someone to sell to them. With rent control, for example, not everyone who wants a cheap apartment can find one.

Taxes and Subsidies

Price floors, price ceilings, and quantity restrictions are important but relatively rare policies. The government intervenes regularly in almost every market in the economy in a different way—by the imposition of taxes. Had you purchased the milk and sugar on our shopping list, for example, you would very likely have paid a sales tax. Sometimes cities levy their own sales taxes as well. On certain goods, such as alcohol or gasoline, you may pay additional taxes.

Taxes

A **tax** is a payment made to the government that is associated with an economic transaction. Although the details of taxes can differ substantially, most taxes come down to one simple point: the price paid by the buyer is higher than the price received by the seller.

Suppose you want to purchase a book for its list price, say, $20. In the United States, if you take this book to the cash register, you will typically be charged a sales tax. If the sales tax is 5 percent, you will have to pay $21 for the book. The store collects the $1 tax on behalf of the government. So who is paying this tax? On the one hand, the amount of the tax is marked right there on the receipt as an amount you have to pay. Yet it is the store that actually sends the money to the government.

Imagine, by contrast, that you had to give the bookstore only $20 but then were personally responsible for sending the sales tax to the government. You would have to file a sales tax declaration each year for every item you bought. That would be both inconvenient and difficult for the government to monitor; for this reason, sales taxes are funneled through the seller. But we are interested in a more fundamental question: would this make a difference on who pays the tax? The answer is no. You would still pay $21, the government would still get $1, and the bookstore would still get $20.

In other words, *it does not make any difference whether the tax is imposed on buyers or sellers*. This is one of the most surprising results that economics teaches us. In our book example, the conclusion may seem obvious. Yet people often to fail to appreciate the far-reaching significance of this insight.

For example, social security taxes in most countries are imposed on both workers and employers. Suppose the government changed its policy and declared that the portion of social security that was previously paid by the employer now had to be paid by the worker instead. Looking at this as employed workers, we might think that we had just been hit with a huge tax increase. Indeed, if nothing else changed, the policy change would make workers worse off. Fortunately, the logic of supply and demand would quickly come to our rescue. At existing wages, firms would no longer be able to hire all the workers they wanted. Wages would be bid up, and before long we would expect to see workers and firms no better and no worse off than they were previously.

Who Pays the Tax?

The key question, then, is not who sends the money to the government. The key question is, *What happens to prices when a tax is imposed?*

To answer this, imagine that the government increases taxes on gasoline by 50 cents a gallon and consider two extreme cases. First, suppose the price of gas increases by 50 cents a gallon. Households are evidently paying the tax; the amount they must pay per gallon has gone up by the full amount of the tax. Now suppose that the price of gasoline at the pump does not change at all. Then firms are paying the tax: they are receiving 50 cents less per gallon once they pay the tax to the government. Most often, we expect

to see the price of gasoline increase but by less than 50 cents. Therefore, the burden of the tax is shared between the gas station and the household. It is the change in the price that tells us who *really* pays the tax.

Figure 12.2.7, and Figure 12.2.8 illustrate the effects of a tax.

- The gap between the buyer's price and the seller's price means that the quantity sold is less than the market equilibrium quantity (Figure 12.2.7).
- There is a deadweight loss: some mutually beneficial trades go unrealized. This is again visible in Figure 12.2.7. There are potential trades where the buyer's valuation exceeds the seller's valuation. However, because the difference in valuations is less than the amount of the tax, these trades are not worthwhile once the tax must be paid.
- There is a reduction in both the buyer surplus and the seller surplus, as can be seen in Figure 12.2.8. The buyer surplus is the area under the demand curve and above the price paid. The seller surplus is the area above the supply curve and below the price received.
- Figure 12.2.8 also shows that some of the surplus generated by these trades now goes to the government in the form of tax revenues. Government tax revenues equal the amount of the tax multiplied by the quantity traded. Graphically, they are equal to the rectangle shown in part (b) of Figure 12.2.8.

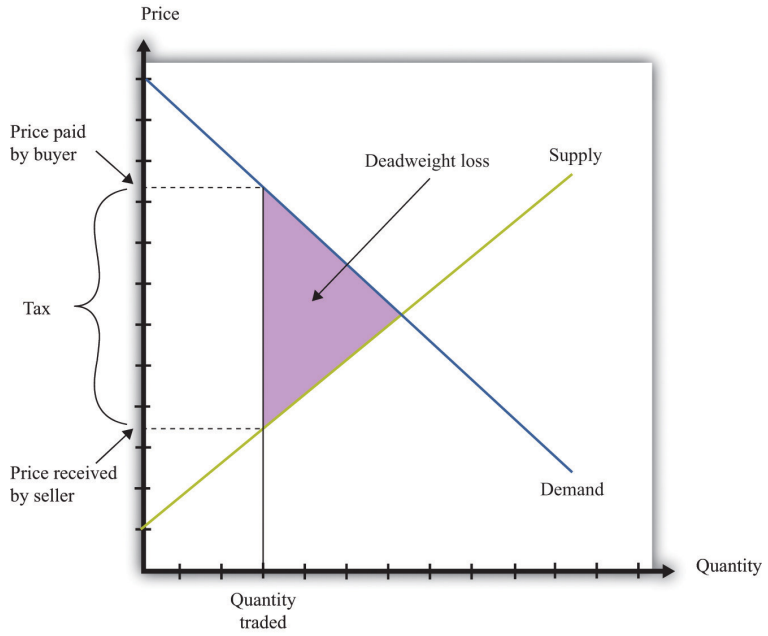


Figure 12.2.7: The Deadweight Loss from a Tax. A tax means that there is a wedge between the price paid by the buyer and the price received by the seller.

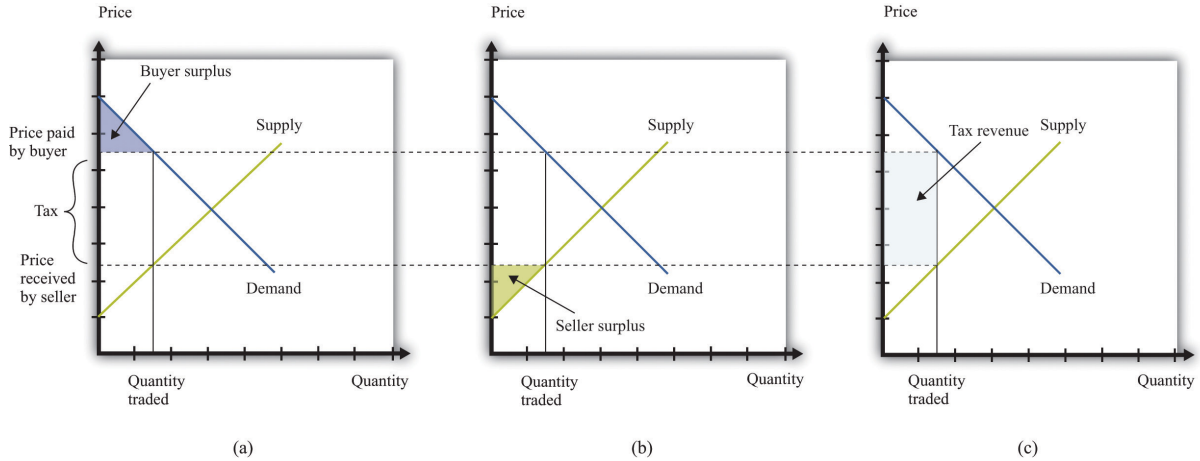


Figure 12.2.8: The Loss in the Buyer Surplus and the Seller Surplus from a Tax. The total surplus is the sum of the buyer surplus (a), the seller surplus (b), and the tax revenue received by the government (c).

Tax incidence is the way in which the burden of a tax is divided between buyers and sellers. In general, the incidence of a tax depends on the **price elasticity of supply** and the **price elasticity of demand**. Figure 12.2.9 shows why tax incidence depends on the elasticity of demand. That figure has two parts. In both parts, we start from the same initial competitive equilibrium and impose a tax of the same size. This means that the gap between the price paid by buyers and the price received by sellers is identical.

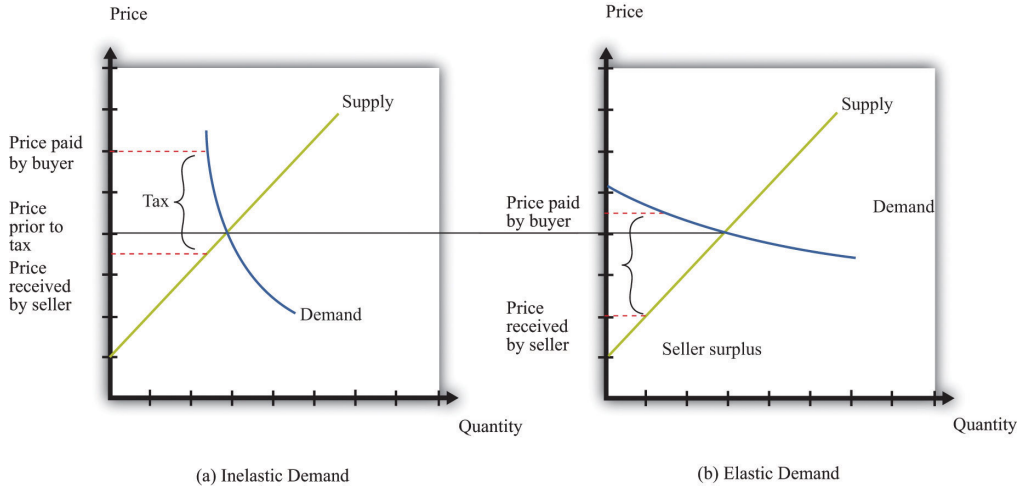


Figure 12.2.9: Tax Incidence with Inelastic and Elastic Demand. When demand is inelastic (a), most of the burden of the tax is borne by buyers, while the opposite is true when demand is elastic (b).

In part (a) of Figure 12.2.9, demand is inelastic. Buyers are not very price sensitive, so even if the price increases, their quantity demanded does not change a great deal. The result is that the price paid by buyers increases a lot. Most of the burden of the tax is borne by buyers. In part (b) of Figure 12.2.9, demand is elastic. As the price increases, the quantity demanded decreases a great deal. In this case, the price paid by buyers increases much less, and the price received by sellers decreases by more. Most of the burden of the tax is borne by sellers.

Keep in mind also that the distortion induced by the tax is smaller when demand is inelastic. The key indicator of the distortion is how much change there is in the quantity traded. When demand is inelastic, the quantity traded changes by less. As a consequence, there is a much smaller deadweight loss in part (a) of Figure 12.2.9 than in part (b) of Figure 12.2.9.

Why Do Governments Impose Taxes?

Given our analysis so far, you might think that governments should not impose taxes at all. After all, taxes reduce the surplus received by buyers and sellers. However, there are several reasons why governments tax households and firms, despite the adverse consequences for the gains from trade.Many of these arguments for taxation are also discussed in other chapters.

Raising revenue. Governments perform certain essential functions, such as maintaining a legal system and defending the borders. Governments also typically supply various goods and services (such as roads, schools, and streetlights) as well as paying out subsidies to certain industries and transfers to individuals. All of these require government revenues. We are not interested right now in which of these things governments *should* do nor with the question of whether governments intervene too much or too little in the economy. It is simply a fact that governments incur a lot of expenses, and these expenses must be paid for through taxation. One key reason for taxes is therefore to raise revenue to fund government activities.

In fact, governments sometimes finance their expenses through borrowing rather than current taxation. But borrowing is the same as deferred taxation: the debt obligation must eventually be paid through taxes levied in the future.

Redistributing income. Taxes are a means by which governments can take money from one group of people and give it to another. Governments often use progressive taxation, meaning that the rich are taxed proportionately more than the poor. Taxation then serves to make the distribution of income more equal.In Chapter 13, we look in detail at the arguments for redistribution in society.

Externalities. In some circumstances, an individual's actions have an influence, either positive or negative, on others in the economy. Economists call such an effect an **externality**. Chapter 14 is all about such externalities. In the presence of externalities, distortions in the market and some type of government intervention may be warranted. Often, that intervention takes the form of taxes and subsidies that alter individual incentives to encourage behavior that promotes economic **efficiency**.

Sometimes externalities are adverse; these are known as negative externalities. The effect of second-hand smoke is an example. Other times there are positive externalities associated with an action. An example is education, which has benefits to society as well as to the individual who obtains the education. When there are negative externalities, the government can impose a tax to discourage the activity in question. When there are positive externalities, the corresponding government response is a subsidy.

Uninformed choices. Economists generally presume that informed individuals will make informed choices. Not everyone agrees with economists about this. One often hears the argument that governments ought to intervene so that individuals do not make the "wrong decisions." Take, for example, the decision to smoke cigarettes. It has been known for a long time that cigarette smoking is harmful to one's health. One reasonable view is that smoking should be purely a matter of individual choice: people can make their own choices about the enjoyment of smoking versus the adverse health effects. As long as individuals make informed choices, there seems to be little basis for government intervention.

But another view is that people are not always capable of informed choice. Perhaps people are not good at making decisions that involve their health 30 years from now. Perhaps people are not good at making decisions about addictive substances. Perhaps it is not appropriate to think of rational individuals making informed choices when many people start smoking as children. An argument can then be made that governments should step in and alter incentives, through taxes and subsidies, to help people make better choices.

Toolkit: Section 31.2

You can review the definition and calculation of elasticities in the toolkit.

Subsidies

A **subsidy** is the opposite of a tax. It is a payment made to a producer to encourage production. A subsidy means that the price paid by the buyer is lower than the price received by the seller. Figure 12.2.10 shows the deadweight loss from a subsidy. Subsidies distort markets not by leading to too small a quantity being traded but by causing too large a quantity to be traded. The deadweight loss lies to the right-hand side of the competitive equilibrium quantity because some trades occur where the cost exceeds the benefit. Figure 12.2.11 shows the buyer surplus and the seller surplus in the presence of a subsidy. Both are increased by the subsidy. However, subsidies mean that the government spends resources rather than taking them in. The figure shows that the cost of the subsidy is greater than the increased surplus received by the buyers and the sellers. The difference between the cost and the increases in surplus is the deadweight loss

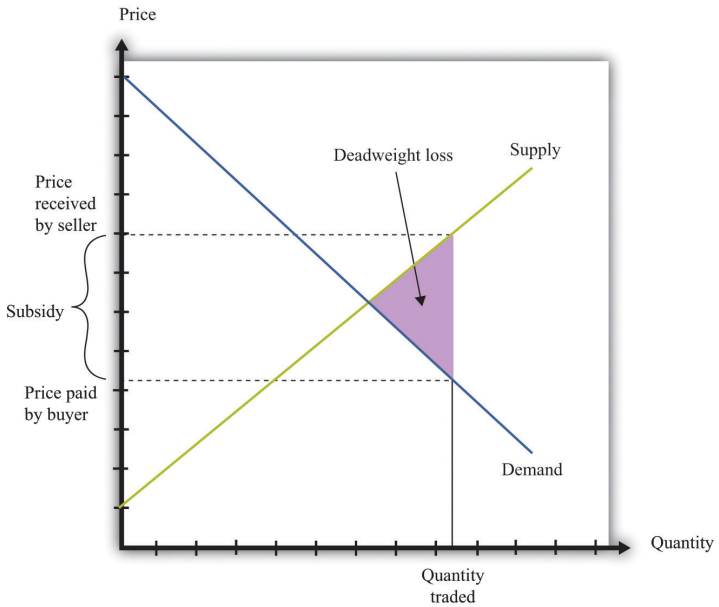


Figure 12.2.10: The Deadweight Loss from a Subsidy. A subsidy means that some transactions are now carried out even though they actually *destroy* value.

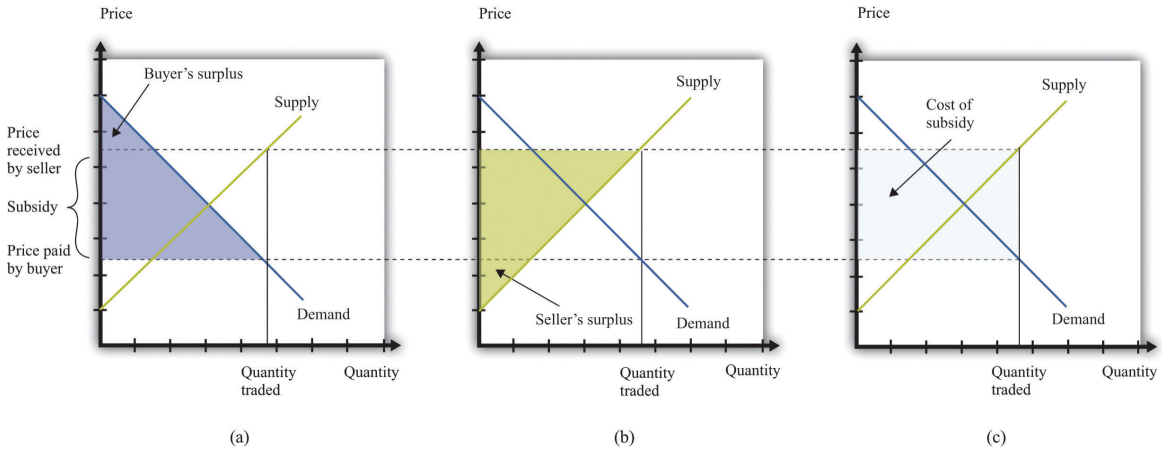


Figure 12.2.11: The Buyer Surplus and the Seller Surplus after the Imposition of a Subsidy. The buyer surplus and the seller surplus are shown in (a) and (b), and the *cost of the subsidy is shown in (c). The total surplus is obtained by adding together the buyer surplus and the seller surplus and then subtracting the subsidy paid by the government.*

Figure 12.2.12 summarizes the different kinds of trade restrictions that we have looked at.

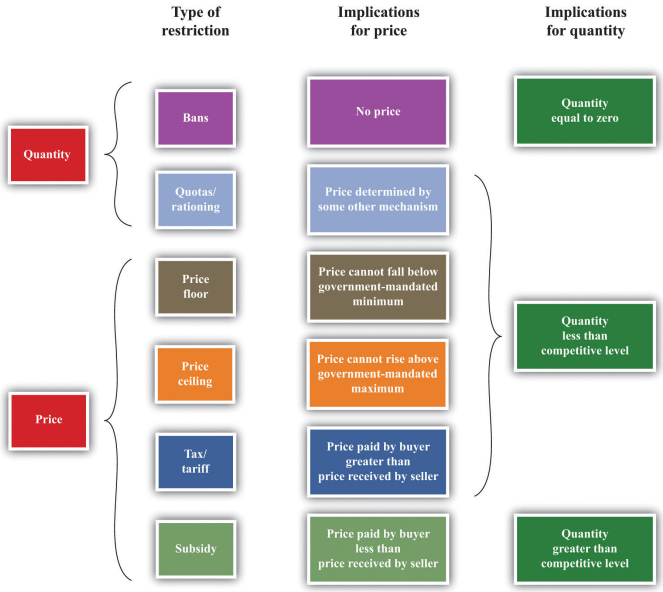


Figure 12.2.12: The Different Ways in Which Governments Intervene in Markets

Key Takeaways

- Government restrictions take a variety of forms, including bans on trades, controls on prices, and the imposition of taxes and subsidies to change incentives. These are summarized in Figure 12.2.12.
- Some of the reasons governments restrict trades are to protect individuals and society from unsafe and unhealthy products, for moral reasons, and for fairness.
- Through these restrictions, some gains from trade may be lost. For example, in the presence of taxes, there are deadweight losses due to the lost gains from trade. If a market is shut down entirely, then all the gains from trade are lost. In some cases, individuals find a way to circumvent government restrictions to realize these gains from trade.

Exercises

1. In what sense is the closing down of a market like a tax?
2. If the government sets a tax rate, how is the quantity of revenue collected determined?
3. Explain why the allocation of the tax burden does not depend on who pays a tax to the government.

This page titled 12.2: How the Government Controls What You Buy and Sell is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

12.3: Limits on Trade across Borders 12.3: Limits on Trade across BordersEdit Update titlecancel

Learning Objectives

1. What is the source of gains to trade across two countries?
2. Why do governments put restrictions on trade across borders?
3. How do governments restrict international trade?

Restrictions do not appear only within a country. We see restrictions on trade across countries as well. In our shopping list at the beginning of the chapter, we mentioned several goods that are imported from other countries, such as Cuban cigars and French cheese. We begin by reviewing the motivations for trade between countries. Just as individuals are motivated to trade by the fact that it can make them better off, countries can also benefit from trading with each other.

Comparative Advantage

The principle of **comparative advantage** provides one reason why there are gains from trade among individuals.We discuss this in Chapter 6. Because different individuals have different skills and abilities, everyone can benefit if people specialize in the things that they do relatively well and trade with others to obtain the goods and services that they do not produce. Such specialization is a cornerstone of our modern economy, in which people are specialists in production but generalists in consumption.

The idea of comparative advantage also provides a basis for trade among countries. In the absence of trade, countries end up producing goods and services that they can provide only very inefficiently. When countries trade, they can instead specialize in the goods and the services that they can produce relatively efficiently. All countries can take full advantage of their different capabilities.

We illustrate comparative advantage in a simple way, with a story about trade between Guatemala and Mexico. If you understand this story, you should also be able to see that we could make the example more complex and yet still keep the same basic insight. Table 12.3.1 "Beer and Tomato Production in Mexico and Guatemala" provides information about the **technology** in each country: how much a typical individual can produce in a 36-hour workweek. The table shows how much time is required in each country to produce two goods: beer and tomatoes.

Table 12.3.1: Beer and Tomato Production in Mexico and Guatemala

	Hours of Labor Required	
	Tomatoes (1 kilogram)	**Beer (1 liter)**
Guatemala	6	3
Mexico	2	2

In both Mexico and Guatemala, people like to consume beer and tomatoes in equal quantities: 1 liter of beer to accompany each kilogram of tomatoes. In Guatemala, it takes 6 hours of labor to produce 1 kilogram of tomatoes, and 3 hours of labor to produce 1 liter of beer. In 9 hours, therefore, it is possible to produce 1 kilogram of tomatoes *and* 1 liter of beer. In a 36-hour week, the worker can enjoy 4 kilograms of tomatoes accompanied by 4 liters of beer.

Mexico is much more efficient at producing both tomatoes and beer. It takes only 2 hours to produce 1 kilogram of tomatoes, and it takes only 2 hours to produce 1 liter of beer. In 36 hours, therefore, a Mexican worker can produce 9 kilograms of tomatoes and 9 liters of beer.

Because Mexico is better at producing both tomatoes and beer—it has an **absolute advantage** in the production of both goods—it would be natural to think that Mexico has nothing to gain from trading with Guatemala. But this conclusion is wrong. Mexico is a bit better at producing beer but a lot better at producing tomatoes. Guatemala has a comparative advantage in the production of beer. One way to see this is through opportunity cost. In Guatemala, the opportunity cost of producing 1 kilogram of tomatoes is 2 liters of beer. In Mexico, the opportunity cost of producing 1 kilogram of tomatoes is only 1 liter of beer. Thus Guatemala should specialize in the production of beer.

In a 36-hour week, Guatemala produces 12 liters of beer. Now suppose Mexico devotes 30 hours to producing tomatoes and only 6 hours to producing beer. Then Mexico will produce 15 kilograms of tomatoes and 3 liters of beer. The two countries produce, in total, 15 kilograms of tomatoes and 15 liters of beer. Previously, they were producing 13 kilograms of tomatoes and 13 liters of beer. Both countries can be better off if they trade and take advantage of comparative advantage. We illustrate this in Figure 12.3.1.If you have read Chapter 6, then this figure should look familiar. A similar figure shows up there for trade between two individuals. It shows the joint **production possibilities frontier** for Guatemala and Mexico. When they produce individually and do not trade, they end up at a point inside the production possibilities frontier. If they specialize and trade, they end up on the production possibilities frontier instead.

Toolkit: Section 31.12, and Section 31.13

You can review the idea of the production possibilities frontier and the concepts of comparative and absolute advantage in the toolkit.

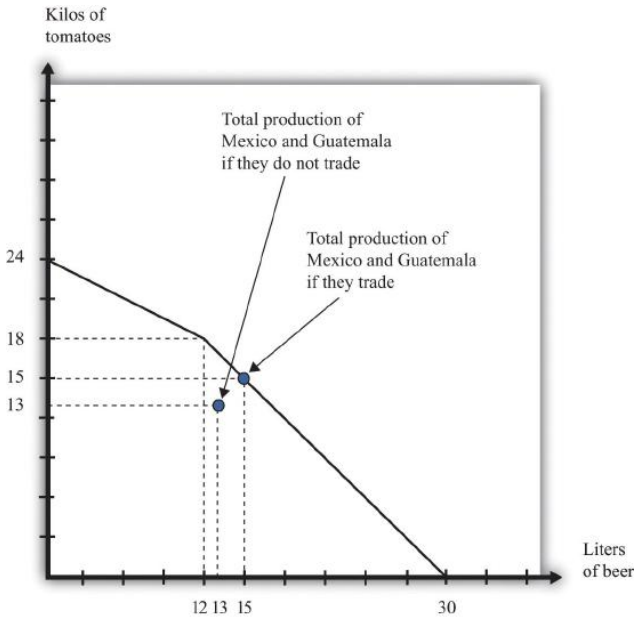


Figure 12.3.1: The Production Possibilities Frontier. Individually, Mexico and Guatemala produce 13 kilograms of tomatoes and 13 liters of beer. Jointly, they can produce 15 kilograms of tomatoes and 15 liters of beer by exploiting comparative advantage.

Why Do Governments Intervene in International Trade?

To economists, the logic of comparative advantage is highly compelling. Yet noneconomists are much less convinced about the desirability of free trade between countries. We see this reflected in the fact that countries erect a multitude of barriers to trade. Where economists see the possibility of free trade and mutual gain, others often see unfair competition. For example, many countries in the developing world have very low wages compared to the United States, Europe, and other relatively developed economies. Economists see this as a source of comparative advantage for those countries. Because labor is cheap, those countries can produce goods that require a large amount of labor. Countries like the United States, by contrast, can specialize in the production of goods that require less labor. The logic of comparative advantage suggests that both countries would be made better off. To noneconomists, however, the cheap labor looks like “unfair” competition—how can workers in rich countries compete with workers in poor countries who are paid so much less?

This concern has some merit. Comparative advantage tells us that a country as a whole is made better off by trade because that country can have more goods available for consumption. Yet comparative advantage, in and of itself, says nothing about *who* gets those benefits or how they are shared.

Hypothetically, it is possible to share these goods so that everyone is made better off.Actually, even this statement carries an implicit assumption that it is possible to share out these goods without distorting economic activity too much. In Chapter 13, we explain that redistribution typically involves some loss in efficiency. As a practical matter, even if the country as a whole has more

goods and services to consume, some individuals within the country are made worse off. There are winners and losers from trade, and there is frequently political pressure to limit international trade from or on behalf of those who lose out.

Another reason governments intervene in international trade is because of political lobbying. Generally, the beneficiaries from trade barriers are a small and identifiable group. For example, the United States provides sugar subsidies that increase the price of sugar. Sugar producers are the clear beneficiaries of this policy and have an incentive to lobby the government to ensure that the subsidies stay in place. The losers from this policy are those who consume sugar—that is, all of us. But there is no lobby representing sugar consumers.

Whatever the reasons, governments frequently intervene in international trade. Sometimes they completely close certain markets. Sometimes they impose limits on how much can be imported from abroad. And sometimes they impose special taxes on imports. We look at each in turn.

Sanctions and Bans

In some cases, governments close down certain categories of overseas trade completely. They may do so in an attempt to further international political goals. An example from our shopping list is the Cohiba cigars. You cannot buy these directly from Cuba due to a ban on the import of Cuban goods into the United States.The sanctions began in 1962 in response to the takeover of US property by the Cuban government of Fidel Castro. The Cuban Democracy Act of 1992 extended the sanctions. US Department of State, "Title XVII—Cuban Democracy Act of 1992," accessed March 14, 2011, http://www.state.gov/www/regions/wha/cuba/democ_act_1992.html. This policy is designed to make it harder for Cuba to function in the world economy and thus puts pressure on the Cuban government.

Governments quite often use international sanctions in an attempt to achieve political goals. These measures can be enacted by individual governments or by international bodies such as the United Nations. Currently, the international community is putting pressure on Iran because of concerns about the development of nuclear capabilities in that country.An embargo on trade with Iran was imposed by the United States in 1987. Details can be found at US Department of the Treasury, Resource Center, "Iran Sanctions," accessed March 14, 2011, http://www.treasury.gov/resource-center/sanctions/Programs/Pages/iran.aspx. From 1990 to 2003, there were international sanctions placed on trade with Iraq.

Governments also ban certain products from overseas for reasons of health and safety. The United States does not allow the importation of cheeses made with raw milk because it argues that such cheeses pose a health risk; thus it is difficult to find the French raw milk camembert on our shopping list. When the United Kingdom had an outbreak of bovine spongiform encephalopathy (better known as mad cow disease), many countries banned the import of beef from that country. More generally, countries have their own health and safety laws, so foreigners who wish to compete in these markets must ensure that they satisfy these standards.

Quotas

Another way in which governments frequently intervene in international transactions is by means of a **quota**—that is, a quantity restriction on imports. Figure 12.3.2 gives an example of how quotas work. Suppose that Australian consumers buy both domestically produced cars (Holdens) and cars imported from the United States (Fords). These cars are not perfect substitutes for each other, so we draw a market for each kind of car. To begin with, both markets are in equilibrium where demand equals supply. Suppose that Australia were then to impose a quota on the import of Fords. The price of Fords is determined by consumers' willingness to pay at the quantity set in the quota—that is, we can find the price by looking at the demand curve. Australian consumers must pay more for Fords. Meanwhile, the fact that fewer Fords are being sold means that Australian households will demand more Holdens. The demand curve shifts to the right. This increases the price of domestic vehicles.

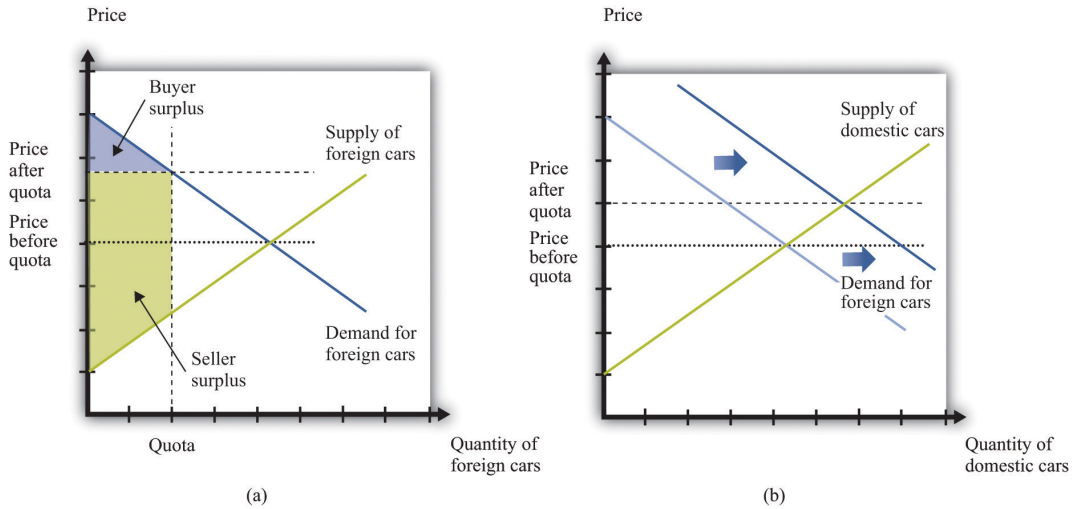


Figure 12.3.2: The Effects of a Quota. After the imposition of the quota, the price of cars increases in the market for foreign cars (a) and the demand for domestic cars increases (b).

Who are the winners and the losers in this process? The clear winners are domestic producers of automobiles. They get to sell more cars at a higher price, and their surplus increases. Australian consumers, meanwhile, are losers. We cannot see this immediately by looking at the buyer surplus because the buyer surplus decreases in the market for foreign cars but increases in the market for domestic cars. However, we can tell that consumers are worse off because both domestic and foreign cars have become more expensive. Finally, the effects on foreign producers are in general ambiguous. They sell fewer cars but at a higher price. American producers might even benefit from the Australian quota.

In general, governments that impose quotas are transferring resources from domestic consumers to domestic producers. This illustrates the point we made earlier: the beneficiaries of this kind of policy are typically a small group—in this case, producers of Holdens. The losers are everyone who wants to buy a car. The producers are likely to have much more political influence than the consumers.

Tariffs

Where quotas are the equivalent of quantity restrictions, applied in the context of international trade, tariffs are the equivalent of taxes. A **tariff** on a good is an amount that must be paid by someone who wishes to import that good from another country. The main implication of a tariff is that the price received by foreign sellers is less than price paid by domestic consumers. We illustrate a tariff in Figure 12.3.3.

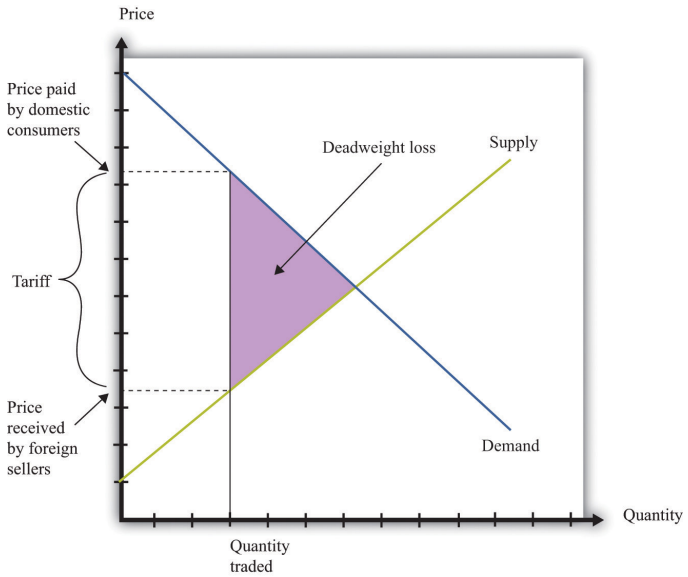


Figure 12.3.3: The Effects of a Tariff. A tariff means that there is a wedge between the price paid by domestic buyers and the price received by foreign sellers. Just as with a tax, the quantity traded is lower because some transactions are no longer worthwhile. There is a deadweight loss.

The main implications are very similar to those of a tax. Consumers are made worse off, as are foreign producers. There is a deadweight loss, as indicated in Figure 12.3.3. As with quotas, tariffs are often designed to protect domestic producers. Thus, as we saw when looking at a quota (Figure 12.3.2), a tariff on foreign goods induces a substitute toward goods produced in the domestic country. This is the law of demand at work: when the price of a good increases, the demand for substitute products will be increased.

One element that distinguishes a tariff from a quota is the collection of government revenue. When a quota is imposed, trade is limited at a particular quantity, but the government collects no revenue. Instead, as shown in Figure 12.3.2, the surplus from the trade is distributed among buyers and sellers. When a tariff is imposed, the government collects revenue equal to the product of the tariff and the quantity traded. Comparing Figure 12.3.2 with Figure 12.3.3, part of the surplus that was shared by buyers and sellers under the quota is now captured as revenue by the government. This parallels the results that we saw when looking at domestic quotas and taxes earlier in the chapter.

Key Takeaways

- There are gains to trade across countries due to comparative advantage.
- Governments place restrictions on trade for political reasons, to protect jobs, and to increase revenue by taxing trade.
- Governments may impose outright bans on trade, place limits on the quantities traded, or put taxes on trade.

Exercises

1. Looking at Table 12.3.1 "Beer and Tomato Production in Mexico and Guatemala", we said that Mexico had an absolute advantage in both goods, but still there were gains from trade. Change two numbers in the table so that Guatemala has an absolute advantage in both goods. Explain how each country still has a comparative advantage in your new example.
2. Do the economic effects of a tariff depend on who pays it?

This page titled 12.3: Limits on Trade across Borders is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

12.4: Government and the Labor Market 12.4: Government and the Labor MarketEdit Update titlecancel

Learning Objectives

1. What are the forms of government restrictions in labor markets?
2. What are the effects of government restrictions on migration?
3. Who bears the burden of a tax on labor income?

Some of the most important sets of markets in the economy are those for different kinds of labor. There are many ways in which governments intervene in these markets.

Toolkit: Section 31.3

You can review the labor market in the toolkit.

Licensing

Some occupations cannot be carried out without licensing or accreditation of some kind. You cannot set up in business as a doctor or a lawyer without any training. Here, the government's reason for intervening is because of information problems: we do not have the knowledge to determine if someone is indeed trained in medicine or law. Chapter 16 has more to say about this.

Migration Restrictions

Other things equal, people want to move to where they can earn a high wage. Within the United States, people are free to move from state to state in search of good jobs and good wages. Workers are likewise free to move among the countries of the European Union. In both places, we see many examples of people moving to where wages are higher. Young Polish students move to the United Kingdom in search of work; workers in Louisiana move to Washington state because wages are higher there. Obviously, many factors influence where people choose to live and work, but wages are one of the most important.

If a firm is willing to pay a worker $15 per hour in New Jersey but firms in Idaho will pay that same worker only $12 per hour, then this is an indication that the worker's time is more valuable in New Jersey than it is in Idaho. The market, through the higher wage, sends a signal to the worker that it is desirable to move. The movement of workers from Idaho to New Jersey will cause the supply curve of labor in Idaho to shift to the left, so wages in Idaho will increase, and the supply curve of workers in New Jersey to shift to the right, so wages in New Jersey will decrease. The movement of workers thus also serves to make wages more equal.

Workers in the United States are permitted to move anywhere in the country. The same is true for workers in New Zealand, Mexico, and most other countries. In some places, however, laws enacted by national or local governments make such migration harder. In China, certain government benefits are highly localized, making it difficult for a worker to move from one town to another.

In a world with no restrictions on labor movement, workers would move across countries as they do within a country. Consider the market for labor within the European Union. Figure 12.4.1 shows the markets for workers in Portugal and France. If labor is unable to migrate, then the equilibrium wage in France is higher than the wage in Portugal. Once labor mobility is allowed within Europe, workers naturally move to the labor market with the higher wage. This forces wages to decrease in France and increase in Portugal.

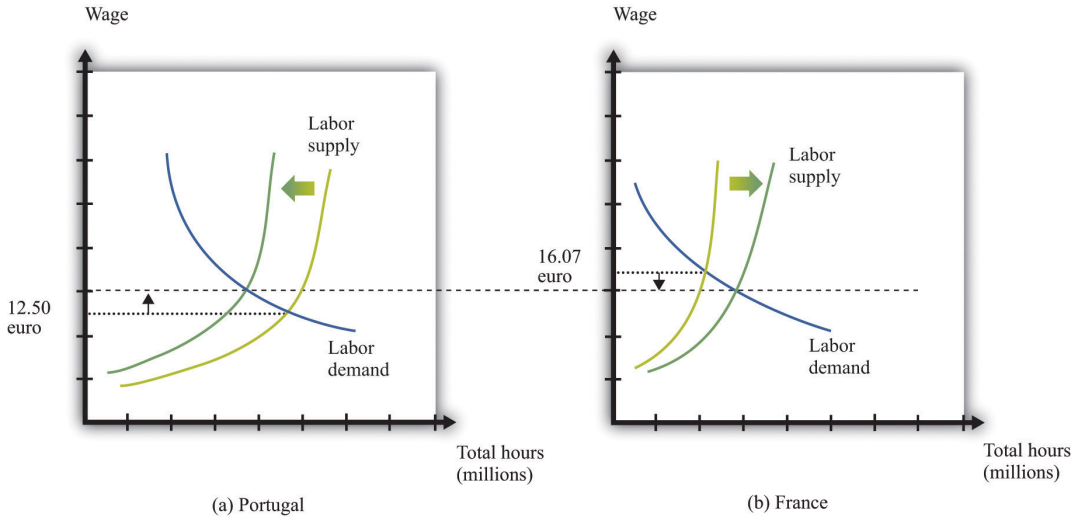


Figure 12.4.1: Migration Eliminates Wage Differences. Workers move from Portugal to France in search of higher wages.

If workers care only about wages, then migration would completely equalize wages in France and Portugal. In practice, some differences in wages persist. For example, if most people think living in Portugal is better than living in France, then the wage rate in Portugal will be lower than that in France. Despite this wage differential, individuals living in Portugal will not move to France. The higher wage is France is an example of a **compensating wage differential**: it is the difference in wages needed to compensate individuals for living and working in France.

Free migration across countries exists in the European Union, but international migration is typically much more restricted. One of the items on our shopping list at the beginning of the chapter was the hiring of an illegal domestic worker. This is yet one more example of a restriction on trade because people are not allowed to work wherever they want. Most countries restrict the amount of immigration permitted into the country; some countries restrict emigration as well.

Income Taxes

Governments also affect the labor market through the imposition of taxes. In most countries, there is an income tax. In some cases, income taxes may also be imposed more locally: some individual states within the United States have an income tax in addition to the federal tax. An income tax works like the taxes we saw earlier. Fundamentally, it means that the amount paid by the employer exceeds the amount received by the worker. Exactly as before, this gives rise to a deadweight loss. Some workers will choose not to work or work fewer hours because of the tax. It follows that some mutually beneficial transactions go unrealized.

Because the market for labor is so fundamental to the economy and because the income tax is, in most economies, a major source of revenue for the government, economists and politicians pay a lot of attention to this market. Figure 12.4.2 summarizes the effects of a tax on wages using a diagram of the labor market. When there is a tax on wages, there is a gap between the wage paid by the firm and the wage received by the worker. As shown in the figure, the effect of the tax is to reduce the quantity of labor traded. The wage paid by the firm is higher than the wage in the original equilibrium, and the wage received by the worker is less than the wage in the original equilibrium.

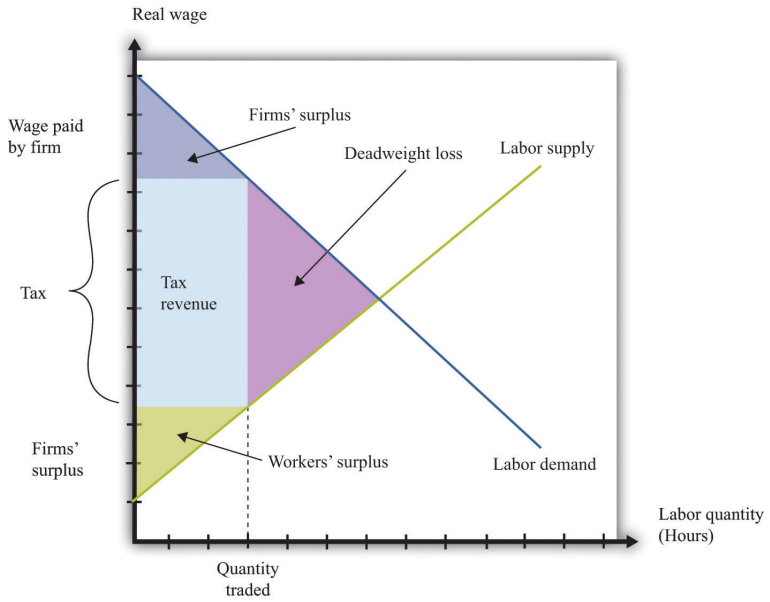


Figure 12.4.2: The Effect of an Income Tax in the Labor Market. An income tax means that there is a wedge between the wage paid by the firm and the wage received by a worker.

As in our earlier example, the incidence of the tax will depend on the elasticity of labor supply and labor demand. We can understand incidence by looking at how the tax on labor income affects the wage. If the wage paid by the firm increases by the amount of the tax, then the firm is paying the tax. This will happen if the demand for labor is very inelastic. If the wage received by the worker decreases by the amount of the tax, then the worker is paying the tax, not the firm. This will happen if the supply of labor is very inelastic.

Tax Evasion in the Underground Economy

In some countries, governments have difficulty collecting income taxes from their citizens. This is partly an enforcement issue: if many people in a country misrepresent their income, it is difficult to hire enough people to enforce the tax laws. In this case, income taxes become ineffective, and governments resort to other forms of taxation, such as sales taxes.

Another form of tax evasion is to conduct trades in the underground economy. When income taxes are very high, small business owners and other individuals may offer to do work "under the table." They will ask to be paid in cash, so there is no record of the transaction and no basis for collecting income tax. In return, they will do the work for a cheaper price. This is illegal, but the likelihood of getting caught is low enough that many people decide that avoiding the income tax is worth the crime. The magnitude of this underground activity can be substantial: "In a report to the Senate in May, Deputy Finance Minister Vincenzo Visco said that the hidden, untaxed economy accounted for around 27 percent of Italy's gross domestic product of nearly $2 trillion."Elisabetta Povoledo, "Italy Changes Rules of Tax-Evasion Game," *New York Times*, June 13, 2007, accessed January 29, 2011, http://www.nytimes.com/2007/06/13/world/europe/13iht-taxes.4.6129007.html?_r=1.

Thus there are two different aspects of the underground economy. There is the exchange of goods and services that cannot be traded legally (drugs, scalped tickets, etc.). And there are trades that are legal but not reported to the tax authorities (illegal).

The underground economy tends to be larger when income taxes are higher and where tax enforcement is difficult, but it exists everywhere. If your neighbor pays you $20 to mow his lawn, and you do not declare this on your taxes, you are participating in the underground economy. Besides allowing you to avoid income taxes, working in the underground economy has an additional benefit. If you do not work a regular job, then you can collect unemployment insurance. This means that you can work and earn income without paying taxes in the underground economy and also collect unemployment insurance.

A recent study by the International Monetary Fund (IMF) concluded, "In the European Union in the late 1990s, 20 million people engaged in shadow [underground] economy activities. In all European OECD countries combined, about 35 million people did so. In some individual countries, the shadow economy labor force was very large: in Italy, 30–48 percent of the total labor force (1997); Spain, 12–32 percent (1997–98); and Sweden, 20 percent (1997–98). In many countries, these high shares coexisted with high official rates of unemployment."Friedrich Schneider and Dominik Enste, "Hiding in the Shadows: The Growth of the Underground Economy," *International Monetary Fund*, March 2002, accessed January 29, 2011,

http://www.imf.org/external/pubs/ft/issues/issues30. According to this study, the underground economy is between 35 percent and 44 percent of gross domestic product (GDP) in developing countries and around 15 percent in the advanced Organisation for Economic Co-operation and Development (OECD) countries.

Finally, people sometimes barter goods and services rather than trade them. If you are a web designer and your next-door neighbor is a plumber, you might agree to build a website for her in exchange for her installing a new shower for you. Again, if you fail to report this "income in kind," you are evading taxes. Bartering schemes can be very sophisticated, involving the creation of groups that set up their own special money to pay for transactions.

Key Takeaways

- Governments often limit the movement of people across borders and tax labor income.
- Restricting the flow of labor across international borders can lead to an inefficient allocation of labor across its productive uses.
- The burden of a labor income tax, like other taxes, depends on the elasticities of supply and demand.

Exercises

1. If the government required licenses to sell fruit in outdoor markets, who would benefit from this restriction and who would lose?
2. If there are no migration restrictions between two states in the United States, must wages be equal in the two states?
3. Use a version of Figure 12.4.2 to show that if labor supply is very inelastic, then a worker is bearing most of the burden of the income tax.

This page titled 12.4: Government and the Labor Market is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

12.5: End-of-Chapter Material 12.5: End-of-Chapter MaterialEdit Update titlecancel

In Conclusion

The underground economy is not new; it has been around for as long as rulers have been levying taxes and banning trades. If you read about the prohibition of alcohol in the United States, for example, you will quickly learn that there was still a thriving market for alcohol and alcoholic beverages, despite the illegality of these trades. This was partly due to the fact that the production of alcohol was legal in nearby countries, such as Canada. Alcohol produced in Canada and elsewhere was imported and sold in the United States.

The establishments that served alcohol at that time were called speakeasies. Today you can find local bars that advertise themselves as having started as speakeasies during the Prohibition years. Of course, while Prohibition was in force, the speakeasies did not advertise so loudly. They were generally run by gangs that were willing to take the risk of being arrested to get the profits from selling alcohol.

Associated with Prohibition are several infamous individuals, such as Al Capone and his competitor, Bugs Moran. They were leaders of gangs in Chicago that provided alcohol to speakeasies. But you can, if you like, think of them as managers of firms that were involved in the importation, manufacturing, production, and sale of a consumer good. In many ways these firms operated according to the same principles as firms in this textbook. They were interested in producing efficiently and maximizing their profits.

Capone was eventually indicted and convicted. But the legal action against Capone was not directed at his violation of Prohibition. Instead, the federal government indicted him for tax evasion. Even if you are a leading producer in the underground economy, you still have to pay your taxes.

This story of Prohibition reminds us that the government does more than simply restrict trades in the economy. The government also provides the framework that allows trades. It provides a system of laws that allows people to enter into contracts, and it provides courts as a mechanism for enforcing these contracts.

Capone and Moran could not turn to the government to enforce their contracts and agreements. The firms in the industry had to create their own mechanisms for settling disputes. You won't be surprised to hear that these mechanisms were not pretty. One famous incident was the Saint Valentine's Day Massacre in 1929 when the Capone gang engaged the rival gang led by Bugs Moran. This was like a strategic interaction between rival producers. In this case, their respective competitive strategies left seven people dead. When the government is not there to enforce contracts, agreements will be enforced by other, often violent, means.

Key Links

- Food and Drug Administration: http://www.fda.gov
- US Citizenship and Immigration Services: http://www.uscis.gov/portal/site/uscis
- Jeffrey Miron's blog: jeffreymiron.com/2010/04/drug-prohibition-and-immigration

exercises

1. List three additional examples of government restrictions on your ability to buy or sell something.
2. The sharing of the burden of a tax also depends on the elasticity of supply. Draw diagrams like Figure 12.5.7 looking at the case of elastic and inelastic supply—that is, draw two diagrams with identical demand curves but different supply curves. How does the elasticity of supply affect the changes in the buyer surplus and the seller surplus? Can you explain why?
3. (Advanced) In some countries, there are restrictions on the length of a contract to rent an apartment. Suppose the restriction is that contracts must last for five years. In response, some people sign private agreements to rent for shorter durations, such as a year. What are the problems that might arise from signing these private agreements? What happens if there is a dispute? What role might reputations play in the enforcement of these private agreements?
4. Suppose there is a forecast that a hurricane will hit in a day. Everyone expects the government to ration the supply of coffee. What will likely happen to the price of coffee once the forecast is announced?
5. How does rent control affect the incentives for an owner to invest in upgrades of an apartment?
6. The payout from a social security system depends on years worked. How would an increase in social security payments affect the choice of workers between jobs in the formal and informal (underground) parts of the economy?
7. If the underground part of an economy is large due to tax evasion, could a tax cut increase tax revenue?

8. If two states have different rates of labor taxation, what can you say about wages before and after taxes in the two states?
9. Can you think of a good or a service that the government *requires* you to consume? Why do you think the government has this policy?
10. There are substantial differences in food and product safety standards across countries. Can you think of reasons why this might be the case?
11. Liquor sales are state controlled in Pennsylvania but not in New Jersey. What effects do you think this has on the buying and selling of liquor near the border of the two states?
12. Suppose that Principles of Economics is a very popular course at your school. More people want to take the course than there are seats available. Do you think it would be a good idea if those initially enrolled in the class were able to sell their seats to those who didn't get a spot? What would be the advantages of such a system? What would be the problems?
13. Sometimes armies are raised by a draft, while in other times armies are volunteer. Which way of raising an army do you think is most efficient in terms of getting the best people to participate in the army? Which way of raising an army is most "fair"?
14. (Advanced) When Question 13 talks about "the best people to participate in the army," does it make a difference whether we are talking about comparative advantage or absolute advantage?
15. Explain how the incidence of a new tax on textbooks, collected at the point of sale, will be determined.
16. Which type of trade barrier creates more revenue for the government—a tariff or a quota? Why would a government ever impose a quota?
17. (Advanced) One benefit of working in the formal labor market in some developing countries is eligibility for both unemployment insurance and retirement pensions. All else being the same, would you predict that wages are higher in the formal or the informal sector of the economy? In addition, workers in the informal sector do not pay income taxes. What is the effect of this on wage levels in the two sectors?
18. Due to mobility restrictions, the labor markets in China are not fully integrated. If restrictions on mobility of workers in China were relaxed, what would happen to wage differences across regions? What predictions would you have for the flow of workers across parts of China?
19. If you are a member of a professional union, would you be in favor of licensing requirements to join that profession? How might you defend the need to have a license?

Economics Detective

1. Suppose you live in Mexico. If you wanted to get a job in Canada, what would you have to do to obtain permission to work? What if, instead, you wanted to work in the United States? Does your answer depend on your occupation?
2. Try to find estimates of the size of the underground economy in two different countries (for example, Portugal and Sweden). Is the underground economy of very different sizes in the two countries? Why?

This page titled 12.5: End-of-Chapter Material is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

13: Superstars

13.1: Rich and Richer
13.2: Facts about Inequality
13.3: The Sources of Inequality
13.4: Distributive Justice
13.5: Government Policy
13.6: End-of-Chapter Material

This page titled 13: Superstars is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

13.1: Rich and Richer 13.1: Rich and RicherEdit Update titlecancel

Table 13.1.1 "Wealthiest Individuals in the United States" shows the top 10 wealthiest people in the United States in 2006 and 2010. These names come from lists compiled each year by *Forbes* magazine of the 400 wealthiest individuals.Forbes has many such lists available for your study (http://www.forbes.com/lists). You almost certainly recognize some of the names, such as Bill Gates and Michael Dell from your dealings with the computer industry. Other names may be less familiar to you.

Table 13.1.1: Wealthiest Individuals in the United States

Rank	2006 List	2010 List
1	William H. Gates III	William H. Gates III
2	Warren E. Buffett	Warren E. Buffett
3	Sheldon Adelson	Lawrence J. Ellison
4	Lawrence J. Ellison	Christy Walton
5	Paul G. Allen	Charles Koch
6	Jim C. Walton	David Koch
7	Christy Walton	Jim C. Walton
8	S. Robson Walton	Alice Walton
9	Michael Dell	S. Robson Walton
10	Alice L. Walton	Michael Bloomberg

Whether or not you know their names, you surely have difficulty conceiving of their wealth. Bill Gates's net wealth in 2010 was estimated at $54 billion, which is $9 billion more than the wealth of financier Warren Buffett. To give some idea of what this means, if Gates were to receive no further income for the rest of his life but wanted to use up all his wealth before he died, he would need to spend it at a rate of about $5 million a day. The person at the bottom of the *Forbes* list—that is, the 400th wealthiest person in the United States—had a net worth of a mere $1 billion.

Comparing the two lists, you can see that some of the names and rankings changed between 2006 and 2010. The top two names are the same in both years, but the rest of the list is different. Sheldon Adelson, Paul Allen, and Michael Dell were in the top 10 in 2006 but not in 2010. In 2010, Charles and David Koch joined the top 10. Even among the very rich, there is some instability within the distribution of wealth.

The *Forbes* list was of the wealthiest Americans. Only the top 3 from the 2010 list are on the list of the world's wealthiest individuals. In 2010, the wealthiest individual in the world was Carlos Slim Helu, a Mexican businessman who made his fortune from real estate speculation and the telecom industry. Others in the world top 10 come from India, France, Brazil, Spain, and Germany. *Forbes* also publishes many other lists, including a list of the most powerful celebrities. At the top of that list in 2010 was Oprah Winfrey, who earned $315 million. (Notice that this is her *income*—the amount she earned in the year—while Table 13.1.1 "Wealthiest Individuals in the United States" is based on the total *wealth* accumulated.) Also on the list were Beyonce Knowles, Lady Gaga, Tiger Woods, Johnny Depp, and others from the entertainment industry.

When *Forbes* published its 2007 list, it also published an article by economist Jeffrey Sachs discussing the other extreme of the wealth distribution: the world's poorest households. Sachs pointed out that there are about a billion households in the world living on about $1 a day. He calls this group the *Forbes* One Billion. Sachs calculates that the richest 946 households have the same earnings as the *Forbes* One Billion. The discussion in *Forbes* and the calculations by Sachs make it clear that there are immense differences in income and wealth across people in the world. This is true both if we look across countries, comparing the richest to the poorest nations, and if we look within countries.

These differences are *persistent*, meaning that an individual's place in the income or wealth distribution is not likely to change significantly from one year to the next. If you are poor this year, you will probably be poor next year. It is not impossible for people to become rich overnight, but it does not happen often. In fact, such differences persist not only from year to year but also from generation to generation. This doesn't mean that everyone is completely stuck in the same place in the economic hierarchy. There

are opportunities for children to become much richer—or much poorer—than their parents. But when we look at the data, we will see that the income level of parents is an important indicator of the likely income of their children.

One goal of this chapter is to document some facts of inequality. This is not a straightforward task. For one thing, it is not even clear what measure of a household's economic success we should look at. Is it more useful to look at inequalities in income, wealth, consumption, or some other variable altogether? We also get a different picture if we look at these differences at a point in time or across time.

Data on inequality matter for discussions about taxation and redistribution. Governments throughout the world levy a number of different taxes, including taxes on the income people earn and the purchases that they make. Some of the revenues from these taxes are transferred to poorer households in the economy. The taxation of some households and the transfer of the resulting revenue to other households make up the redistribution policies of the government. We are interested in documenting facts about inequality in large part because we need these facts to have a sensible discussion about how much redistribution we—as a society—would like.

In this chapter, we therefore consider the following questions.

What determines the distributions of income, wealth, and consumption?

Is the market outcome "fair" or is there a need for government intervention?

What are the consequences of government redistributions of income and wealth?

Road Map

A road map for this chapter is shown in Figure 13.1.1. We begin with some facts about inequality and introduce some techniques to help us describe the amount of inequality both in a country and across countries. Then we consider some explanations of why we observe inequality in society. We observe first that people have different abilities, which translate into differences in income. Then we consider how individual choices—about education, training, and effort—are a further source of difference.

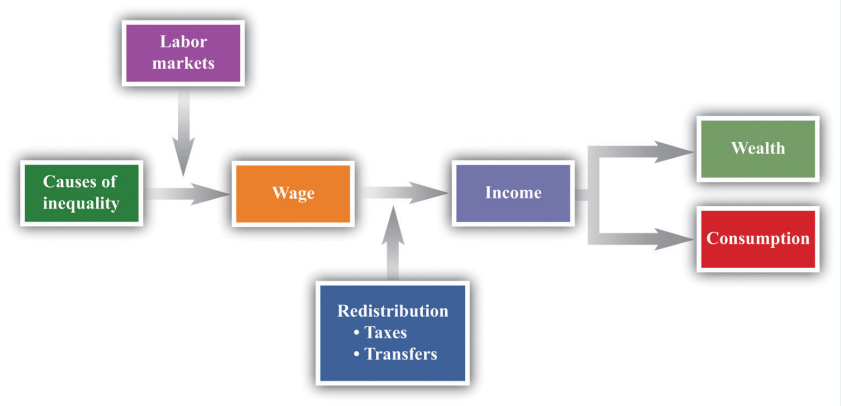


Figure 13.1.1: Road Map. This figure shows a plan for this chapter. We investigate the different underlying causes of inequality and explain how these translate, through labor markets in the economy, into differences in wages. We then explain how government policies affect the distribution of income in the economy. We also look at what determines the distribution of income, consumption, and wealth.

We then turn to a more abstract discussion of some different philosophical views of inequality. These different views influence current thinking about the distributions of income, wealth, and consumption and help us understand why people have such different opinions about equality and redistribution. We consider how redistribution might affect people's incentives to work, study, and cheat. Finally, we turn to economic policies that affect inequality.

This page titled 13.1: Rich and Richer is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

13.2: Facts about Inequality 13.2: Facts about InequalityEdit Update titlecancel

Learning Objectives

1. What is a Lorenz curve?
2. What is a Gini coefficient?
3. What has happened to income distribution in the United States?

There is no single, simple measure of the amount of inequality in a society. For example, we could study the distribution of consumption, income, or wealth, but each will tell us something different about the amount of inequality in our economy. These differences matter for the debate about inequality and our evaluation of policy.

The Lorenz Curve and the Gini Coefficient

Suppose you want to document the distribution of income in an economy. You could begin by asking every household its level of income. In many countries, the government already collects such data. In the United States, for example, this investigation is carried out by the US Census Bureau (http://www.census.gov). If everyone on the list had exactly the same level of income, you would conclude that income was equally distributed. If all but one person on the list had zero income and the remaining person had all the income, then you would conclude that income was very unequally distributed. In reality, of course, you would find that different households have all sorts of different levels of income.

The **Lorenz curve** provides a useful way of summarizing the distribution. It plots the fraction of the population on the horizontal axis and the percentage of income received by that fraction on the vertical axis. We construct a Lorenz curve as follows.

1. Take the list of incomes and order them from the lowest to the highest.
2. Calculate the total income in the economy.
3. Calculate the income of the lowest 1 percent of the population. Then calculate the income of the lowest 1 percent of the population as a percentage of total income.
4. Calculate the income of the lowest 2 percent of the population. Then calculate the income of the lowest 2 percent of the population as a percentage of total income.
5. Continue for all income levels.
6. Plot these points on a graph with fraction of the population on the horizontal axis and fraction of income on the vertical axis.

We know that 0 percent of the population earns 0 percent of the income, so the Lorenz curve starts at the origin. We also know that 100 percent of the population earns 100 percent of the income, so the other end of the Lorenz curve is at that point. If income were exactly equally distributed, then any given fraction of the population would earn that same fraction of income. The lowest 28 percent of the population would earn 28 percent of the income, the lowest 74 percent of the population would earn 74 percent of the income, and so on. In this case, the Lorenz curve would be a 45-degree line connecting the two endpoints. The closer the Lorenz curve to the 45-degree line, the more equal the distribution of income.

Table 13.2.1 "Example of Income Distribution" illustrates how to calculate the points on a Lorenz curve. The table shows four households, ordered by their income levels. The total income earned is $2,000. The lowest household (25 percent of the population) earns 5 percent of the total income because = 5 percent. If there were complete equality, this number would be 25 percent. So the lowest income household accounts for one quarter of the population but only one twentieth of the income. The first and second households together account for 50 percent of the population (see the last column of the table). They earn $500 in total, which is 25 percent of the total income. The first, second, and third households account for 75 percent of the population and 50 percent of the total income. Finally, if we look at all four households (100 percent of the population), this group earns $2,000, which is, of course, 100 percent of the total income. This Lorenz curve is illustrated in Figure 13.2.1.

Table 13.2.1: Example of Income Distribution

Household	Income Level ($)	Percent of Total Income Earned by Household	Percent of Total Income Earned by All Households with This Income or Lower	Percentage of Population with This Income or Lower
1	100	5	5	25

Household	Income Level ($)	Percent of Total Income Earned by Household	Percent of Total Income Earned by All Households with This Income or Lower	Percentage of Population with This Income or Lower
2	400	20	25	50
3	500	25	50	75
4	1,000	50	100	100

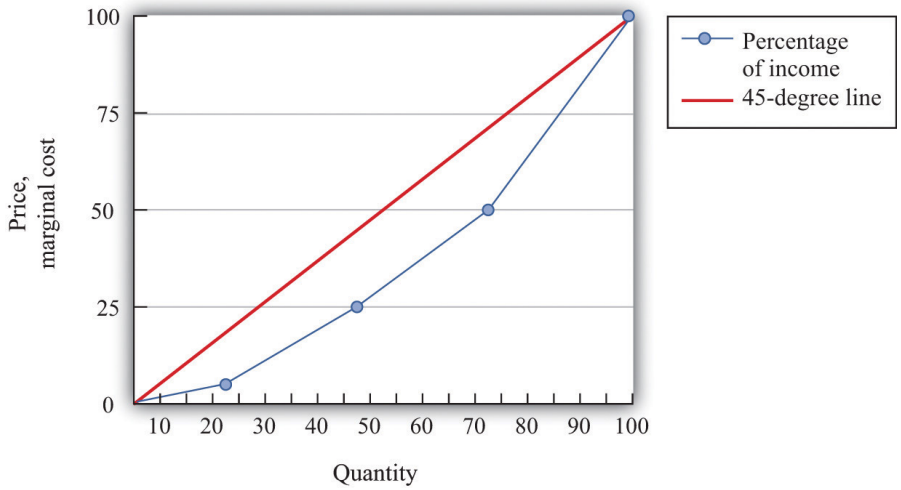


Figure 13.2.1: The Lorenz Curve. The more equal the distribution, the closer is the Lorenz curve to the 45-degree line.

We explained that the Lorenz curve coincides with the 45-degree line if there is complete equality. There is also a Lorenz curve for the case of complete inequality—in which a single person earns all the income. In this case, the Lorenz curve lies along the horizontal axis until the final household (that is, at 100 percent on the horizontal axis). At that point, the Lorenz curve lies along the vertical line at the right of the figure because the last person has all the income. Real economies exhibit neither complete equality nor complete inequality; a typical Lorenz curve lies below the 45-degree line and above the horizontal axis.

If we want to compare inequality over time or across countries, then we need something even simpler than the Lorenz curve. For this, we use the **Gini coefficient**, which is equal to the area between the 45-degree line and the Lorenz curve divided by the area below the diagonal. Figure 13.2.2 shows how the Gini coefficient is related to the Lorenz curve.

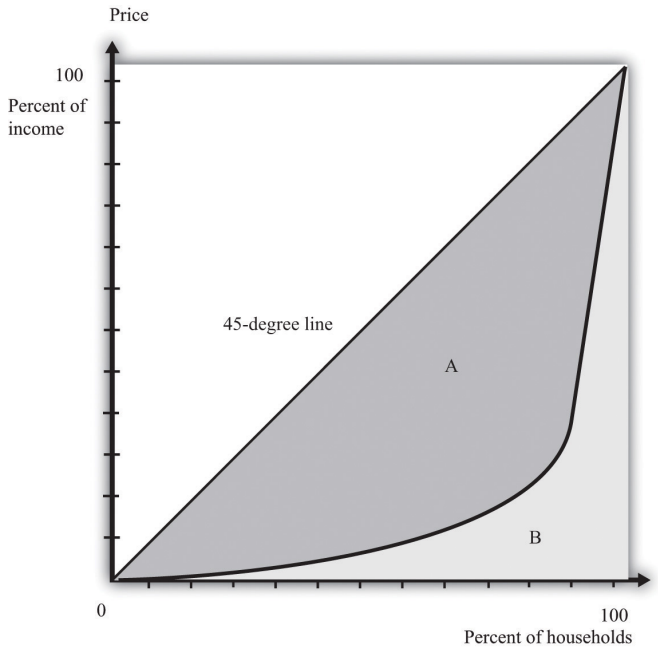


Figure 13.2.2: The Lorenz Curve and the Gini Coefficient. The Gini coefficient is calculated as the area between the Lorenz curve and the 45-degree line divided by the area under the 45-degree line—that is, it equals *A/(A + B)*.

If the Lorenz curve is exactly the same as the 45-degree line, then the Gini coefficient is zero. In this case, there is no area between the Lorenz curve and the 45-degree line. At the other extreme, if the Lorenz curve coincides with the horizontal axis until the final household, then the area above the Lorenz curve and the area below the diagonal are exactly the same. With complete inequality, the Gini coefficient is one. A higher Gini coefficient therefore means more inequality in the distribution of income.

Data on Inequality

We now use the Gini coefficient and other data to look at some facts about the distributions of income and wealth.

The Distribution of Income

Table 13.2.2 "Household Income by Quintile" presents data from the US Census Bureau on the distribution of various measures of income from 2003 to 2005. There are three measures of income given for each of the three years:

1. **Market income.** A measure of income earned from market activity, such as labor income and rental income.
2. **Postinsurance income.** Market income plus transfers received from the government.
3. **Disposable income.** Market income less taxes paid to the government plus transfers received from the government.

This table tells us how government redistribution affects the link between wage earnings and income.

Table 13.2.2: Household Income by Quintile

Quintiles	Market Income			Postinsurance Income			Disposable Income		
	2003	**2004**	**2005**	**2003**	**2004**	**2005**	**2003**	**2004**	**2005**
Lowest	1.5	1.5	1.5	3.3	3.3	3.2	4.6	4.7	4.4
Second	7.5	7.4	7.3	8.9	8.6	8.6	10.3	10.3	9.9
Third	14.5	14.1	14.0	14.8	14.5	14.3	15.8	16.1	15.3
Fourth	24.2	23.6	23.4	23.5	23.0	22.8	23.8	24.0	23.1
Highest	52.5	53.4	53.8	49.6	50.6	51.0	45.6	44.9	47.3
Gini coefficient	0.492	0.496	0.493	0.446	0.449	0.447	0.405	0.400	0.400

Source: US Census Bureau, Current Population Survey, 2004–2006 Annual Social and Economic Supplements.

These measures of income for each of the three years create the columns of the table. The rows of the table are quintiles (fifths) of the population. As in the construction of the Lorenz curve, the population is ordered according to income. This means the first quintile is the bottom 20 percent of the population in terms of income. The fifth quintile is the top 20 percent of the population in terms of income. To see how these quintiles are created, imagine taking 100 people and arranging them by their income, starting at the lowest level. Then create five groups of 20 people each where the first 20 people in the income distribution are in the first group, the second 20 in the income distribution are in the second group, and so on. Each group of 20 is a quintile of this population.

For each measure of income and for each year, there is an entry in the table showing the fraction of income in that year for a particular quintile. For example, looking at disposable income in 2004, the third (middle) quintile had 16.1 percent of the disposable income, and the highest quintile had 44.9 percent.

There are two striking features of this table. First, there is substantial inequality in the US economy. Looking at market income, the lowest 20 percent of the population receive about only 1.5 percent of the total market income. Contrast this with the highest quintile, which receives more than 50 percent of the total market income. This inequality is reflected in the Gini coefficient of about 0.49. If we look at the very top of the income distribution, the inequality is even more marked: the top 5 percent of the population in 2005 received about 30 percent of income after taxes and transfers, and the top 1 percent received about 16 percent of income.These figures come from Congressional Budget Office, *Historical Effective Federal Tax Rates, 1979 to 2005*, table 4C, accessed March 14, 2011, www.cbo.gov/ftpdocs/88xx/doc8885/Appendix_wtoc.pdf; the definitions of income therefore differ slightly from the US Census Bureau numbers in the table.

Second, the Gini coefficient decreases if we look at postinsurance income relative to market income and at disposable income relative to postinsurance income. This is because transfers represent—on average—a flow from richer to poorer households, and taxes are *progressive*: they redistribute from the rich to the poor. Government policies bring about some redistribution from richer households to poorer households. That said, there is still substantial inequality even after this redistribution: the lowest quintile receives less than 5 percent of total income, while the highest quintile receives about 45 percent.

Table 13.2.3 "Gini Coefficient over Time" shows changes in the Gini coefficient over time. (The data are on household incomes and come from the Census Bureau. See http://www.census.gov.) This table shows that inequality in the United States, as measured by the Gini coefficient, has increased steadily over the last few decades. In fact, if you go back to the end of World War II, the end of the 1960s represents a turning point in the income distribution.Thomas Piketty and Emmanuel Saez, “Income Inequality in the United States, 1913–98,” *Quarterly Journal of Economics* 118 (2003):1, together with their updated data set available at Emmanuel Saez’s faculty home page, http://www.econ.berkeley.edu/~saez/TabFig2008.xls. From 1940 through the 1960s, the income share of the top 10 percent fell from about 45 percent to about 33 percent. But starting in the 1970s, the pattern reversed, so that by 2007, the share of the top 10 percent exceeded 45 percent of total income.

Table 13.2.3: Gini Coefficient over Time

Year	Gini Coefficient
2009	0.469
2001	0.466
1997	0.459
1992	0.434
1987	0.426
1982	0.412
1977	0.402
1972	0.401
1967	0.399

Figure 13.2.3 focuses on the top of the income distribution: the top 1 percent. In part (a) of Figure 13.2.3, we can see that the real income of the bottom 99 percent of the population increased dramatically between the 1930s and the 1970s, increasing from $9,000

in 1933 to over $40,000 in 1973. (These numbers are adjusted for inflation and are in 2008 dollars.) Income over this period, for this group, grew an average of 3.7 percent per year. Over the next 35 years, the real income of this group hardly grew at all: the average growth rate was 0.2 percent per year. By contrast, the income of the top 1 percent grew only 1.7 percent per year on average between 1913 and 1973 but grew at an average 2.8 percent from 1973 to 2008. As a consequence, the top 1 percent of the income distribution roughly doubled their share of total income over this period.

At the very top of the income distribution, we have the true superstars: rock stars, movie stars, sports stars, top CEOs, and so on. The top 0.01 percent of the population—that is, the richest 30,000 or so people—has seen their share of income increase sevenfold since 1973.

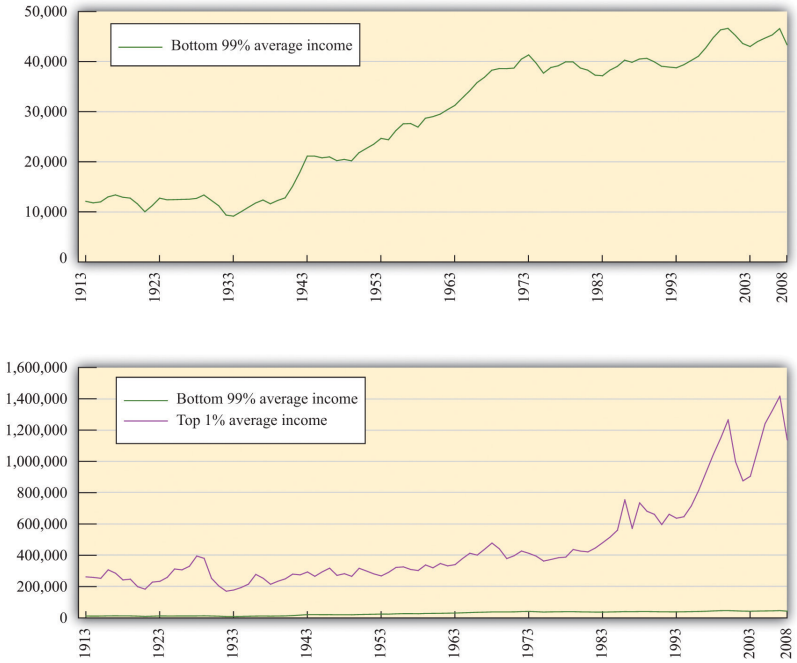


Figure 13.2.3: The Distribution of Income from 1913 to 2008. (a) The average real income in 2008 dollars for the bottom 99 percent of the population rose substantially between the 1930s and the 1970s but has been much flatter over the past few decades. (b) The top 1 percent has seen substantial income growth in recent decades.

The Distribution of Wealth

Table 13.2.4 "Gini Coefficients for Net Worth" looks at wealth data for a cohort of individuals between 1989 and 2001. At the beginning of the study, this group was between 34 and 43 years old. Arthur B. Kennickell, *A Rolling Tide: Changes in the Distribution of Wealth in the U.S., 1989–2001* (Washington, DC: Federal Reserve Board, 2003). Wealth is defined as assets minus liabilities. We can see that the Gini coefficients for wealth are considerably larger than the ones we saw earlier for income. There is more equality in income than in wealth.

Table 13.2.4: Gini Coefficients for Net Worth

Year	Gini Coefficient
1989	0.74
1992	0.75
1995	0.75
1998	0.76
2001	0.78

Income is a *flow*, meaning that individuals receive labor income on a weekly or monthly basis. Wealth is a *stock*: it is a measure of the assets that an individual or a household has accumulated and is measured at a particular point in time. Wealth comes partly from what people inherit and partly from decisions they make about allocating income between consumption and saving. The table also shows that wealth inequality increased for this group. There are two reasons that this could happen: (1) it may reflect greater inequality as a whole in society and (2) it may be due to inequality increasing as people become older.

Dynamics of Inequality

The position of a household in the income distribution is not static. A household in the lowest quintile of income one year will not necessarily be there the following year. Households can move up and down in the income distribution. For example, suppose you are fortunate enough to win the lottery or publish a hit song. Your income and thus your position in the income distribution will change quickly. For others without a hit song or luck with the lottery, changes in income can take more time. Perhaps you invest in a college education; after graduation and with a new job, you begin a climb through the income distribution. Bad luck can send you in the opposite direction. If your skills become less valuable, perhaps because of changes in technology, you may find that you have to move from a higher-paying to a lower-paying job, or you may become unemployed. There are many routes from rags to riches and from riches to rags.

One reason for mobility is the changes in income that most people experience in their lifetimes. For most people, the income they earn in their first job after school pays a lot less than the job they retire from. Thus most individuals experience a profile of income over their lifetime that takes them from one part of the income distribution to another. For most people, income also decreases in retirement.

Table 13.2.5 "Dynamics of Income in the United States" illustrates these dynamics over a five-year period. The top part of the table refers to earnings and the lower part to wealth. The data come from looking at distributions of earnings and wealth in two years: 1989 and 1994.

Table 13.2.5: Dynamics of Income in the United States

Measure	1989 Quintile	1994 Quintile				
		Highest	Fourth	Third	Second	Lowest
Earnings	Highest	90	7	2	1	0
	Fourth	27	34	30	6	2
	Third	9	14	45	25	6
	Second	5	6	15	51	23
	Lowest	5	5	6	17	68
Wealth	Highest	63	26	7	3	2
	Fourth	27	45	17	8	3
	Third	7	22	45	20	6
	Second	3	5	26	45	21
	Lowest	1	3	5	25	67

Source: Santiago Rodríguez, Javier Díaz-Giménez, Vincenzo Quadrini, and José-Víctor Ríos-Rull, “Updated Facts on the U.S. Distributions of Earnings, Income and Wealth,” *Federal Reserve Bank of Minneapolis Quarterly Review*, Summer 2002.

Under “Earnings,” there are five rows indicating the quintiles of the distribution in 1989. Along the top, there are five columns indicating the quintiles of the distribution in 1994. The entries refer to the percentage of people who go from one quintile in 1989 to another quintile in 1994. For example, 27 percent of the households in the second highest quintile in 1989 were in the top quintile in 1994, while 34 percent of the households in the second highest quintile in 1989 stayed there. A similar interpretation is given for the wealth part of the table.

The two parts of this table give a sense of income and wealth mobility through the distribution. If there were no mobility over time, so that households stayed in the same income and/or wealth quintiles), then the table would have 100 on the diagonal and 0

everywhere else. Mobility is indicated by the fact that the numbers along the diagonal are less than 100. From the part of the table referring to earnings, 90 percent of the people in the top income group in 1989 were there in 1994 as well. This means that very high income is extremely persistent. In contrast, only about two-thirds of the people in the lowest income class in 1989 remained in that group in 1994, while 17 percent moved up one quintile. As time passes, those who moved up will then move on to other parts of the income distribution.

Table 13.2.5 "Dynamics of Income in the United States" shows income and wealth dynamics over a relatively short period of time. It is also useful to look at dynamics across generations, though data are more difficult to obtain. One approach that researchers use over longer periods of time is to follow families. If your family was in the middle income group, we can see the likelihood that you will be in that same income group or in another income group. These dynamics take a longer amount of time because they are affected by things like parents' choices about the education of their children.

One way to study intergenerational income mobility is to take a group of individuals at a point in time and see how much of their current income can be "explained" by the income of their parents. (*Explained* is in quotation marks because it is difficult to disentangle the effects of family income from other influences. There are many factors associated with parents' income, such as the quality of schools and schoolmates, which are correlated with family income.)

One study reports an **elasticity** of 0.5 on the relationship between family and child income. This means that if parents' income is 1 percent higher, the child's income will be higher by about 0.5 percent. So if two families have an income difference of $100,000, then the prediction is that their children will have a difference of $50,000.The estimate is reported in Thom Hertz, *Understanding Mobility in America* (American University, Center for American Progress, April 26, 2006). This number is higher for the United States than for almost all the other (mostly European) countries studied. This same elasticity in Denmark is only 0.15, for example

> **Toolkit: Section 31.2**
>
> You can review the concept of elasticity in the toolkit.

The same study also looked at the mobility of families across the quintiles of income. A child whose family was in the middle quintile income had about a 40 percent chance of moving down the income distribution to a lower quintile and a 36.5 percent change of moving up. But 47 percent of the children born to a family in the lowest quintile remained there.

Inequality in Other Countries

Table 13.2.6 "Gini Coefficients in Different Countries" presents some evidence on the distribution of income in different countries. There are some significant differences across countries in income inequality. Eastern European countries, such as Hungary and Albania, and Western European countries, such as Sweden and France, have relatively equal distributions of income. At the other extreme, countries like Namibia and Brazil are highly unequal. The United States is about in the middle of these distributions.

Table 13.2.6: Gini Coefficients in Different Countries

Country	Gini Coefficient in 2005
Namibia	0.71
Brazil	0.59
South Africa	0.58
Mexico	0.55
Zambia	0.53
Argentina	0.52
Malaysia	0.49
Philippines	0.46
China	0.45
Thailand	0.43
United States	0.41

Country	Gini Coefficient in 2005
United Kingdom	0.36
France	0.33
Russian Federation	0.31
Ethiopia	0.30
Albania	0.28
Hungary	0.27
Sweden	0.25

Source: “Inequality in Income—Selected Countries and Regions,” *United Nations 2005 Human Development Report*, figure 3, accessed January 30, 2011, http://hdr.undp.org/en/reports/global/hdr2005.

When we compare countries, remember that some countries have much higher income than others. Looking at Table 13.2.6 "Gini Coefficients in Different Countries", low-income countries generally seem to have more inequality than high-income countries. This is suggestive of a link between inequality and stages of development. Economist Simon Kuznets suggested that inequality would increase in the early stages of the development process but decrease in later stages. This became known as the *Kuznets hypothesis*. One story was that as a country grows, the labor force is split between a relatively high-income industrial sector and a relatively low-income agricultural sector. As a country grows, more labor is allocated to the more productive manufacturing sector, and thus inequality is reduced over time.

Whatever the mechanism, world inequality appears to be decreasing significantly. A recent study found that the Gini coefficient for the world had declined from about 0.58 in the 1970s to about 0.51 in the late 2000s.See Maxim Pinkovskiy and Xavier Sala-i-Martín, “Parametric Estimations of the World Distribution of Income” (National Bureau of Economic Research Working Paper 15433, October 2009), accessed March 14, 2011, http://www.nber.org/papers/w15433.pdf.

There are also some fascinating differences in the dynamics of inequality. The decline in inequality in the middle of the 20th century was common throughout much of the developed world. The more recent increase in equality that we have documented in the United States is also visible in some other countries, such as Australia, New Zealand, and the United Kingdom. By contrast, most of Western Europe has not seen the same kinds of increases in inequality.

Key Takeaways

- The Lorenz curve shows the distribution of income in an economy by plotting the fraction of income on the vertical axis (after households have been ranked by their income) and the fraction of the population on the horizontal axis. The closer the Lorenz curve to the 45-degree line, the more equal the distribution of income.
- The Gini coefficient is a statistic that indicates the degree of inequality by looking at how far the Lorenz curve is from the 45-degree line.
- A given household’s position in the distributions of income, wealth, and consumption changes over time. This is partly due to education and work experience and partly due to luck. Another dynamic element of the income distribution comes from transfers across generations of a household.

Exercises

1. If you have two countries, what does it imply about the Lorenz curves for the two countries if the Gini coefficient on income is higher in the first country compared to the second?
2. Is it possible for disposable income to be distributed more equally across households in a country than market income? How could this happen?
3. How do taxes influence the distribution of disposable income?

This page titled 13.2: Facts about Inequality is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

13.3: The Sources of Inequality 13.3: The Sources of InequalityEdit Update titlecancel

Learning Objectives

1. Where do differences in income come from?
2. Why might the marginal product of labor differ across people?
3. What is the skill gap?
4. What is a winner-takes-all market?

We have provided some facts about differences in income across households. We now turn to a discussion of where those differences come from.

From Ability to Earnings

We begin by looking at earnings, by which we mean the income that households obtain from their work in the **labor market**. Figure 13.3.1 shows the labor market. The **real wage** is on the vertical axis, and the number of hours worked is on the horizontal axis. The labor demand curve indicates the quantity of labor demanded by firms at a given real wage. As the real wage increases, firms demand less labor. The labor supply curve shows the total amount of labor households want to supply at a given real wage. As the real wage increases, the quantity of labor supplied also increases. See Chapter 4, Chapter 8, and Chapter 9 for more discussion. Here we are interested in what the labor market can tell us about how much people earn.

Toolkit: Section 31.3

You can find more details about the labor market in the toolkit.

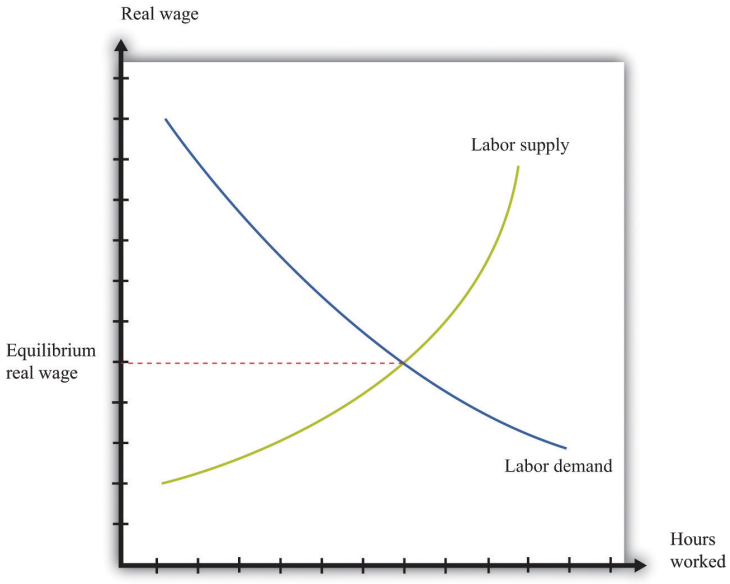


Figure 13.3.1: Labor Market Equilibrium

When firms are deciding how many hours of work to hire, they use this decision rule: hire until

$$real\ wage\ =\ marginal\ product\ of\ labor.$$

The left side of this equation represents the cost of purchasing one more hour of work. The right side of this equation is the benefit to the firm of one more hour of work: the marginal product of labor is the extra output produced by the extra hour of work. If the marginal product is higher than the real wage, a firm can increase its profits by hiring more hours of work.

We use this equation as a starting point for thinking about distribution and inequality. Different individuals in the economy are paid different real wages. This reflects, among other things, the fact that there is not a single labor market in the economy. Rather, there

are lots of different markets for different kinds of jobs: accountants, barbers, computer programmers, disc jockeys, and so on. We can imagine a diagram like Figure 13.3.1 for each market. In all cases, the firms doing the hiring will want to follow the rule given by the equation. And if firms follow this hiring rule, then two individuals who earn different real wages must differ in terms of their marginal product. The worker who earns the higher wage is also the worker who is more productive.

But why would workers have different marginal products? One reason is that people differ in terms of their *innate abilities*. For any individual, we could come up with a long list of the skills and abilities that he or she is born with—natural talents. Some are good at mathematics, some are particularly strong, some are good at music, some are good at building things, some are very athletic, some are good at managing other people, and so on. Abilities that tend to make someone have a high marginal product allow that person to earn higher real wages. Differences in innate abilities, then, are the first explanation we can suggest for why there are differences in earnings when we look across individuals.

The possession of innate ability is not enough to guarantee someone a high marginal product; the market must value the individual's talents as well. The demand for particular abilities or skills is high if they can be used to produce something that people want to buy. Think about a talented quarterback: his talents translate into an ability to draw paying customers to games, which in turn translates into a willingness to pay a lot for his labor. Or think about a skilled manager: her ability to make good business decisions translates into higher profits for a firm, which in turn translates into a willingness to pay for her labor. If an ability is valued in the market, then there will be high demand for the labor of people with that ability.

What is valuable changes over time and from place to place. Being a skilled quarterback is valued in the modern-day United States. The same innate talent was worth much less 50 years ago in the United States and is still worth little today in a village in the Amazon. Rock stars who can earn hundreds of millions of dollars today would have had very little earning power in 19th-century Australia. The same holds for more mundane skills. The innate abilities that make for a good software designer are more valuable than in the past; the innate abilities that make for a good clockmaker are less valuable than in the past.

Labor supply matters because the value of your innate abilities also depends on how many other people have similar talents. Another reason that highly talented quarterbacks command such high earnings is because their abilities are in short supply. Being a good taxi driver also requires certain skills, but these are much more common. As a result, the supply of taxi drivers is larger, so the real wage earned by taxi drivers is smaller.

Education, Training, and Experience

Star quarterbacks have innate abilities that most of us don't possess. But they also have more training and experience in this role. Just about every one of us could be a better quarterback than we are now, if we were willing to train several hours a day. Indeed, most occupations require some skills and training. Computer programmers must learn programming languages, engineers must learn differential equations, tennis players must learn how to play drop shots, and truck drivers must learn how to reverse an 18-wheeler.

As well as such specific skills, an individual's general level of education is usually an indicator of his or her marginal productivity and hence the wage that can be earned. Basic literacy and numeracy are helpful—if perhaps not absolutely necessary—for nearly any job. A high school education typically makes an individual more productive; a college education even more so. So the distribution of labor income is affected by the distribution of education levels. People also learn on the job. Sometimes this is through formal training programs; sometimes it just comes from accumulating experience. Generally, older and more experienced workers earn higher wages.

Education and experience affect both labor demand and labor supply. More highly skilled workers are typically more valuable to firms, so the demand curve for such workers lies further to the right. At the same time, experienced and trained workers tend to be in more limited supply, so the supply curve lies further to the left. Both effects lead to a higher real wage. Just as a worker's real wage depends on how valuable and scarce are her abilities, so also does it depend on how valuable and scarce are her education and training.

The influence of experience on earnings is a reminder of an observation that we made when discussing the data. Even in a world where everyone is identical in terms of abilities and education, we would expect to see some inequality in earnings and income simply because people are at different stages of life. Younger, inexperienced workers often earn less than older, experienced workers.

The Skill Gap

In recent years, economists have looked closely at the differences in wages among skilled and unskilled workers. Loosely speaking, skilled workers are more educated and in occupations that rely more on thinking than on doing. So for example, an accountant is termed a skilled worker, and a construction worker with only a high-school diploma is an unskilled worker. Data on wages suggest that the return to skill, as measured by the difference in wages between skilled and unskilled workers, has widened dramatically since the mid-1970s. Many economists think that this is an important part of the explanation for the increasing inequality in the United States.

One way to measure the increased return to skills is to look at the financial benefit of education, given that more educated workers are typically skilled rather than unskilled. Table 13.3.1 "Relationship between Education and Inequality in the United States" summarizes some evidence on the distributions of earnings, income, and wealth from 1998. The table indicates that there is a sizable earnings gap associated with education. According to this sample, completing high school increased earnings by nearly $20,000, and a college degree led to an additional $34,000 in average annual income. Education is an important factor contributing to inequality. One way to decrease inequality is to improve access to education.

Table 13.3.1: Relationship between Education and Inequality in the United States

Education	Earnings	Income (1998 $)	Wealth
No high school	14,705	21,824	78,548
High school	34,211	43,248	189,983
College	68,530	88,874	541,128

Source: Santiago Rodríguez, Javier Díaz-Giménez, Vincenzo Quadrini, and José-Víctor Ríos-Rull, “Updated Facts on the U.S. Distributions of Earnings, Income and Wealth,” *Federal Reserve Bank of Minneapolis Quarterly Review*, Summer 2002. Here earnings come from both labor and business activities. Income includes transfers.

Effort

So far we have said nothing about how hard people choose to work, in terms of either the number of hours they put in on the job or their level of effort while working. Those who are willing to work longer hours and put in more effort will typically obtain greater earnings.

Effort is a matter of individual choice. Some other factors that can influence your earnings are likewise under your own control. Training and education are largely a matter of choice: you can choose to go to college or take a job directly out of high school. By contrast, the abilities you are born with are, from your point of view, a matter of luck. We have more to say about this distinction later when we evaluate the fairness of the distribution of income.

The Gender Gap

Study after study indicates that the gender of a worker also influences real wages. Figure 13.3.2 shows the wage gap and the participation rates for married women in the United States.We are grateful to Michelle Rendell for this figure. The discussion in this section is drawn in part from her PhD dissertation research. The participation rate for married women—the fraction of married women in the labor force—has increased from slightly above 20 percent in 1950 to about 70 percent in 2000. Meanwhile, the ratio of wages paid to married women relative to married men displays an interesting pattern over this period. From 1950 to 1980, the ratio fell from 65 percent to 60 percent—that is, the wages of married women fell relative to married men. Thereafter, the ratio rose substantially, to about 80 percent in 2000. At the end of the 20th century, in other words, married women were earning about four-fifths of the wages of married men.

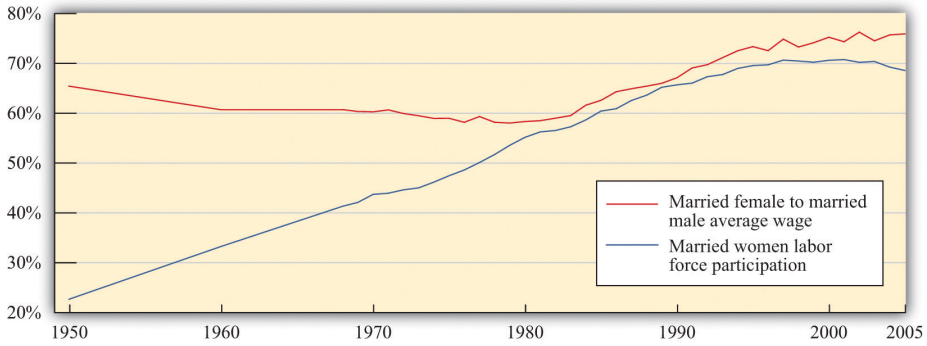


Figure 13.3.2: Labor Market Outcomes for Women

Economists and other social scientists are interested in understanding these facts. What was the source of the increased participation in the labor force by women and what factors increased their wages relative to men? One tempting approach is to use a **supply-and-demand** diagram like Figure 13.3.1, thinking specifically about women's labor. For example, we could explain the overall shift between 1950 and 2000 by a rightward shift of the labor demand curve. A shift to the right in the demand curve increases the real wage. The higher real wage would also induce women to supply more hours: this is the corresponding movement along the labor supply curve. More women would be induced to move away from work at home and toward work in the market, given the higher return for market work. To explain the increase in women's wages *relative* to men's, we would need to see a larger increase in the demand for women's labor than for men's labor.

But this is a somewhat odd story. There is no reason to think that there should be a separate labor market for women and men. Women and men can and do perform the same jobs and thus compete in the same labor market. Any supply-and-demand explanation needs to be subtler. One possibility is that there has been a shift in the kinds of jobs that are most important in the economy and hence a shift in the kinds of skills needed. Suppose, for example, that women are more likely to be accountants than construction workers. A shift in labor demand toward accountancy and away from construction will increase wages in accountancy relative to construction work and will therefore increase women's wages, on average, relative to men's. Researchers looking closely at the data see some evidence of such effects when they look at wages and employment patterns across jobs that require different skills.

There is another, perhaps even more basic question: why are women's wages consistently lower than men's wages? Researchers have also devoted a great deal of effort to this problem, looking to see in particular if differences in education and skills can account for the difference in wages. Typically, these studies have found that such differences can explain some—but not all—of the gap between wages for men and women. The remaining difference in wages is very possibly due to discrimination in the labor market. If this is the case, then recent increases in women's wages relative to men's wages could be due to a reduction in discrimination.

Of course, women are not the only group that has been subject to discrimination in the labor market. In the United States, African Americans and other minority groups have suffered from discrimination. In many other countries, there are similarly different groups that have been unfairly punished in the labor market. Economists point out that supply and demand is actually a positive force for combating discrimination. Discrimination against women workers, for example, means that women are being paid less than their marginal product. Nondiscriminatory employers then have an incentive to hire these workers and make more profit, which in turn would tend to increase women's wages.

Economic forces can mitigate discrimination, but this is not an argument that discrimination is not or cannot be a real problem. First of all, discriminatory attitudes might make employers incorrectly perceive that the marginal product of women (or other groups) is lower than it actually is. Second, even if employers are not actively discriminating against women, coworkers may be discriminatory, and this could lead to lower productivity among women in the workforce. Research in social psychology tells us that such discrimination—by employers or colleagues—can occur even if people have no explicit discriminatory intent.

Winner-Takes-All Markets

There are some markets where compensation reflects ability in a very extreme way. These are often called **winner-takes-all markets**. In such a market, the person with the highest ability captures the whole market, and everyone else gets nothing. You can think of this as a race where the winner of the race gets all the prize money. The phrase *winner takes all* is not meant literally. The idea is more that a small number of people earn very large returns. Think, for example, of the professional golf or tennis circuits,

where perhaps a few hundred people obtain the winnings from the tournaments—and the bulk of the winnings go to a small number of top players.

In these markets, we cannot assume that the wage equals the marginal product of labor. In a winner-takes-all market, you get a wage that depends not on your productivity in isolation but on how your productivity compares with that of others. If you are the most productive, you win the entire market.

Many markets have at least some aspects of a winner-takes-all market. Think of the market for rock musicians. If there were one group that everyone liked more than all the others, then that group would sell CDs and MP3s, give concerts, and completely dominate the music scene. Other groups would disappear. The actual music market is not this extreme. There are many groups who produce songs, give concerts, and so on. But there is a clear ranking between the first-class groups and the others. So even though there is not a single winner who takes all the market, there are a relatively small number of big winners who together take most of the market.

Why does the market for rock musicians have winner-takes-all characteristics? A good way to understand the phenomenon is to think about the market for musicians centuries ago—before recording technologies. Good musicians might still be rewarded well—perhaps they would play for the king or queen—but there was room for, relatively speaking, a large number of good musicians because each would be serving only a relatively small local market. Today, though, the very best musicians can record their music and sell it all around the world. A single group, at relatively low marginal cost, can serve a very large market. (This is particularly true for CDs or MP3 files. It is less true for concert appearances because these do not have such low marginal cost.)

In winner-takes-all markets, there is a very skewed distribution of income relative to ability. Small differences in ability can translate into substantial differences in income. Moreover, winner-takes-all forces may be becoming stronger as a result of technological advances. The most popular rock stars, sport stars, and movie stars are now worldwide celebrities. Lady Gaga is famous in Thailand and Toledo; Brad Pitt is known from Denver to Denmark. This is perhaps one reason the very rich are getting relatively richer.

From Income to Consumption and Wealth

We are interested not only in the distribution of income but also in the distribution of consumption and wealth. To connect these three, we use the following equation: See Chapter 5 for more discussion.

$$wealth\ next\ year\ =\ (wealth\ this\ year\ +\ income\ this\ year\ -\ consumption\ this\ year)\ \times\ interest\ factor.$$

The first term on the right-hand side is the wealth you have at the start of a given year. To this wealth you add the income you earn in the current year and subtract your consumption. Because $income\ -\ consumption\ =\ savings$, this is the same as saying that you add your savings to your wealth. You earn interest income on your existing wealth and your new savings. Your initial wealth plus your savings plus your interest income gives you the wealth you can take into next year.

Suppose you currently have $1,000 in the bank. This is your wealth this year. You receive income of $300 and spend $200 of this income. This means that you save $100 of your income. So wealth this year plus income this year minus consumption this year equals $1,100. With an interest rate of 5 percent, your wealth next year would be $\$1,100\ \times\ 1.05\ =\ \$1,155.$

This equation tells us several things.

- Wealth, income, and consumption are interconnected. A household's decisions about how much it wants to save and how much it wants to consume determine what its consumption and wealth will look like. Imagine two otherwise identical households with different preferences about consuming this year versus the future. The impatient household consumes a lot now and saves little. It has high consumption early in life, low consumption later in life, and relatively low wealth. A more patient household has a very different pattern of wealth and consumption. It has lower consumption early in life, higher consumption later in life, and higher wealth on average.
- Differences in earnings cumulate over time to generate a distribution of wealth. High-ability households are more productive and thus earn more income. Some of this income is saved, and the rest consumed. Higher-income households thus tend to have higher wealth than lower-income households because the higher-income households have higher levels of saving each year.
- Inherited wealth can be a source of differences in income and consumption. Some individuals start out their working lives as beneficiaries of inheritances from their parents (or others). These people can enjoy higher consumption. They also obtain more income in the form of interest earnings on their wealth.

The equation also *conceals* at least one relevant fact for inequality: wealthier households typically enjoy higher returns on their wealth. The interest rate is not the same for all households. There are several reasons for this, such as the fact that richer individuals find it worthwhile—and can afford—to hire professionals to manage their portfolios of assets or the fact that richer people may be able to purchase assets that are riskier but offer higher returns on average. It is not surprising that, as we saw, the wealth distribution is more unequal than the income distribution.

Figure 13.3.3 brings together all the ideas we have discussed so far. It shows us three things. (1) Discrimination and winner-takes-all situations can break the simple link between the marginal product and the wage. (2) Government policies can break the simple link between wages and income. (3) Household decisions about how much to consume and save affect the observed amounts of income, consumption, and wealth. The figure also makes it clear that some of the forces leading to inequality are under the control of the individual, while others are outside the individual's control.

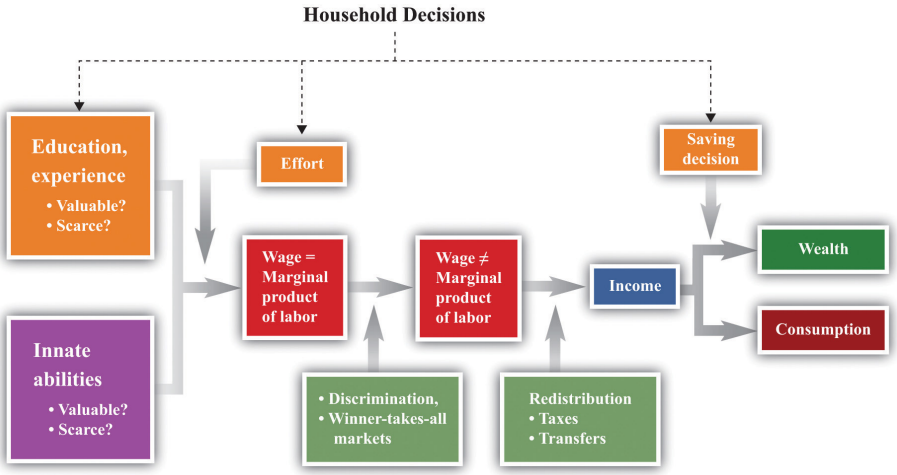


Figure 13.3.3: The Different Sources of Inequality

Key Takeaways

- Differences in income can reflect, among other things, differences in ability, education, training, and gender.
- Wage differences across people reflect differences in marginal products across people.
- The skill gap shows the differences in earnings from differences in education. This gap has widened in recent years.
- In a winner-takes-all market, the most talented individual captures all (or almost all) of the market.

exercises

1. Draw two versions of the labor market: one for lawyers and one for taxi drivers. How would you use these labor markets to explain the differences in labor income between lawyers and taxi drivers? Are these two labor markets related in any ways?
2. Does Figure 13.3.2 imply that as more women participate in the market, there are increases in the ratio of wages earned by women relative to men?
3. Where do differences in wealth come from?

This page titled 13.3: The Sources of Inequality is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

13.4: Distributive Justice 13.4: Distributive JusticeEdit Update titlecancel

Learning Objectives

1. What is the evidence from economic experiments about “fairness”?
2. What are some of the leading theories about “fairness”?

So far we have described some facts about inequality in the United States and the world, and we have offered some explanations of why we observe these inequalities. In this section, we take a more philosophical perspective on the distribution of income and wealth. We ask questions of a kind that economists generally ignore, such as the following: “Is the distribution of income fair?”

As you might expect, questions like this are extremely contentious. Different people have very different ideas about what is fair and just, and this topic is highly politicized. It is not our job, nor is it our intention, to tell you what is and is not fair. What we can do is give you a (very brief) introduction to some of the ways that philosophers, economists, political scientists, and others have thought about these very hard questions. More particularly, we can give you some “thought experiments” to help you determine your own views on these topics. Hundreds of books have been written on these issues, however, so we simply scratch the surface here.

Experimental Evidence on Fairness

Noneconomists frequently speak about a “fair wage” or a “fair price” for a particular product. To economists, this language is unfamiliar, even confusing. Economics provides no theory about what is fair or unfair; it gives us no basis to ask whether particular prices in the economy are fair.

Yet ideas about fairness motivate people in many economic transactions. As one example, some people are willing to pay extra for “fair trade” goods, such as coffee or chocolate bars. The idea of these goods is that the seller makes some guarantees about payments to producers, working conditions, or other variables that are not intrinsic to the good itself. As another example, people are often willing to take part in boycotts, meaning that they voluntarily forgo a good that they like to send a message to the producer of the good.

Experimental economists have conducted many studies to try to understand some of these ideas of fairness. Sometimes they have used a **dictator game**. This game has two players. Player A, the dictator, is given a sum of money and decides *how much of that money to give* to player B. Player B keeps the money he is given, and player A keeps the rest. From the perspective of economic reasoning, this game is completely trivial. Suppose you are the dictator, and you are given $100 to allocate. The self-interested thing to do is to keep all the money for yourself and give nothing to player B.

Yet study after study has shown that people typically give away some of their money, often dividing it up in equal shares. You may be able to think of several reasons why people behave this way. Perhaps they are worried about what the other person will think about them. Perhaps they are worried about what the *experimenter* will think about them. Researchers have gone to great lengths to design studies where no one except player A can possibly know her decision. Even in this case, most people do not keep all the money. It is hard to dismiss the view that people’s decisions are motivated in some way by what they think is the fair thing to do.

A related but slightly richer game is known as the **ultimatum game**. It also has two players. Player A is given a sum of money and then decides *how much of that money to offer* to player B. Player B then decides whether to *accept or reject* player A’s offer. If player B accepts that offer, he keeps the amount offered, and player A keeps the rest. If player B rejects the offer, then both player A and player B receive nothing.

Toolkit: Section 31.18

You can read more about these games and others in the toolkit.

The difference between the ultimatum game and the dictator game is that player B has the right to veto the offer. If he vetoes the offer, then both players get nothing. Economic theory again has a simple prediction about what completely self-interested players will do. Player B is better off accepting *any* positive offer than he is rejecting the offer. Suppose player A starts with $100 and offers $1 to player B. If player B accepts, he gets $1. If he rejects, he gets $0. Because $1 is better than nothing, player B should accept the offer. Knowing this, player A should offer the smallest amount possible. For example, if player A has $100 to allocate, she should offer $0.01. Player B should accept the offer ($0.01 is bigger than $0.00), and player A will then end up with $99.99. In fact, this is not what usually happens. People in the role of player A typically offer much more than the minimum amount. One

reason is the risk that if player B is made a stingy offer, he will reject it out of spite. Another reason, like in the dictator game, is that people may care about fairness when making their offers. The evidence suggests that both factors seem to matter in this game.

Hundreds of studies have been conducted using different variants of these two games. The big conclusion from all these studies is that people seem to be motivated by more than just narrow self-interest when they play games such as these in the laboratory. Instead, they care about allocating the rewards from the experiment in a way that is fair. Understanding exactly what underlies these ideas of fairness is an exciting area of research in experimental and behavioral economics, as well as in psychology and other disciplines.

Meritocracy

We begin with a very simple framework. Imagine an economy in which there are two kinds of people: high ability and low ability. Half the people in the economy are high ability: they can produce 100 chocolate bars in a year. The other half are low ability: they can produce only 50 chocolate bars in a year. In this economy, productivity and ability are the same thing. High-ability people are more productive than low-ability people. We use this simple economy to think about different approaches to the allocation of society's resources.

Libertarianism

One view of distribution is summarized by the statement "you are entitled to whatever you earn." In this world, the distribution of income and consumption will be the same as the distribution of output. High-ability people have income of 100 chocolate bars. If our fictional economy were to last for only one year, their consumption would also be 100 chocolate bars. Similarly, low-ability people will have income and consumption of 50 chocolate bars. This economy has an unequal distribution of income and consumption.

If we were to associate this position with a particular philosophical school of thought, it would be *libertarianism*. Libertarians generally believe that people are entitled to whatever they can earn, the state should intervene as little as possible, and the state should not actively seek to redistribute resources. The fact that there is inequality in this society is simply a reflection of differing abilities, which is not any reason for the government to get involved. (To be clear, libertarians have no objection to people making charitable contributions. If the high-ability people in the economy wanted to voluntarily give money to the low-ability people, libertarians would have no complaint about this.)

Consumption, Saving, and Insurance

Now let us consider a slight variant on this economy. Suppose the economy lasts for two periods: in each period, every individual has a 50-50 chance of being either high or low ability. If we measured income in either period, we would see the same amount of inequality as before.

Consumption, however, is a different story. Suppose you are a high-ability person in the first period. You know that you face a risk of being low ability in the second period. Should you eat your entire 100 chocolate bars in the first period? Most people prefer to keep their consumption at least somewhat smooth, so they will "save for a rainy day." We expect that high-income people in this economy will consume less than their income.

Similar reasoning applies to low-ability people. They earn only 50 bars in the first year but have a 50-50 chance of higher income next year. By the same consumption-smoothing argument, they would like to somewhat increase consumption today. Thus low-ability people will consume more than their income in the first period. There will be a **credit market (or loan market)** in which high-income people lend money to low-income people in the first year, and those loans are repaid the following year.

> **Toolkit: Section 31.6**
>
> You can review the idea of the credit market in the toolkit.

This example of borrowing and lending driven by the desire for smooth consumption affects the distributions of income and consumption. Economic theory tells us that consumption will be more equal than income. This is consistent with the evidence: consumption is indeed more evenly distributed than income. Again, believers in a libertarian philosophy would see no reason for any intervention in this economy.

If this economy were more sophisticated, it might even develop an insurance market. All the individuals in the first year would recognize that their future income was uncertain. If they are **risk-averse**, then they would all prefer to eliminate this uncertainty. Being risk-averse means you prefer the average of a gamble to the gamble itself. Suppose a person is faced with a choice between

- 100 chocolate bars with a **probability** of 0.5 and 50 chocolate bars with a probability of 0.5
- 75 chocolate bars

A risk-averse person prefers the option that delivers 75 chocolate bars with certainty. The first option also yields 75 chocolate bars on average (more technically, it has an **expected value** of $0.5 \times 100 + 0.5 \times 50 = 75$), but this option has uncertainty that risk-averse people will want to avoid. In this economy, there would be some redistribution of income in the second year. However, it would be a voluntary redistribution based on the insurance contract that everybody agreed to in the first year. Again, there would be no role for government.

Toolkit: Section 31.7

You can review the concepts of probability and expected value in the toolkit.

The Rawlsian Veil of Ignorance

One of the most famous approaches to the questions of fairness and justice was pioneered by the philosopher John Rawls in his celebrated book, *A Theory of Justice*.John Rawls, *A Theory of Justice* (Cambridge: MA, Harvard University Press, 1971). Rawls's work is rich, complicated, and much debated, and the presentation here is very simplified and stylized. For example, Rawls focused more on the institutions that people behind the veil would want, rather than on the actual distribution of income. Rawls introduced a powerful thought experiment to help people decide how they feel about different distributions of society's resources.

It is difficult for any of us to think about redistribution without framing it in terms of our own personal circumstances and interests. Rawls's idea was designed to help us shed those considerations. He proposed thinking about redistribution from behind a veil of ignorance. Behind this veil, you know what the distribution of resources and abilities will look like in society, *but you do not know where you will be in this distribution*. You might be born rich, or you might be born poor. You could end up as Bill Gates, or as a homeless person in New York. If you want to play this game globally, you might end up as a member of a royal family in Europe or as someone scavenging for food on a garbage heap in Cambodia. If we frame this in terms of our previous example, then, behind the veil of ignorance, you know that 50 percent of the people will be high ability, and 50 percent will be low ability, but you do not know which you will end up being.

Now suppose that decisions on how to allocate chocolate bars across households are made before people know whether they are high or low ability. Rawls suggested that people behind the veil would adopt a **social contract** in which they agree to the following.

- Once born, they will produce chocolate bars according to their ability and then put the chocolate bars they produce into a big pile.
- Each individual will take out an equal share of 75 bars.

This contract involves taxation and redistribution. The high-ability people are each taxed 25 bars, and the low-ability people receive a transfer of 25 bars. The taxes are sufficient to finance the transfers.

Figure 13.4.1 shows a taxation and transfer scheme that could be used with this social contract. On the horizontal axis is production, which is income. On the vertical axis is the tax paid by each income group. With this scheme, anyone with income above 75 bars pays a tax of 25 bars. Anyone with income below 75 bars gets a transfer of 25 bars. Because there are an equal number of high- and low-ability households, taxes collected equal transfers. The government's budget balances.

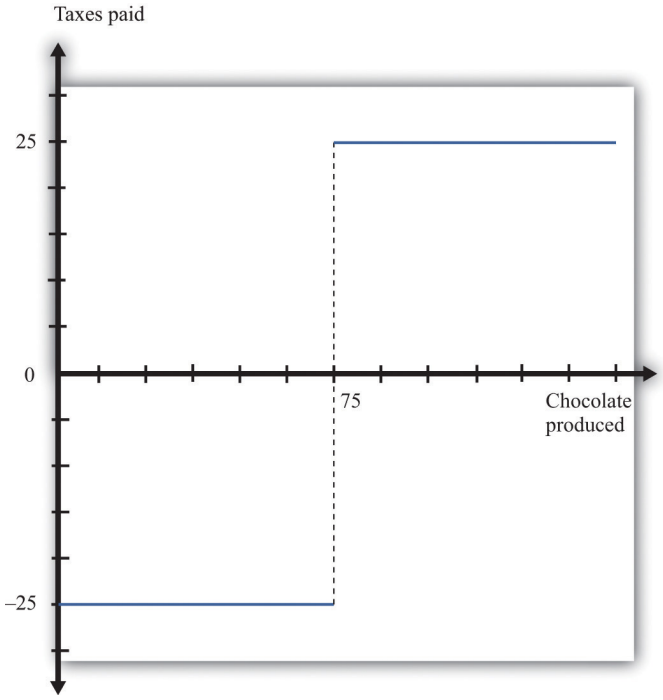


Figure 13.4.1: Taxes and Transfers in a Rawlsian Social Contract. With this tax scheme, households producing more than 75 chocolate bars pay a tax of 25 bars, and those producing fewer than 75 bars receive a transfer (negative tax) of 25 bars.

Because everyone is risk-averse, all will prefer this deal to the allocation that gave the high-ability people 100 bars and the low-ability people 50 bars. Though additional chocolate bars are not produced, the redistribution of the contract is preferred to everyone *before they know their ability*. The key, emphasized in the previous sentence, is that the contract is agreed on before people know their ability. Because of this timing, the risk sharing through the redistribution of the chocolate bars makes everyone better off, compared to the—imaginary—initial condition.

You have almost certainly noticed that this Rawlsian social contract very closely resembles the insurance contract that we described in 13.3 Section "Meritocracy". In effect, Rawls suggested that people behind the veil would want to write the same kind of insurance contract that they would write in a similar situation in real life. But because we obviously can't write contracts before we are born, Rawls thought that we should agree to government policies that would mimic these kinds of insurance contracts. Notice that, in the Rawlsian world, the distribution of income has a higher Gini coefficient than does consumption. In fact, in this example, there is no inequality in consumption.

From Each According to His Ability and to Each According to His Needs

Karl Marx, the famous philosopher and social theorist, suggested that society should distribute its resources as follows: "From each according to his ability, to each according to his needs." Marx's prescription recognizes that individuals differ in their ability to produce and in their consumption needs. He said that workers should produce at a rate commensurate with their ability, so high-ability individuals would be expected to produce more output than low-ability individuals.

In the Marxian view, there is a complete disconnect between production and consumption. There is no sense that those who *produce* more of society's resources should be entitled to *consume* more of those resources. It stands in complete contrast to the libertarian view that individuals have a right to whatever they produce. The distribution of production is independent of the allocation of income and consumption.

How would the Marxian view work in our chocolate bar economy? "From each according to his ability" means simply that the high-ability individuals should produce 100 chocolate bars and the low-ability individuals should produce 50 chocolate bars. Meanwhile, "to each according to his needs" means that the total number of chocolate bars produced in the economy ought to be allocated in a way that reflects the needs of the individuals. In our simple example, individuals do not differ in their **valuation** of a chocolate bar. All individuals like chocolate bars the same amount. Therefore, the allocation that satisfies the Marxian prescription is that everyone should have the same number of chocolate bars.

In our simple example, Marx and Rawls agree on how to allocate chocolate bars. We can imagine, however, ways in which individuals might differ in terms of their needs. For example, some people are fortunate enough to be healthy and fit, while others suffer from illness or disease. A Marxian prescription would allocate more of society's resources to the sick, on the grounds that their needs were greater. (It is also possible, of course, that people behind the Rawlsian veil of ignorance would make a similar allocation.)

Luck versus Merit

In all of our examples so far, we have supposed that people differed only in terms of their abilities, which are—by assumption—completely outside their control. In our earlier discussion of the sources of inequality, however, we listed many different possible reasons why people might have different earnings. Some of these factors were outside any individual's control; others were not. Table 13.4.1 "Luck versus Merit" provides a partial listing.

Table 13.4.1: Luck versus Merit

Outside an Individual's Control	Within an Individual's Control
Innate abilities	Effort and hours worked
Demand for these abilities	Education (in part)
Supply of these abilities by others	Experience and training
Discrimination	Consumption/saving decisions
Inherited wealth	

An individual does not control his or her basic abilities. Some are lucky, possessing the abilities that allow them to be great basketball players, pianists, authors, or scientists. Abilities that are scarce are likely to be more valuable. The value associated with a particular set of abilities is also heavily dependent on time and place—for example, being a great rock drummer would not have been worth much in the Roman Empire, and an ability to throw a spear hard and accurately is not especially valuable in modern-day San Francisco. All of these come down to luck when viewed from the perspective of any individual.

We have hinted at many other factors that are also a matter of luck. Those born of wealthy parents in wealthy countries are likely to attend high-quality schools and receive inherited wealth. They may also be able to earn higher real interest rates on their savings. Meanwhile, those who are subject to discrimination will earn lower incomes.

There are also many factors that influence the distribution of income, consumption, and wealth that *are* under the control of an individual. Individuals can choose how hard to work and how many hours to work. They can choose whether to sacrifice current earnings to go to college. They can decide to go back to school to earn a master's degree. They can choose careers that allow them to develop skills and experience on the job.

Why does this distinction matter? Most people would agree that there is little or no problem with inequalities that result from people's choices. There is nothing self-evidently unfair about one person having a higher income than another because he works harder or chose to take time off from work to pursue a graduate degree. But opinions differ much more about the fairness of inequalities that result from luck or chance. Tiger Woods is an immensely talented golfer, but is it fair that he should earn so much on the basis of his genetic luck? Is it fair that someone who struggles in school and possesses little in the way of valuable skills should earn only minimum wage? These are not questions that we can answer, but thinking about these questions should help you form your own opinions on what is a fair and just distribution of society's resources.

Equality of Opportunity versus Equality of Outcome

The distinction between luck and merit gives us a more nuanced view of equality. It is closely related to another distinction that is often made when discussing the distribution of society's resources: equality of opportunity versus equality of outcome. Here is an example to help make the distinction clear.

At major soccer tournaments, such as the World Cup, teams often line up behind banners proclaiming "fair play." The international soccer association, FIFA (Fédération Internationale de Football Association), places a lot of emphasis on this idea. Fair play means that players should always play within the rules, and these rules provide equality of opportunity on the soccer field. At the start of any game, both teams line up with the same number of players, try to score in the same sized goal, and enjoy the benefits of

impartial referees. This does not mean that soccer games always end in a tie: FIFA's rules do not mean that there is equality of outcome. The outcome depends on the two teams' abilities. So although the opportunity to win is shared equally by the teams, the outcome is not: the winner takes all.

Equality of opportunity without equality of outcome is pervasive in the economy as well. Institutions exist to enhance equality of opportunity with no guarantees about outcomes. For example, going to a public school is an option for everyone (though there are significant differences across schools in terms of their quality). But there is no guarantee that two people graduating from the same school will have the same outcome. When you apply for a job, you have an opportunity to compete along with anyone else for that job, but the outcome is different for the person who is hired compared to those who are not.

It is tempting to identify equality of opportunity with the view that merit should be rewarded but luck should not. There is certainly a connection. Both imply that discrimination should not affect the distribution of income in the economy. But equality of opportunity still allows those with high abilities to get higher rewards, even though those abilities are a matter of luck. If your college soccer team were to play Real Madrid, either team would have the chance to win the game, according to the rules. That equality of opportunity would be of little consolation to your team's goalkeeper as he picked the ball out of the net for the 20th time.

Yet there is one very good reason why equality of opportunity is so important. Imagine what would happen if FIFA started instructing referees to ensure that every soccer game ended in a draw. To ensure equality of outcome, the referee would alter the rules of the game to help the side that was losing. Fair play would be gone, together with lots of other things: teams would have no incentive to play hard, they would have no incentive to find quality players, and fans would not enjoy the game as much. We get the best from a team because it knows that if it performs well, under the rules, it will win and receive financial and emotional rewards. These provide the incentives for team members to train and play hard, within the rules of the game. Combining equality of opportunity with the ability to compete for a prize strikes the right balance—at least for soccer—between equality and incentives. In the next section, we will examine why incentives matter so much for decisions about redistribution.

Key Takeaways

- In experimental bargaining games, players seem to be motivated by more than narrow self-interest. In many cases, they give money to the other player.
- Equality of opportunity argues that everyone should have an equal chance of succeeding without guaranteeing that success. It contrasts with the view that everyone should work as hard as they can, and goods and services should be allocated according to need.

Exercises

1. Does the grading in your economics class exhibit equality of opportunity? Why or why not?
2. If you think about the allocation of resources within a household, which of the theories of distributive justice best applies?

This page titled 13.4: Distributive Justice is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

13.5: Government Policy 13.5: Government PolicyEdit Update titlecancel

Learning Objectives

1. What actions does the government take to influence the distributions of income, wealth, and consumption?
2. What is the rationale for these government interventions?
3. What limits the effects of government redistribution?

Governments play a significant role in the distribution of income, consumption, and wealth. The argument for government intervention usually takes the form that the market outcome is too inequitable, relative to, for example, a Rawlsian view. We now look at various forms of redistribution through government actions, paying particular attention to their effects on incentives.

Incentives

Redistribution is more than setting **taxes** and **transfer payments** to give money from one person to another. The problem is that redistribution can affect people's incentives in various ways.

The Incentive to Be Truthful

Go back once more to our chocolate bar economy. We proposed a scheme whereby high-ability individuals would be taxed 25 chocolate bars, with this being paid to low-ability individuals. A tax-and-transfer scheme of this kind would allow us to achieve the equitable outcome mandated by the Rawlsian or Marxian view.

Low-ability households evidently have an incentive to participate in this scheme: they give up 50 bars and get back 75 bars. The redistribution is in their favor. The story is different for high-ability people. They give up 100 bars and get 75. Before abilities are known, everyone likes this social contract. But once ability is known, high-ability people prefer not to participate. If they can produce and then hide some of their chocolate bars, they have an incentive to

- produce 100 chocolate bars,
- pretend to be a low-ability person and declare production of 50 bars, and
- take the transfer of 25 bars and consume 125 chocolate bars.

High-ability people can get away with this if chocolate bar production cannot be monitored. They have an incentive to rip off the system by pretending to be low ability. Because all high-ability people behave this way, the contract will fail: no one will pay taxes, and everyone will demand a transfer.

In this extreme example, the incentive problem completely destroys the redistribution policy. In reality, there is some redistribution through taxes and transfers because the government, acting through the taxation authority, is able to tax households at different rates: low-income households face lower tax rates than higher income households. In addition, low-income households receive transfers from the government. Governments can carry out such policies because they have access to information about the income households earn. Yet incentive problems like the one we have outlined pose very real difficulties for governments. Rich people have an incentive to hide their true income and do so through legal and illegal means. For example, a recent story in the *New York Times* began as follows: "In the wealthy, northern suburbs of [Athens, Greece], where summer temperatures often hit the high 90s, just 324 residents checked the box on their tax returns admitting that they owned pools. So tax investigators studied satellite photos of the area—a sprawling collection of expensive villas tucked behind tall gates—and came back with a decidedly different number: 16,974 pools."See Suzanne Daley, "Greek Wealth Is Everywhere but Tax Forms," *New York Times*, May 1, 2010, accessed January 30, 2011, http://www.nytimes.com/2010/05/02/world/europe/02evasion.html?hp.

The Incentive to Work

As the real return to working increases, households will generally work more. Labor supply is upward sloping: increases in the real wage lead to more people participating in the labor market and individuals' choosing to work more hours. Households care about the real wage *after taxes*—that is, they decide how much to work based on the wage they receive after paying tax. Everything else being the same, an increase in the tax rate on labor income reduces the real wage received by households, and they will work less in response.

Contrast high-ability and low-ability workers. High-ability workers are more productive. From society's point of view, it is better for them to work more. But if tax rates are higher for higher-income people, then these people will have an incentive to work less,

so total output for the economy will be lower. This lost output is the efficiency loss from the progressive tax system.

The Incentive to Train

Redistribution can also affect the incentive to study and acquire additional skills. Once again, we use our chocolate bar example. We still have two types of individuals: high ability and low ability. Which type you are when you are born is completely beyond your control; it is just a matter of luck. But the actions you take, given your ability, are something you control.

Suppose that high-ability people can only produce 100 chocolate bars if they first go through some training. Further, assume that this training is not fun: everything else being the same, people would prefer not to spend time training. Instead, they would prefer to use their leisure time in other ways. Under the social contract, the efficient way to organize society would be for high-ability people to incur the cost of training to produce more output.

If the tax-and-transfer system completely equalizes incomes, however, high-ability people will not think it worthwhile to train. This highlights a problem with the Marxian view of "from each according to his ability, to each according to his needs." The incentives needed to induce people to produce according to their ability may be inconsistent with allocating goods according to need.

Assuming that a little inequality is better than a lot of lost chocolate bars, the social contract needs to be amended to create an incentive for high-ability people to train. The solution is to give them some extra chocolate bars as an inducement to train and thus produce more for society. The result is inequality in consumption.

The Leaky Bucket

The incentive problems that we have discussed so far result in an **equity-efficiency trade-off**. Arthur Okun, a famous economist in the 1960s, proposed a very useful thought experiment for thinking about such trade-offs. He imagined that redistribution from the rich to the poor is like carrying a bucket of water from one person to another. Unfortunately, the bucket leaks. So the process of transferring water from one person to another also means that there is less total water available.

At one extreme, if the bucket does not leak, then there is no trade-off. You can redistribute water evenly in society without any loss in efficiency. At the other extreme, all the water gets lost in the transfer. The only way to achieve equality in this society is simply by destroying the wealth of the rich. Okun invited his readers to contemplate how much leakage they would be willing to tolerate to make society more equal. If you are in favor of a more equal society, then you too should think about the extent to which you think it is worth sacrificing some of our output to share the rest out more equally.

The Inheritance Tax

At the beginning of this chapter, we listed the wealthiest people in the United States in 2006 and 2010. Do you think that 50 years from now, the families of these people will appear on the *Forbes* list of the wealthiest people in the United States in 2060? The answer to this question partially depends on the choices of these wealthy people: how much of an estate will they decide to leave to their families? It also depends on how much of the estate the government will tax.

When we talked earlier about the dynamics of inequality, we noted that there were links across generations of a family. Some of those links come directly from expenditures on children. Everything else being the same, richer families have more income to spend on their children's education, and thus their children are likely to be more productive. The transfer of wealth is a second link that leads income (earned on financial investments) to be higher for children of wealthier families.

According to the current tax code in the United States, the tax rate applied to an estate appears to be progressive, with higher tax rates levied on larger estates. But there is an exclusion of $5 million, and only estates above this level are taxed at a 35 percent tax rate. So if you were left an estate valued at $6 million, you would pay a tax of $350,000 (= 0.35 × [$6,000,000 − $5,000,000]). Not surprisingly, the inheritance tax is hotly debated. Opponents of the tax argue that individuals ought to have the right to spend their lifetime income on whatever they want, including their children. Proponents of the tax see it as a way to increase mobility within the wealth distribution and argue that it promotes equality of opportunity.

Transfers

The government redistributes across households using taxes and transfers. This redistribution is reflected in the difference between the Gini coefficient for market income and postinsurance income in Table 13.1.3 "Household Income by Quintile". Transfers arise through unemployment insurance payments to unemployed workers, government-financed health care to the poor and the elderly, and other government schemes. Chapter 16 returns to the topic of government transfers associated with health care.

Transfers, like taxes, can affect incentives. Suppose the government makes transfers of $100 to everyone in the economy with income less than or equal to $1,000. Think about an individual who works 40 hours at a wage of $25 per hour to earn a weekly income of $1,000. What are the gains to working 41 hours? If the individual works an hour more, then her income (before taxes and transfers) will increase by $25 to $1,025. But by working an extra hour, she no longer qualifies for the transfer of $100. So she would lose $100 in transfers: the extra hour's work would reduce her income by $75.

Not all transfers are public; some are private. Many of the wealthiest people in the world are also some of the most generous in terms of setting up private foundations. For example, the Bill and Melinda Gates Foundation (www.gatesfoundation.org/Pages/home.aspx) was created in 2000 "to help reduce inequities in the United States and around the world." The reported value of the trust endowment is $34.6 billion, which includes $1.6 billion from Warren Buffett, the number two person on the 2006 and 2010 *Forbes* lists. Another common form of private transfers comes from tuition reductions from private universities. As a leading example, Princeton University replaced student loans, which had to be repaid, with outright grants to qualified students. Other universities provide both grants and subsidized loans.

Key Takeaways

- Governments use a variety of tools, such as income taxes and inheritance taxes, to influence the distributions of income and wealth.
- Governments are motivated by the view that market outcomes are not equitable enough.
- Actions to redistribute are limited by the adverse incentives created by taxes and transfers.

Exercises

1. Does a progressive income tax lead to a more or less equitable distribution of disposable income?
2. Will an inheritance tax create an incentive for people to work more or less?

This page titled 13.5: Government Policy is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

13.6: End-of-Chapter Material 13.6: End-of-Chapter MaterialEdit Update titlecancel

In Conclusion

Most of the time in the study of economics, we focus on efficiency. We ask if there are better or worse ways for society to organize its production of goods and services, and we ask if society has institutions in place that allow people to obtain all the available gains from trade. Although these can be complex questions, there is broad agreement among most people that efficiency is a desirable goal.

In this chapter, we tackled a rather different and more contentious set of issues: what is fair and just? Economists can (relatively) easily explain how society ends up distributing its resources, but the tools of economics do not allow us to say whether a given distribution is fair or not. People certainly seem to care about fairness and hold strong opinions about how society should share out its resources. Unfortunately, different people have very different ideas about what is fair.

The questions we address here go beyond economics; they vex philosophers and political scientists as well. They go to the heart of what we think of as right and good. They also force us to think about the appropriate role of the state and how that matters for the distribution of resources. Is the role of the state simply to provide an environment where people are free to pursue their own self-interest and to keep what they earn? Or does the existence of the state mean that we are all in a social contract, so all have some rights to the output of society as a whole?

As we have said previously, we cannot and do not want to answer these questions for you. Indeed we, as authors of this book, do not even agree among ourselves on the answers. Instead, we have given you some tools so you can think about these questions—which are some of the most important you will ever confront—yourself.

Key Links

- US Census Bureau: http://www.census.gov
- *Forbes* lists: http://www.forbes.com/lists
- The World Bank on poverty: web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTPOVERTY/0,,menuPK:336998~pagePK:149018~piPK:149093~theSitePK:336992,00.html
- Gates Foundation: www.gatesfoundation.org/Pages/home.aspx

exercises

1. Draw a Lorenz curve for the data given in Table 13.1.2 "Example of Income Distribution".
2. Often income data are reported by household. How does the US Census Bureau define a household? Is this the same as a family?
3. Draw the Lorenz curve for the wealth of the top 10 people in the United States for 2006 and 2010 using the data in Table 13.1.1 "Wealthiest Individuals in the United States".
4. Can you think of two other markets with significant winner-takes-all elements?
5. During the past 100 years, there has been tremendous technical progress in creating machines to run in the household, such as dishwashers, washing machines, clothes dryers, and so on. How do you think these inventions have affected the labor participation decisions of women and the wages they are paid?
6. Suppose that the cost of training is 20 chocolate bars. Assume high-ability people produce 100 bars if they get training and 50 bars if they don’t. Low-ability people produce 50 bars regardless of training. If under the social contract you decide to provide an incentive for high-ability people to train, what is the distribution of consumption in the economy? Is society better off with inequality in consumption or is it better to have equal consumption and no training by high-ability people?
7. Start with the example of the social contract given in Question 6 but suppose that 75 percent of the people are high ability and 25 percent are low ability. What does the social contract look like for this economy? How much is produced by high- and low-ability people? What is the total amount of output per capita? What is the consumption per capita?
8. Identify two institutions that provide equality of opportunity but not outcome. Identify two institutions that favor equality of outcome over equality of opportunity.
9. Use college admissions to illustrate the difference between “equality of opportunity” and “equality of outcome.”

10. Suppose a household holds a share of stock in a particular company and receives a dividend from that share. Which of these is a stock, and which is a flow? Which is part of income, and which is part of wealth?
11. (Advanced) Consider two countries: one has a higher Gini coefficient and the other has less mobility across income groups over time. Which country has greater equality?
12. If the return to education depends on innate ability, then what is the point of going to college?
13. Do you think that trading in the stock market exhibits equality of opportunity? Why or why not?
14. Can you come up with your own example of a trade-off between equity and efficiency?

Economics Detective

1. Pick one of the wealthiest people in the United States. How did this person get his or her wealth? How much do you think this person earns each year from his or her assets?
2. Find a list of the world's wealthiest people. What countries are these people from? Pick one person and see how the person got his or her wealth. Are the wealthiest people in the world distributed across lots of countries or isolated in a just a few?
3. The Rockefeller family was one of the wealthiest in the United States around 1900. How did the family accumulate its wealth? Where did the wealth go??
4. Try to find data on the share of income of the bottom 20 percent of the income distribution in two different countries. Also try to find the Gini coefficients for the two countries. How might you explain the differences in income distribution between the two countries you chose?
5. Go to the website of the Internal Revenue Service. Find the tax rates currently in effect for different income levels in the United States. Are these progressive?

Spreadsheet Exercise

1. Create (or find on the Internet) data on income. Input the data into a spreadsheet and plot the Lorenz curve.
2. Create a spreadsheet to follow the income and wealth of two households. Suppose the first household earns 50 chocolate bars each year, and the second household earns 100 chocolate bars each year. Suppose that each household saves a fixed fraction of its income (you can vary this in the spreadsheet). Follow these households for 50 years. Calculate familial wealth year by year using the equation at the beginning of Chapter 13, 13.2 Section "The Sources of Inequality".6. To do this, you will have to specify the interest rate (which you can also vary). In what sense is the distribution of wealth more unequal than the distribution of income? What if the high-income households also had a higher return on saving? What if households sometimes produced 50 chocolate bars and other times produced 100 bars? As a very advanced topic, can you build this uncertainty into your spreadsheet program? What happens to wealth?

This page titled 13.6: End-of-Chapter Material is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

14: Cleaning Up the Air and Using Up the Oil

14.1: Dirty Travels
14.2: The Economics of Clean Air
14.3: Externalities
14.4: Renewable, Nonrenewable, and Accumulable Resources
14.5: End-of-Chapter Material
14.6: Appendix- An Example of the Hotelling Rule in Operation

---

This page titled 14: Cleaning Up the Air and Using Up the Oil is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

14.1: Dirty Travels 14.1: Dirty TravelsEdit Update titlecancel

Here are some places you probably would have difficulty finding on a map:

- Sumgayit, Azerbaijan
- Linfen, China
- Tianying, China
- Sukinda, India
- Vapi, India
- La Oroya, Peru
- Dzerzhinsk, Russia
- Norilsk, Russia
- Chernobyl, Ukraine
- Kabwe, Zambia

These 10 places have the dubious distinction of being the world's most polluted cities, according to a nongovernmental organization called the Blacksmith Institute. See http://www.worstpolluted.org. As a child, one of the authors of this book lived for a year in one of these towns. Figure 14.1.1 shows these cities plus 20 more, giving us the 30 cities in the world with the worst pollution. In some of these places, mining and smelting industries have contaminated the air or the groundwater. In some, dangerous chemicals have been improperly disposed of—often illegally. In some, there is radioactive contamination. In some, garbage and sewage pollute the groundwater, or automobile emissions pollute the air. In sum, you would not want to live in any of them.



Figure 14.1.1: The 30 Most Polluted Cities in the World. This map shows the locations of the 30 most polluted cities in the world, according to the Blacksmith Institute.

The consequences of such pollution are tragic. Pollution of this magnitude causes severe health problems, birth defects, and high mortality rates. For example, according to the Blacksmith Institute, life expectancy in Dzerzhinsk, Russia, is 42 for men and 47 for women. Lead pollution is directly linked to a reduction in children's intelligence and has also been linked to increased violence.

The map reveals that many of these highly polluted cities are found in India, China, and the countries of the former Soviet Union. The richer countries of the world, such as the United States, Canada, Japan, Australia, New Zealand, and all of Western Europe, are not featured. Neither, for the most part, are the very poorest countries of the world, such as much of Africa. Severe industrial pollution seems to be at its worst in middle-income, developing countries. This does not mean that rich countries did not recently have—or do not still have—pollution problems of their own. Only a few decades ago, the Cuyahoga River in Ohio was so polluted that it caught fire; indeed river fires were once relatively commonplace in the United States. The US Environmental Protection Agency lists well over 1,000 sites as eligible for Superfund cleanup (http://www.epa.gov/superfund) because environmental contamination is judged hazardous to health. Little more than 50 years ago, air pollution killed an estimated 4,000 people in London, England, during the so-called Great Smog.

We emphasize in many places in this book that market transactions generate value in an economy. Firms produce things that people want to buy, so both firms and consumers benefit. People voluntarily work for companies, earning money they can then use to purchase goods and services while simultaneously allowing firms to produce the products that people want. These claims are correct, yet the citizens in Linfen, China, or La Oraya, Peru, could be forgiven for thinking that this is a very rosy view of how economies function in practice. Those who live in these communities around the world obviously do not like living in such polluted environments. So what is going wrong? How is it that voluntary trades made by individuals and firms can sometimes lead to such unpleasant and dangerous outcomes?

To begin our answer to this question, let us take a particular example: Mexico City, a city that also makes the list of the worst 30. The air in Mexico City contains particulate matter (think of this as soot and smog) that can cause lung disease and other bronchial problems. This pollution largely comes from automobile emissions, which are a severe problem in part because so many cars in Mexico City are old. According to the Blacksmith Institute report,Blacksmith Institute, "The World's Worst Polluted Places," September 2007, accessed March 14, 2011, www.blacksmithinstitute.org/wwpp2007/finalReport2007.pdf. even a moderate reduction in the amount of particulate matter in the Mexico City air could save thousands of lives each year.

But if everyone dislikes the pollution, why is it being produced? After all, no one is forcing the residents of Mexico City to drive their cars. They could all decide to drive much less, and if they did so, the result would be a cleaner city. Indeed, not everyone is a polluter. Particularly in richer countries, more and more people are driving electric cars or hybrids, which use a combination of electricity and fossil fuels. Such cars emit less pollution from their tailpipes. According to one study, the main reason that people purchase these cars is because they "want an environmentally friendly car." This reason was cited by 66 percent of the respondents. The next most popular response was "I want to save money on gas," which was cited by 16 percent of the respondents. The survey also found that half of all hybrid owners also donated money to environmental causes.Jonathan Klein, "Why People Really Buy Hybrids," *Topline Strategy Group*, accessed January 31, 2011, www.toplinestrategy.com/green_form.htm. Of course, people can have multiple motivations for purchase. The same study concludes that only about 27 percent of hybrid users do not have a financial motivation for their purchase.

Some of these reasons are clearly self-motivated: when gasoline costs $4 or more a gallon, fuel-efficient vehicles look very attractive. However, the desire to behave in an environmentally conscious way is rather different. People like to feel that they are behaving responsibly, even if they understand that their impact on total pollution is negligible. But there is another aspect of this desire to be green that is even more intriguing. If you interview one of these individuals, you will typically learn that he sees two ways in which hybrids are a good choice for the environment: (1) they generate less pollution, and (2) they consume less oil.

Both are environmental concerns. Both address how we use up different natural resources: clean air and fossil fuels. Yet they are very different problems. In this chapter, we consider both pollution and our consumption of natural resources—including but not limited to oil—and ask,

Can we rely on markets to deal with pollution and natural resources?

Road Map

We start our analysis with a familiar idea: the gains from trade. Figure 14.1.2, which also appears in other chapters in this book, illustrates one of the biggest insights of economics: voluntary transactions create value. In every voluntary transaction, both buyer and seller obtain surplus from trading. Even more striking, if these transactions take place in a competitive market, then buyers and sellers reap *all* the gains from trade.

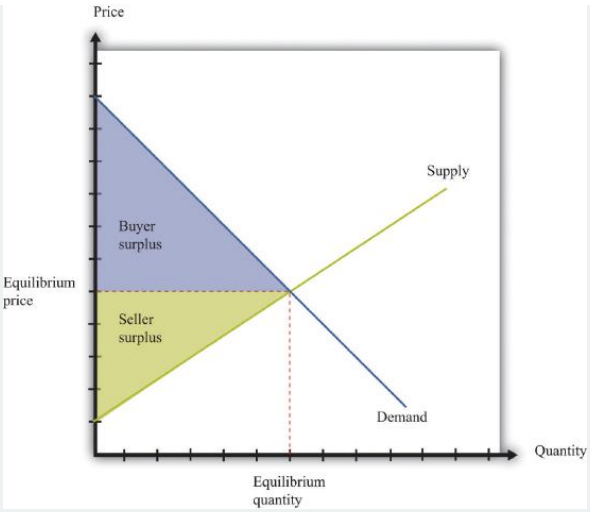


Figure 14.1.2: The Gains from Trade. In a competitive market, total surplus (the sum of the buyer surplus and the seller surplus) is maximized.

The first section of the chapter looks at the use of clean air. To start off, we tackle this in a small-scale situation: we consider what happens if a smoker and a nonsmoker share an office. We ask under what circumstances they might be able to resolve their disagreement without outside assistance. We then explain that air pollution in Mexico City is really the same problem, albeit much larger. We show that the problem of pollution has two related aspects: (1) we cannot easily force polluters to pay for their "use" of clean air, and (2) as a result, there is a divergence between the cost of an action to an individual polluter and the cost to society as a whole.

We discuss different kinds of policies that are used to address these problems. Then we turn to our use of natural resources such as oil. We consider various kinds of resources and consider what economic theory can teach us about how these resources are likely to be used. Finally, we consider the implications for economic policy.

This page titled 14.1: Dirty Travels is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

14.2: The Economics of Clean Air 14.2: The Economics of Clean AirEdit Update titlecancel

Learning Objectives

1. What is the Coase theorem?
2. Why is the Coase theorem important?
3. What is a social dilemma?

If we lived in a world where all economic transactions took place in competitive markets and in which there were "enough" markets, then we would obtain all the possible gains from trade. This logic falls down in reality because markets sometimes fail, for various reasons.

- Something prevents the economy from reaching the competitive outcome.
- Something prevents trade altogether, so the market is missing.
- Some people other than the buyer and seller are affected by the transaction.

We will get to Mexico City shortly. We begin, however, by thinking about a more isolated case of air pollution: cigarette smoke in an office.

Smokers, Nonsmokers, and the Coase Theorem

Cigarettes are sold and smoked almost everywhere. Yet in most countries around the world, you are not able to smoke when and where you please. Governments around the world place limitations on who can buy cigarettes, where they can be bought, and where they can be consumed. From an economic point of view, governments are deliberately restricting the ability of individuals to engage in voluntary transactions.There are many different ways in which governments intervene in market transactions. Chapter 12 contains more discussion. Why do governments restrict an individual's ability to smoke where and when that person wants?



Figure 14.2.1. Mixed messages.

Our answer to this question begins by imagining two people who must share an office. One is a nonsmoker who dislikes the smell of cigarette smoke, while the other likes to smoke while working. On the way to work one day, the smoker purchases a pack of cigarettes that she plans to smoke at work. We can reasonably deduce from this that her valuation of these cigarettes is greater than the price she has to pay. She gets **buyer surplus** from the purchase. To be concrete, suppose a pack of 20 cigarettes costs $4 and her valuation of a pack of cigarettes is $10. Her surplus is then $6.

We can also reasonably assume that the seller's cost is less than the price, otherwise he would not choose to make the sale. He gets **seller surplus**. For example, if his wholesale price for a pack of cigarettes is $2, then he earns $2 surplus (= $4 − $2) on every pack that he sells. The total surplus from the sale is the buyer surplus plus the seller surplus—that is, $8. So far so good.

The problems begin when the worker smokes her pack of cigarettes in the office. She obtains her $6 worth of enjoyment. However, a third party has now been affected by her decision to consume cigarettes: her office mate. The office mate dislikes the smell of smoke and may even face health risks from second-hand smoke. Thus even though the smoker and the store that sold the cigarettes are both better off, the office mate has been made worse off.

We should not automatically assume that the best thing is to ban smoking in the office just because the office mate is adversely affected. We need to know *how much* the nonsmoker is inconvenienced. Suppose the most the nonsmoker would be willing to pay for a smoke-free office is $2 per day. In this case, the $6 gain to the smoker exceeds the $2 loss to the nonsmoker. It seems like it should be easy enough for the two individuals to find a way for everyone to be happy. For example, imagine that the smoker agreed to pay the nonsmoker $4 a day for the right to smoke in the office. Both would then get $2 surplus per day.

On the other hand, suppose a smoke-free office is worth $10 per day to the nonsmoker. In this case, his valuation of clean air ($10) exceeds the smoker's gain ($6). The smoker would be unwilling to pay the nonsmoker enough to compensate for dirtying the air in the office. It would be better not to allow smoking in the office.

We have assumed here that the default situation is that the office should be smoke-free. In the language of economics, the nonsmoker owns the **property rights** to the clean air in the office. Property rights over a resource mean that, by law, the owner can make all decisions regarding the use of the resource. Because of this, the smoker must pay the nonsmoker compensation if she wishes to be allowed to smoke in the office.

We could imagine the opposite situation, where the smoker starts off with the right to smoke in the office. Would we expect a different result from their negotiations? If the smoke-free office was worth only $2 to the nonsmoker, then he would not be willing to pay enough to persuade his office mate not to smoke: the most he would pay is $2, which is less than the smoker's surplus. If, on the other hand, the nonsmoker valued the smoke-free office at $10, then he values a smoke-free office more than the smoker values smoking in the office. The nonsmoker could pay the smoker not to smoke. For example, imagine he pays her $8 per day to not smoke. He pays $8 for the clean air, which is worth $10 to him, so he gets $2 of surplus. The smoker receives $8, which exceeds the surplus she would get from smoking in the office. Again, they would both be happy.

Thus if the smoker "owns" the clean air, the nonsmoker must pay the smoker if he wants a smoke-free office. If the nonsmoker has the property rights, it is the smoker who must pay. In either case, the basic outcome will be the same: there will be smoking in the office if the smoker's valuation exceeds the nonsmoker's valuation; there will be no smoking if the nonsmoker's valuation exceeds the smoker's valuation. But the property rights are valuable. It is the owner of the property rights—whoever that may be—who gets compensation from the other.

As long as they know who has the property rights, it seems likely that the two individuals will be able to come to an agreement that benefits them both. You might imagine, though, that they would find it far harder to come to an agreement if it was ambiguous who had the property rights in the first place. The smoker would likely claim the right to smoke in the office, while the nonsmoker would assert his right to clean air. If they could not settle this basic question, it is unlikely that they would be able to reach a more complicated agreement involving compensation payments.

Let us imagine a further twist. Suppose the nonsmoker has property rights but values clean air at only $7 per day. This is still greater than the smoker's surplus, so, as before, we expect that they would agree to a smoke-free office. But the nonsmoker's valuation for clean air is *less* than the total surplus of the smoker and the store that sold her the cigarettes (recall that the smoker gets $6 surplus and the store gets $2 surplus). If the storekeeper, the smoker, and the nonsmoker all got together, they should again be able to find an arrangement that benefits everyone. For example, the storekeeper and the smoker could jointly give the nonsmoker $7 and still have $1 of surplus to bargain over.

It seems perfectly reasonable to imagine that two people who share an office could come to a mutually beneficial agreement about smoking in the office. It seems much more far-fetched, though, to imagine that they would come to an agreement together with the storekeeper who sold the cigarettes. Economists say that the difference between the two cases is due to **transaction costs**—the costs of making and enforcing agreements.

We began this section by observing that a transaction may affect not only the buyer and the seller but also third parties. When this is the case, we cannot be sure that trade benefits everyone. Even if the buyer and the seller are both made better off, third parties may be made worse off. However, if property rights are clearly established and transaction costs are low, then we can expect that

private negotiations could solve these problems. This idea was first articulated by the Nobel prize–winning economist Ronald Coase (http://www.coase.org) and is called the **Coase theorem**: if property rights are clearly established and transaction costs are low, private bargaining will lead to efficient outcomes.

It is notable that in reality, we do not see office workers buying and selling the right to smoke in an office. Instead, blanket bans on smoking have been enacted throughout the United States and in many other countries throughout the world. Why do we get this government response? One argument for these bans is that smoking poses health risks. In this case, antismoking campaigns are based on an idea that individuals are not always capable of making good choices for themselves. Chapter 5 has more to say about whether individuals are good judges of their own actions, particularly when making decisions with long-term consequences. But another reason is a recognition of the transaction costs involved in these private negotiations. Even if you think two coworkers could reach an agreement, imagine an office with 10 people—perhaps there are 3 smokers and 7 nonsmokers—all of whom place a different valuation on clean air in the office. If we knew everybody's true valuations, then in theory it would be possible to create a system of payments that made everybody better off. In practice, however, people might lie about their valuations. Finding the right system of payments would be very hard indeed and would take a lot of time and effort.

Perhaps you can now see the parallel with Mexico City. The major pollution problem is the emissions from the cars that people drive. Individual residents of Mexico City make decisions to buy gasoline and drive. These transactions create value for the drivers and the sellers of gasoline—but third parties are adversely affected. The Coase theorem can work when the parties involved are easily identifiable and small in number. In contrast, it is impossible to imagine the 20 million residents of Mexico City all meeting and coming to some kind of private agreement to limit their collective driving behavior.

Social Dilemma

Air pollution and second-hand smoke have a common structure, which we explain in this section. Once you understand these common elements, you will probably be able to think of many other examples. We show the interactions between an individual and the rest of society in Table 14.2.1 "The Payoffs in a Social Dilemma Game". This is called a **social dilemma** game.

Table 14.2.1: The Payoffs in a Social Dilemma Game

	Everyone Else Drives (Air Is Polluted)	Everyone Else Takes Public Transportation (Air Is Clean)
$0; $1 > −$1). Regardless of the actions of others, your payoffs are always higher if you drive ($0 > −$1; $2 > $0)." class="lt-socialsci-45271">You Drive	$0; $1 > −$1). Regardless of the actions of others, your payoffs are always higher if you drive ($0 > −$1; $2 > $0)." class="lt-socialsci-45271">$0	$0; $1 > −$1). Regardless of the actions of others, your payoffs are always higher if you drive ($0 > −$1; $2 > $0)." class="lt-socialsci-45271">$2
$0; $1 > −$1). Regardless of the actions of others, your payoffs are always higher if you drive ($0 > −$1; $2 > $0)." class="lt-socialsci-45271">You Take Public Transportation	$0; $1 > −$1). Regardless of the actions of others, your payoffs are always higher if you drive ($0 > −$1; $2 > $0)." class="lt-socialsci-45271">−$1	$0; $1 > −$1). Regardless of the actions of others, your payoffs are always higher if you drive ($0 > −$1; $2 > $0)." class="lt-socialsci-45271">$1
Regardless of which action you choose, your payoffs are higher if everybody else takes public transportation ($2 > $0; $1 > −$1). Regardless of the actions of others, your payoffs are always higher if you drive ($0 > −$1; $2 > $0).		

In Mexico City, it is very possible that people would agree that they would prefer a situation where everybody drives less. Yet this is not what happens. To see why, look at Table 14.2.1 "The Payoffs in a Social Dilemma Game" and imagine you are a car owner in Mexico City. You have to decide how to get to work—by driving or taking public transportation. There are two rows in the table. One is labeled "You Drive," and the other is labeled "You Take Public Transportation." The rows represent your possible choices. The columns refer to everybody else's choices. Everyone else who owns a car similarly chooses between driving and taking public transportation. To keep things simple, we suppose that everyone else makes the same choice. If everybody chooses to drive, then the air is polluted. If everybody chooses to take public transportation, then the air is clean. The current situation in Mexico City is that you and all other car owners are driving to work. You enjoy the convenience of driving rather than taking public transportation, but you suffer from the polluted air.

The numbers in the table refer to your payoffs from the different possible combinations. As in our smoking example, we can think of these as the valuations per day that you place on different outcomes. What matters is how these different possibilities compare

with the status quo, where you and everyone else drive. We therefore begin by setting your payoff at the status quo (the number in the top left cell) at $0. Suppose also that the following are true.

- You value clean air rather than dirty air at $2 per day—that is, you would be willing to pay $2 per day to have clean air rather than dirty air.
- After taking into account the relative costs and inconveniences of driving versus public transportation, you value driving compared to public transportation at $1 per day—that is, you would need to be compensated $1 per day to make you just as happy to take public transportation rather than drive.

Based on these conditions, we can calculate the payoffs in the other three cells of the table:

- The top right cell is the case where you drive and everyone else takes public transportation. Compared to the status quo, you are better off because you get to enjoy clean air. This is worth $2 to you, so your payoff is $2.
- The bottom left cell is the case where you take public transportation and everyone else drives. You still have to breathe polluted air, and you no longer get the convenience of driving (which is worth $1). Compared to the status quo, you are worse off. Your payoff is −$1.
- The bottom right cell combines the two previous cases. In this case, you and everyone else take public transportation. You get the benefit of clean air (worth $2), but you lose the benefit of driving (worth $1). Your payoff is $\$2 - \$1 = \$1$.

What would you do in this situation? Suppose you think that everyone else is going to drive. You are better off if you drive (payoff is $0) rather than take public transportation (payoff is −$1). What if everyone else takes public transportation? Then you still prefer to drive (payoff is $2) rather than take public transportation (payoff is $1). We conclude you will drive regardless of what others in society choose to do.

Here is the crux of the problem: *the situation looks the same to everybody else as it does to you*. As you evaluate these relative payoffs and choose to drive, so too does everyone else. We therefore expect that everyone will follow their individual incentives and choose to drive. We end up in the top left cell, where your payoff is $0.

What is striking, though, is that you would prefer the outcome where everybody—including you—uses public transportation. Your payoff in the bottom right cell is $1, which is better than the current situation. Everyone else would prefer this outcome as well. Society ends up in a bad situation, with everybody driving, even though everyone agrees that there is a better option out there. This is the essence of the social dilemma.

You are one of many people. Although you may value clean air, you are powerless as an individual to keep it clean. And because you are only one person, your decision has a tiny effect on the overall quality of the air. Thus you choose to drive, without paying attention to your effect on the environment. But because everybody makes the same decision, the cumulative effect is that there is a lot of air pollution.

The social dilemma is also sometimes known as the “tragedy of the commons.” This refers to the time when cattle farmers had access to common grazing land. Because every farmer had the right to graze his cattle on this land, no one was in a position to ensure that the land was well managed. Every farmer paid attention only to the health of his own cattle and did not worry about the effect of his cattle on the overall quality of the grazing land. Because every farmer made the same decision, the result was overgrazing, which destroyed the land for everyone.

How do we solve problems such as this? To avoid the bad outcome of the social dilemma, we must find some way of changing the payoffs of the game. We have already seen that, if transaction costs are low, people may be able to negotiate privately. In the case of Mexico City smog, however, such negotiation is impractical. In this case, one possible solution is for the government to alter the payoffs. For example, in the run-up to the 2008 Olympics in Beijing, the Chinese government wanted to improve air quality in the city. It therefore allowed cars to drive in the city only every other day (whether a car was permitted on a given day depended on whether the last digit of the license number was odd or even). If you tried to drive on the wrong day, you might have to pay a large fine, so the payoff to driving changed. We have more to say about government policy later in the chapter.

Toolkit: Section 31.18

You can review the social dilemma and other games in the toolkit.

Private Benefits and Social Costs

When people face a social dilemma, the actions that are the best for all individuals lead to an outcome that is bad for everyone. Much of the rest of the chapter addresses why this happens.

We begin by remembering our theory of how people make consumption choices. In general, people consume a good up to the point where their **marginal valuation** from the last unit of that good equals the price of that unit. This is a specific statement of a more general principle for decision making: “consume until marginal benefit equals **marginal cost**.”

In the social dilemma of Table 14.2.1 "The Payoffs in a Social Dilemma Game", you had two choices: driving or not driving. Let us now expand that to think about a situation where you are deciding how much to drive. Driving your own car brings private benefits (that is, benefits that are obtained only by you), such as comfort and convenience. Driving also brings private costs, such as the costs of gasoline and maintenance of your vehicle. If this were all that was going on, there would be no problem: you would drive up to the point where your marginal (private) benefit from driving was equal to your marginal (private) cost of driving.

When you drive, though, you also impose costs on other people. Your decision to drive one more mile has a **marginal social cost** as well as a marginal private cost. Marginal social cost is the cost to society of consuming or producing one more unit of a good or a service. By “society,” we simply mean “you and everybody else.”

When you choose to drive your car, you contribute to air pollution. This is a cost to the rest of society. However, you have no incentive to worry about this. You care only about the extra cost to you. The same is true when the smoker smokes one more cigarette in the office: She imposes a cost on her office mate, but—unless they make an agreement otherwise—she does not pay this cost. She takes into account the marginal private cost to her (that is, how much she must pay for one more cigarette), but she ignores the cost to other people.

Ignoring for a moment some tricky questions about how to measure the cost of your actions to society, we can set out a principle for how much you *should* consume if your concern were the overall well-being of society: “consume until marginal benefit equals marginal social cost.” The marginal social cost of your driving is the extra cost if you drive more. It is the cost both to you *and* the cost that you impose on others in society. There can be many components to this cost. One component is the marginal private cost to you. In addition, you pollute the air, contributing to public health problems. You add greenhouse gases to the atmosphere, contributing to the risks of climate change. You make the roads more congested, thus wasting the time of other drivers on the roads. You cause wear and tear on the roads, which will ultimately be paid for by taxes on all drivers.

You might be puzzled at this point. If you drive one extra mile, does that really have any appreciable cost on society? The extra emissions from your driving obviously have a tiny effect on pollution and greenhouse gases. The wear and tear you impose on the roads is minimal. How can these minuscule effects possibly matter? The first and more obvious answer to this question is that the quality of the atmosphere and the roads is affected by everybody’s decisions—not only yours. Hundreds of thousands of cars around the world are polluting the atmosphere and damaging the roads. The second, more subtle, answer is that though your individual influence on air and road quality is very small, you are affecting a very large number of people. If you drive an extra mile in Mexico City, you are affecting the air that is breathed by 20 million people. The marginal social cost of your driving includes the effect on every single one of these people. Likewise, one more mile of driving may add only a tiny amount to the greenhouse gases in the atmosphere, but that tiny effect must be added up over the entire population of the world.

As a result, the marginal social cost of your driving is greater than your marginal private cost. Figure 14.2.1 shows the implications of this. Think of consumption in this diagram as referring to the amount of driving you do. The marginal private cost of your driving is the cost of fuel, depreciation of your car, and so on. But the marginal social cost also includes the pollution of the air and the congestion of the roads. Because the marginal social cost is greater than the marginal (private) cost, you will drive too much, from the perspective of society as a whole. This is exactly what we saw in the social dilemma. People choose to drive and pollute the air, even though all members of society could be happier if everyone were to take public transportation and generate less pollution.

The gap between private costs and social costs means that too much driving is undertaken, from the perspective of society as a whole. The outcome is inefficient because people only have an incentive to take account of the private costs of their actions.

Toolkit: Section 31.11

You can review the concept of efficiency in the toolkit.

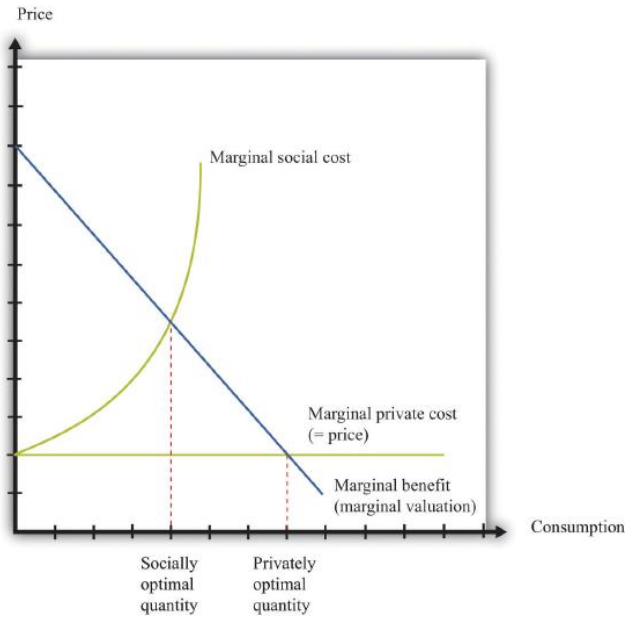


Figure 14.2.1: A Divergence between Marginal Private Cost and Marginal Social Cost. The gap between marginal private cost and marginal social cost is a measure of the impact that an individual has on the rest of society.

Figure 14.2.1 also illustrates another point. Economic analysis tells us that there is “too much” pollution, from a social point of view. This observation probably comes as no surprise. Economic analysis also makes it clear, though, that it is possible to have too little pollution as well as too much. There is an optimal amount of driving for each individual, to be found where marginal social cost and marginal benefit are equal. At this amount of driving, there will be some pollution: the optimal amount of pollution for society as a whole. If we were to ban driving altogether, we would have less pollution, but we would also lose all the benefits from driving.

Our discussion here has been about decisions of consumers, but firms also are sources of pollution. Firms use trucks and other vehicles that, like cars, impose costs on the rest of society. Some firms pollute the air or the water. Exactly the same principles still apply. Any individual firm has no incentive to take into account the costs that it imposes on the rest of society. As a result, firms pollute too much from a social point of view.

Key Takeaways

- The Coase theorem states that if property rights are well defined and transaction costs are low, then bargaining will lead to an efficient outcome.
- The Coase theorem provides the rationale for a market solution to pollution and other similar social problems.
- A social dilemma arises when there are many individuals each making choices that are in their self-interest but leading to an outcome that is bad for society.

Exercises

1. Why does the Coase theorem require that transaction costs be low?
2. Consider Table 14.2.1 "The Payoffs in a Social Dilemma Game". Suppose your payoff from taking public transportation when everyone else drives is −$2 rather than −$1. Would this change the basic message of this example?
3. Look at Figure 14.2.1. How would you modify this figure to make the difference between social and privately optimal quantities larger? Explain your reasoning.
4. Suppose you are in the library and there are two people making out between the shelves. Describe this situation in terms of social versus private costs. How would you use the Coase theorem to find an efficient allocation?

This page titled 14.2: The Economics of Clean Air is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

14.3: Externalities 14.3: ExternalitiesEdit Update titlecancel

Learning Objectives

1. What is an externality?
2. What are the ways in which problems caused by externalities can be solved?
3. What are some of the difficulties in designing policies to deal with externalities?

At the heart of the social dilemma is a divergence between private costs and social costs. Individuals and firms take into account only private costs when making decisions. But the social costs should matter as well. Thus actions that are individually optimal are damaging to society as a whole. Now that we have diagnosed the problem, how do we fix it? The economist's answer is that we must change people's incentives. A social dilemma arises when individual incentives are not well aligned with the interests of society as a whole. Economic policies focus on how to adjust those incentives so that there is a better match between individual and social aims.

Before discussing these policies in detail, we look again at the problem of the social dilemma, focusing now on the *actions* that people choose to take. In our Mexico City example, people decide whether or not to drive. If they choose to drive, this action affects the well-being of others. Economists say that there is an **externality** associated with the action of driving.

Toolkit: Section 31.19

An externality occurs when one person takes an action that directly affects another's welfare, but the effect does not operate through prices.

An externality must come from an *action*—something that somebody does. Good weather is not an example of an externality. Nor is an earthquake. The action could be taken by an individual (say, smoking a cigarette) or a firm (dumping toxic waste into a river). In most cases, the action is associated with production by a firm or consumption by a household.

In addition, the action must directly affect another individual's well-being or a firm's profits. It could be something that affects the health or happiness of an individual. It could be something that affects the profits of a firm. ("Directly" here means that the effect doesn't come about because of an induced change in behavior. Suppose, for example, that a firm offers you a job, but to get to that job you now must spend a longer time commuting. The extra commute is *not* an externality imposed on you by the firm.)

Finally, the effect must not operate through prices. Whenever we take part in market transactions, we have effects (usually tiny effects but effects nonetheless) on market prices. These changes in prices make others in the market better or worse off. But they are not externalities.

In our earlier example of driving, the marginal social cost was larger than the marginal private cost. The gap between the two is a measure of the size of the externality. Because the action of driving imposes a cost, we call this a **negative externalities**. Pollution is the classic example of a negative externality, but there are others. Congestion of public roads or public parks is another instance of a negative externality.

By contrast, there are also occasions when an action bestows an external benefit on third parties. We call this a **positive externalities**. For example, writers of open-source software create a social benefit that is in excess of the private benefit that they personally obtain. As another example, suppose that a firm engages in research and development and creates new knowledge. If others are also able to benefit from that knowledge without paying for it (for example, after the expiry of a patent), they are beneficiaries of a positive externality.

Toolkit: Section 31.19

- A positive externality arises when one person's or firm's action bestows benefits on others. When there is a positive externality, too little of the action is undertaken.
- A negative externality arises when one person's or firm's action imposes costs on others. When there is a negative externality, too much of the action is undertaken.

Although negative externalities sound bad and positive externalities sound good, positive externalities are also a source of inefficiency. The logic exactly parallels the case of negative externalities. Suppose a firm is deciding how much output to produce. To maximize its profits, it sets marginal cost equals to marginal private benefit (that is, marginal revenue). But if the firm's production generates a positive externality, the **marginal social benefit** exceeds its marginal private benefit. The firm produces insufficient output from a social point of view, as illustrated in Figure 14.3.1. The principle for socially efficient production is for the firm to produce up to the point where marginal cost equals marginal social benefit.

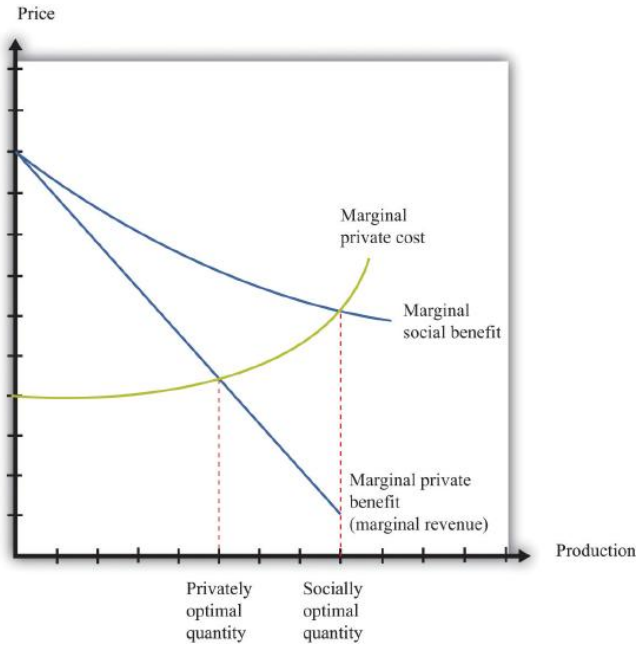


Figure 14.3.1: A Divergence between Marginal Private Benefit and Marginal Social Benefit. From a social point of view, however, the firm should produce up to the point where marginal social benefit equals marginal cost.

Solutions to Externality Problems

The definition of an externality makes it clear that the fundamental problem is one of behavior—actions by a firm or a household. The behavior reflects a difference between private costs or benefits and social costs or benefits. These observations also point us to a solution. We need to change incentives so as to align private costs or benefits and social costs or benefits. For example, if the private marginal cost of pollution to a firm were somehow equal to the social marginal cost, then a firm acting in its own self-interest would produce the socially optimal amount of pollution. The challenge for policymakers is to find a way to adjust the incentives so that the firm takes into account social marginal costs in addition to private marginal costs.

From this perspective, inefficiency arises because there are no market signals that force the polluter to take into account how its actions are affecting others. The goal of government policy in the presence of externalities is to provide incentives for firms and households to *internalize* their effects on others. These policies include direct restrictions on what people can do (for example, banning smoking in public buildings), taxes and subsidies that affect prices in an economy, and the introduction of markets that force polluters to pay for the right to pollute. Because externalities involve a divergence between private costs and social costs (or private benefits and social benefits), the goal in all cases is to adjust the incentives so that the actor internalizes the externality.

Creating Markets

We said that externalities are a source of inefficiency, but we should be more precise: externalities are a source of inefficiency *unless they are compensated for*. Think back to the smoker and nonsmoker who shared an office. The smoker's actions impose a negative externality on the nonsmoker. Without any compensating payments, we end up with an inefficient outcome. But when the smoker pays the nonsmoker for the right to use up the clean air, we end up with an efficient outcome. Negotiation between the smoker and the nonsmoker in effect creates a market for the clean air. Once this market is in place, the inefficiency disappears.

Building on this insight, governments can actively try to create markets to solve pollution and other externality problems. A good example of this is the 1990 Amendments to the Clean Air Act in the United States. Much of the air pollution in the United States is caused by utility companies (think of power stations), particularly those that generate electricity from coal. Such power stations

pump sulfur dioxide into the atmosphere, which causes acid rain and other environmental problems. The amendments to the Clean Air Act created **tradable emission permits** that were allocated to utility companies.

Such permits are licenses to emit a specified amount of pollution. A firm must own or purchase a permit if it wishes to emit pollutants into the atmosphere. These permits can be traded in a market. A firm that wishes to emit more pollution than allowed by its existing permits can purchase permits from others. A firm with more permits than it needs can sell them to other firms.

The first response of many people to a policy such as this is moral outrage. At first hearing, it may seem odd that the government is granting a license to pollute. In fact, this can be a very effective way to control pollution. To see how such a system works, suppose that there are two power stations.

1. GreenPower has installed pollution reduction measures and is emitting 100,000 tons of sulfur dioxide per year. Because it is already an environmentally friendly power station, it is very costly for it to reduce emissions further. To be concrete, suppose it will cost $20 per ton to reduce its emissions.
2. Atmosfear has no pollution control devices in place and is currently emitting 300,000 tons of sulfur dioxide per year. Because it has not yet installed any pollution reduction devices, it is able to reduce its emissions relatively cheaply, at a cost of $10 per ton.

Suppose the government decides that it wants to restrict the amount of pollution to a total of 200,000 tons per year, down from the current 400,000. It doesn't matter which power station emits the pollution; either way, the sulfur dioxide ends up in the atmosphere.

One approach is for the government to simply instruct each power station to cut its emissions by 50 percent. The trouble with this is that GreenPower already has an environmentally friendly system in place. It will cost $1 million ($= 50,000\ tons \times \$20\ per\ ton$) to reduce its emissions from 100,000 tons to 50,000 tons. Atmosfear has to reduce its emissions by 150,000 tons, which costs $1.5 million ($= 150,000\ tons \times \$10\ per\ ton$). The total cost of reducing emissions to meet the target is $2.5 million.

There is a better alternative. Suppose the government gives each power station a license to emit a certain quantity of pollution. For example, suppose it gives GreenPower a license to emit 50,000 tons and Atmosfear a license to emit 150,000 tons. So far, this is identical to the previous situation. Crucially, though, the government also allows the power stations to trade these licenses. If GreenPower can buy the right to emit sulfur dioxide for less than $20 per ton, it will prefer to do this rather than reduce its own emissions.

Because it costs Atmosfear only $10 per ton to reduce its emissions, the cheapest way to achieve a 200,000-ton reduction is for Atmosfear to carry out the entire reduction in emissions, down to 100,000 tons. If it does so, then it will have 50,000 unused permits that it can sell to GreenPower. The total cost of emissions reduction in this scenario is only $2 million. For example, suppose they agree on a price of $15 per permit. Then the cost to GreenPower is $\$50,000 \times \$15 = \$0.75\ million$ (instead of $1 million). The cost to Atmosfear is $1.25 million: $200,000 × $10 = $2 *minus* the $0.75 million it collects from GreenPower.

In some ways, this is like our smoking example. The power stations are able to trade, so they can both be better off. Emissions are reduced to the required level of 200,000 tons, and this reduction is achieved in the most cost-effective manner. Notice also that the government gets to decide on the total amount of acceptable pollution; the approach is consistent with very tight or very lax environmental standards. However, once the government has determined its desired level of pollution, the trading of permits allows the necessary reductions in pollution to be achieved at the lowest possible cost to society.

With tradable permits, pollution is controlled in the most efficient way, without regulators' needing detailed knowledge on different power stations. The trading system has some other advantages as well. Firms have an incentive to pollute less because they can sell excess permits for a profit if they do not need to use them. Environmental groups can even purchase emissions permits and take them out of circulation to reduce pollution below the level mandated by the government.

A more contentious question has to do with how the permits are allocated in the first place. One approach is to do what we did in our example: the government can simply allocate permits based on the existing level of emissions. A problem with this policy is that it effectively punishes firms that have already engaged in environmentally responsible changes. It also creates an incentive for lobbying by firms: because permits are valuable, firms will invest resources in trying to persuade policymakers to give the permits to them rather than to their competitors. Instead of giving away the permits, the government can instead auction them. This increases costs for the firms but has the advantage that it generates funds for the government. See Chapter 6 for a discussion on auction mechanisms.

Taxes and Subsidies

Remember that the problem with externalities is that private incentives do not reflect social costs and benefits. Another approach to fixing these incentives is through taxes and subsidies.

Consider first the case of a negative externality. The problem in this case is that the marginal social cost exceeds the marginal private cost. As we saw in Figure 14.3.1, this leads to overconsumption of a good. One way to fix the problem is to impose a tax that equals the difference between the marginal private cost and the marginal social cost. In effect, this converts the condition for private optimality (consume until marginal benefit equals marginal private cost) into the condition for social optimality (consume until marginal benefit equals marginal social cost).

The case of a positive externality is entirely analogous. The problem is that the marginal social benefit exceeds the marginal private benefit. As we saw in Figure 14.3.2, this leads to underproduction of a good. One way to fix the problem is to impose a subsidy that equals the exact difference between the marginal private benefit and the marginal social benefit. This converts the condition for private optimality (consume until marginal private benefit equals marginal private cost) into the condition for social optimality (consume until marginal social benefit equals marginal cost).

We see such policies in practice. Taxes on gasoline exist in part to compensate for externalities such as the pollution caused by automobiles and the wear and tear on roads. Subsidies to universities and think tanks exist in part to encourage the production of knowledge, which is a good with positive externalities.

Command and Control

There are also other kinds of environmental policies. The government can simply mandate that certain levels of pollution must not be exceeded. In some cases, such **command and control** regulation may be easier to implement and monitor.

Under most circumstances, economists favor either taxes or the creation of a permit market to command and control. The reason is that command and control is relatively inflexible and requires a lot of knowledge of how much pollution is generated by each individual firm. Taxes, subsidies, or permit markets are more flexible and do a better job of changing the incentives faced by polluting firms.

Encouraging Altruism

A rather different approach to externalities is to appeal to people's altruism. In economics, we typically assume that people consider only their own self-interest when making decisions. Moreover, we usually think of this self-interest in fairly narrow terms. This may seem a rather embittered view of human nature. People often do things that are directed toward others' happiness rather than their own—that is, people are sometimes altruistic rather than selfish. In the environmental context, people sometimes purchase products from environmentally responsible firms even if those products are more expensive. People sometimes purchase carbon offsets to compensate for the carbon generated by their driving or air travel. Companies often find it worthwhile to advertise the fact that they are environmentally responsible.

Governments also provide environmental information. The European Union has recently proposed measures requiring that new cars display their impact on global climate change, just like health warnings on cigarettes.See Ian Traynor and David Adam, "It's Lean and Mean, but Is It Green? EU Plans Clampdown on Car Ads," *Guardian*, June 5, 2008, accessed February 28, 2011, http://www.guardian.co.uk/environment/2008/jun/05/carbonemissions.carbonfootprints. Unlike health warnings on cigarettes, this labeling does not help people live healthier lives. And unlike efficiency ratings on appliances, this labeling does not help people make better decisions that will save them money. Climate change labels appeal purely to people who want to act in a way that will lessen their environmental impact. As we saw earlier in the chapter, purchases of hybrid cars seem to be primarily motivated by people's desire to be more environmentally responsible.

Psychologists and economists have studied why people sometimes behave in such ways. Perhaps it is because they are genuinely altruistic—they care about the well-being of others as well as themselves. Perhaps it is because of persuasion and peer pressure—if your friends behave in an environmentally conscious way, then there is pressure for you to do the same thing. Perhaps it is because of the feeling of satisfaction (sometimes called a "warm glow") associated with such behavior. For our purposes in this chapter, however, the exact reasons why people behave this way are less important than the behavior itself. There are several ways of thinking about altruistic behavior, but in the environmental context, altruism really amounts to people deciding to internalize some of the externalities that they impose. They incorporate some of the social cost into their own private costs. Governments, through advertising campaigns, can encourage such altruistic behavior.

Difficulties with Environmental Policy

Our discussion of externalities makes it seem that it is straightforward for policymakers to design effective environmental regulation. Government regulators simply need to calculate the difference between marginal private costs and marginal social costs and between marginal private benefits and marginal social benefits. Then they can put in place the appropriate taxes, subsidies, or both.

In practice, a major difficulty is knowing how to place values on externalities. Environmental policies to combat air or water pollution require the government to monitor the amount of emissions effectively and accurately. If emissions cannot be monitored, then tax or permit schemes are impossible to implement. Effective environmental policies also require the government to measure the damage incurred by the victims of the pollution.

As with many economic policies, questions also arise concerning the distribution of resources. In an environmental context, the key point of debate is often whether compensation should and will be paid. If a factory pollutes the river running through a town, imposing negative externalities on the town's residents, then is it enough to adjust the incentives so that there is the "right" amount of pollution—that is, so that the marginal benefit to the firm equals the marginal social cost? Or should the firm also be required to compensate the residents of the town? This is again a question of property rights, for we are really asking who has the initial right to the clean water in the river. It is also closely related to the question of whether pollution permits should be given away (implying that firms have the property rights) or sold at auction.

Whenever government steps in and enacts policies to affect behavior, it also must worry about whether there will be perverse responses to those incentives. As an example, the disposal of solid waste is a significant environmental problem in many countries. It is in part an economic problem: people do not usually pay directly for the removal of their garbage, so they do not have an incentive to recycle or avoid waste in other ways. A solution—favored by many economists and adopted by some municipalities—is to charge people a fee per bag for garbage removal. A study by Don Fullerton and Tom Kinneman revealed that this policy can bring its own problems.Don Fullerton and Thomas Kinneman, "Household Responses to Pricing Garbage by the Bag," *American Economic Review* 86, no. 4 (1996): 971–84. When such a scheme was introduced in Charlottesville, Virginia, citizens responded by putting much more garbage in each bag than they used to (Fullerton and Kinneman called this the "Charlottesville stomp"), so the reduction in the waste stream was lower than the authorities had anticipated. More seriously, some people also responded by dumping their trash illegally.

Valuing the Environment

The measurement of environmental harm is complicated. Suppose, for example, that pollution brings with it increased risk of disease or death. How do we place a cost on these risks? We can perhaps attempt to value health in terms of the costs of treatment and lost working hours, but it is much harder to place a value on the distress and suffering caused by ill health. Economists, lawyers, and others have even come up with varying ways to place a dollar value on human life, but—as you can surely imagine—these techniques are contentious. One reason is that all such approaches tend to rely, at least in part, on estimates of lost earnings. This means that the lives of skilled and high-paid individuals may end up being valued more than the lives of unskilled, lower-paid individuals.

A second set of issues has to do with how we value damage to the natural environment. Take, for example, the oil leak in the Gulf of Mexico in 2010, which caused substantial harm to birds, fish, and ocean ecosystems. Most of us are upset by the sight of seabirds with their feathers clogged with oil. But how should we assess the value of such damage? Similarly, what is it worth to ensure the survival of a particular endangered species?

Where possible, economists look to market prices to provide some indication of the value that society places on goods and services. In the case of environmental goods, though, we typically cannot look to markets. In such cases, we may need to use surveys and other methods for inferring household valuations. This is known as **contingent valuation**.

For several reasons, it is very difficult to carry out reliable contingent valuation surveys. In such a study, a household might be asked, "What would you be willing to pay to ensure that the whooping crane (for example) does not go extinct?" Such surveys often give implausible answers. For example, following an oil spill, people in Washington state and the province of British Columbia were supposedly willing to pay over $11,000 for each seabird that was saved—even though the seabird population would recover naturally in a decade or so.

The main problem, as you can perhaps guess, is that people do not face a real budget constraint when asked such questions, so they have no particular incentive to give a truthful answer. A related issue is that people are typically presented with an issue in

isolation. A household that claims to be willing to spend $50 to save one species from extinction might not be willing to spend $10,000 to save 200 species at a cost of $50 each. A third issue is that people may not have the information they need to make good decisions. We are not used to making purchases of environmental quality, so it is harder for us to give our valuation of a clean river than it is for us to give our valuation of, say, a bar of soap.

Yet, for all these objections, we do need some way of placing a value on environmental resources. The fact that we find it difficult to measure such things does not mean that they have no value. Contingent valuation studies are now very sophisticated, and researchers are very active in the search to make such studies better and more accurate.

Whose Welfare Should Be Included?

Another difficulty with environmental policymaking is that it is not always clear whose opinions should be taken into account when making environmental policy. For example, suppose there were a proposal to allow a major resort development at the Grand Canyon. Should that be the concern of residents of the area, the state of Arizona, or the entire United States? For that matter, should residents of other countries be entitled to a voice? After all, the Grand Canyon is one of the most spectacular sites in the world, visited by thousands of foreign tourists every year.

Perhaps you think it self-evident that foreigners should have no say in US environmental policy. Yet environmentalists in the United States and Europe have often voiced their opinion on the environmental consequences of policies in other countries, such as the construction of the Three Gorges Dam in China. Similarly, much of the world was outraged when, in 2001, the Taliban in Afghanistan destroyed two giant Buddhas that had been carved in the sixth century.

An even bigger problem has to do with the treatment of future generations. When you drive your car today, you are imposing costs not only on the living but also on those as yet unborn. The people most likely to be adversely affected by global climate change are not yet alive. How should we take into account their welfare and well-being? Scientists are largely in agreement that carbon emissions from the burning of fossil fuels will have an effect on global climate. Most environmental economists are convinced by this evidence, yet there continues to be disagreement among economists about the appropriate policy response. Most of that disagreement in the end comes down to different views about how to account for the welfare of future generations.

Uncertainty

One of the biggest difficulties with designing environmental policy is uncertainty. Global climate change is the clearest example. Although there is widespread agreement that we face some risk of climate change from carbon emissions, there is debate about the size of the effect. It is possible that we are facing only a small change in global climate, in which case it might not be worth spending a lot of resources now on reducing emissions. It is also possible that there could be catastrophic effects on global climate, in which case we should spend a lot of resources right now on reducing carbon emissions. And it is probable that we are facing something in between these extremes. Policymakers must try to get good estimates of how likely these scenarios are and then must decide how **risk-averse** they want to be when setting policy.

International Cooperation

Because environmental problems are often not confined to a single country, international agreements are sometimes needed for effective environmental policy. The *Montreal Protocol on Substances That Deplete the Ozone Layer*United Nations Environment Programme, “Section 1,” *Handbook for the Montreal Protocol on Substances that Deplete the Ozone Layer—7th Edition (2006)*, accessed March 14, 2011, ozone.unep.org/Publications/MP_Handbook/Section_1.1_The_Montreal_Protocol. is an example of successful international cooperation on environmental policy. In the 1970s, scientists recognized that certain chemicals known as chlorofluorocarbons (CFCs) were leading to a reduction in the atmospheric ozone layer. The ozone layer filters out dangerous radiation, so its destruction was linked to increases in skin cancer and other problems. The Montreal Protocol came into force in 1989 and has been signed by almost every country in the world. It mandated a gradual phaseout of CFCs, and current research shows that the atmospheric concentrations of CFCs have decreased as a result.

To date, however, there has been much less progress on the even bigger problem of global climate change. Although there has been much negotiation, there is still no international agreement on climate change that is comparable to the Montreal Protocol. In 1997, countries signed the Kyoto Protocol,United Nations Framework Convention on Climate Change, *Kyoto Protocol to the United Nations Framework Convention on Climate Change*, 1998, accessed March 14, 2011, http://unfccc.int/resource/docs/convkp/kpeng.pdf. which was the first major international agreement to address the accumulation of greenhouse gases in the atmosphere. Some countries—notably Canada, Australia, New Zealand, and most of Europe—committed

to specific targets, while others made more general commitments. However, the United States did not ratify the agreement. Meanwhile, several of the largest emitters of greenhouse gases, such as China, India, and Indonesia, did not have any specific commitment to greenhouse gas reductions. As a result, the impact of the protocol is greatly limited. One interesting feature of the Kyoto Protocol is that it included provisions for emissions trading.

In 1999, the United Nations held another summit in Copenhagen, in an attempt to make more progress on this topic. The outcome of the summit was the Copenhagen Accord:United Nations Framework Convention on Climate Change, "Copenhagen Accord, Draft decision -/CP.15," Copenhagen, December 7–18, 2009, accessed March 14, 2011, http://unfccc.int/resource/docs/2009/cop15/eng/l07.pdf. a declaration that climate change was a problem. However, no binding commitments on greenhouse gas emissions resulted. A follow-up meeting in Cancún, Mexico, in 2010, delivered the Cancún Agreement,United Nations Framework Convention on Climate Change, "Outcome of the work of the Ad Hoc Working Group on long-term Cooperative Action under the Convention, Draft decision -/CP.16," accessed March 14, 2011, http://unfccc.int/files/meetings/cop_16/application/pdf/cop16_lca.pdf. which—although still not a binding treaty—was judged by many observers to represent significant progress over the Copenhagen Accord. Still, despite many meetings and fine-sounding commitments, the world is a long way from having an agreement on greenhouse gas emissions to match the Montreal Protocol.

Other Kinds of Externalities

Our discussion here has been about externalities arising from either the production or the consumption of goods and services. However, the idea of externality is used much more broadly in economics because it is a very helpful way of diagnosing inefficiencies.

One example comes from the working of labor markets. In general, both firms and workers spend time and resources trying to make a good "match." Firms have human resource departments that spend resources advertising jobs, going through résumés, interviewing job applicants, and so on. Workers spend time preparing their résumés, interviewing at different firms, and so on. Both are willing to do this because a good match can be very beneficial: the firm gets a highly productive worker for which it is willing to pay a good wage. Chapter 9 discusses search and matching in the labor market in more detail.

Now, the more effort the worker and the firm expend, the more likely they are to come up with a good match. And though each benefits individually from this effort, some of the benefit also flows to the other side of the market. In other words, if a worker tries harder to make a good match, this bestows a positive externality on a firm, and if a worker tries harder to make a good match, this bestows a positive externality on the worker.

As another example, is it a good thing if a new fast-food restaurant opens in your town? To answer this, we can think about the externalities that arise because of the decision to enter. When a new restaurant opens, it is providing a product that is similar but not identical to some of the offerings already in the market. It thus steals business from existing restaurants and reduces their profits. We can think of the entry of a new restaurant as imposing a negative externality on existing restaurants. Conversely, the entry of a new restaurant bestows a positive externality on consumers because they benefit from the increased choice. Thus the entry of a new restaurant leads to both positive and negative externalities.

We said previously that if an action has positive externalities, then there will be too little of that action from a social point of view. If an action imposes negative externalities, there will be too much of that action. In this example, the action is the entry of a new firm or product. If the positive externalities are more important, then there is too little entry, from a social point of view. If the negative externalities are more important, then there is too much entry. In general, we cannot draw any conclusions about whether there are too many firms (or products) or too few.

Key Takeaways

- An externality arises when an action taken by one person directly affects another's welfare. These operate outside of markets.
- One solution to an externality problem is to create a market so that the effects of one person's actions on others will be reflected in the market price of taking that action. Another solution is to put in place taxes or subsidies so that private incentives are aligned with social goals.
- One challenge for policy design is that the valuation of environmental goods is difficult to measure. Moreover, external effects do not respect borders, so international agreements are often required.

Exercises

1. If someone living nearby you is playing loud music, is that a positive or negative externality?
2. In a system with tradable permits, how is the total quantity of permits determined?
3. Suppose you invent a new product. After the patent expires, others can freely copy your design. Is your act of invention an example of an externality?

This page titled 14.3: Externalities is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

14.4: Renewable, Nonrenewable, and Accumulable Resources 14.4: Renewable, Nonrenewable, and Accumulable ResourcesEdit Update titlecancel

Learning Objectives

1. What is the Hotelling rule for the use of resources?
2. What is a nonexcludable good?
3. What is an accumulable resource?

Some individuals take deliberate actions to limit or reduce their environmental impact. For some people, the choice of a greener lifestyle can affect almost all aspects of their consumption, including housing, transportation, and the food that they eat. There are, of course, many different ways in which our actions affect the environment. Our choices affect the amounts of pollutants that are emitted into the air and the water. They affect the amount of greenhouse gases that enter the atmosphere, which in turn has an impact on climate change. And they affect the rate at which we use up natural resources such as oil.

We explained that air pollution from automobiles is an example of an externality, which means that there is a divergence between marginal private costs and marginal social costs. Because of this externality, decisions to drive in a city can generate a social dilemma. It is not obvious, though, that consumption of gasoline generates an analogous problem. Some people argue that, if you are willing to pay for gasoline, you have the right to buy as much or as little as you please. If you want to drive a gas-guzzling Hummer or sport utility vehicle (SUV) rather than a fuel-efficient Prius, then—in their view—that is your right. Others make a different claim. They argue that because oil is a scarce resource, we ought to be conservative with its use. Driving a gas guzzler—in their view—wastes the earth's limited natural resources and thus is environmentally irresponsible. We now evaluate these two views.

Nonrenewable Resources and the Hotelling Rule

Oil is a **nonrenewable (exhaustible) resource**—that is, a resource that does not regenerate over time. Obviously, if we keep using up a nonrenewable resource, we will eventually run out of it, which is why it is called exhaustible. We most often hear this concern voiced about oil. For example, as oil prices increased to record levels in 2008, there was a great deal of discussion in the media about "peak oil"—the point at which world oil production is at its maximum.

Given that total oil resources are limited, we will eventually go beyond the peak oil point. Many commentators believe that we are at or close to this point already and have expressed concern about the implications of this for the world's economy. Consider, for example, the following quotation from a 2005 report:

World oil demand is expected to grow 50 percent by 2025. To meet that demand, ever-larger volumes of oil will have to be produced. Since oil production from individual reservoirs grows to a peak and then declines, new reservoirs must be continually discovered and brought into production to compensate for the depletion of older reservoirs. If large quantities of new oil are not discovered and brought into production somewhere in the world, then world oil production will no longer satisfy demand. That point is called the peaking of world conventional oil production.Robert L. Hirsch, Roger Bezdek, and Robert Wendling, "Peaking of World Oil Production: Impacts, Mitigation, & Risk Management, " Minnesotans for Stability, February 2005, accessed January 31, 2011, http://www.mnforsustain.org/oil_peaking_of_world_oil_production_study_hirsch.htm; also quoted in James Hamilton, *Limitations of the Hirsch Report on Peak Oil*, *Econbrowser*, August 9, 2005, accessed January 31, 2011, www.econbrowser.com/archives/2005/08/limitations_of.html.

When economists read a quotation like this, however, they typically think that something is missing. The quotation says that, at some point "world oil production will no longer satisfy demand." Economists respond that this does not make sense: the price of oil will adjust to ensure that supply equals demand. Similarly, they would say that it is very unclear what "oil demand is expected to grow 50 percent" means. Does this mean a shift in demand (a 50 percent increase in the quantity demanded at every price) or an increase in the equilibrium quantity (the quantity purchased will be 50 percent higher than it is now)? Fundamentally, economists would say that we cannot talk sensibly about the oil market without discussing what will happen to the price of oil. We turn to this question next.

The Hotelling Rule

Suppose you own 1,000 barrels of oil. You want to decide when you should sell them. You could sell them now or hold onto them in the hope that the price of oil will increase. This is a difficult problem, so we begin by making some simplifications. First, imagine you already have the oil in storage, so we can ignore the costs of extracting it from the ground. Second, suppose you can store it for free. Third, suppose you can buy and sell oil in a competitive market.

Even with these simplifying assumptions, it seems as if you are facing a hard decision. In fact, this problem is easier than it looks. If you sell one barrel this year, you get this year's price. You can invest this at the market rate of interest and get (price this year × interest factor). Alternatively, you can store your oil for a year and then sell it next year at next year's price. If this year's price multiplied by the interest rate is greater than next year's price, then you should sell all of your oil this year. If next year's price is higher, you should store all of your oil for sale in the future. By following this rule, you can determine how to make the most money from your stock of oil.

Of course, this decision looks the same to anyone else who is in the same position as you. So if this year's price multiplied by the interest rate is greater than next year's price, then *everyone* would try to sell their oil this year. There would be a huge supply of oil to the market, which would tend to reduce the price this year. If next year's price is higher, then there will be little to no oil supplied to the market this year, which would tend to increase the current price. The only way the oil market will be in equilibrium is if a condition known as the **Hotelling rule** holds. This is an **arbitrage** condition for the use of resource stocks. In the case of a nonrenewable resource sold in a competitive market, with no costs of extraction, no costs of storage, and no uncertainty, the rule states that

The Hotelling rule tells us that the price of the resource increases at the rate of interest. For example, if the nominal interest rate is 6 percent (the **nominal interest factor** is 1.06), then the Hotelling rule says that the price of oil will increase 6 percent per year. The rule has a remarkable implication: it does not matter whether you sell your oil or hold onto it. If you sell your oil, you can earn the market rate of interest on your savings. If you hold onto your oil, it as an asset that yields a rate of return equal to the market rate of interest. Either way, you should expect to get the same return.Oil can be thought of as an asset. In Chapter 10, we discuss how the prices of different kinds of assets are determined. The discussion in that chapter also explains how we can take into account the fact that you don't know the future price of your asset—in this case, next year's price of oil—with certainty.

Complications

The basic Hotelling relationship underlies the pricing of all renewable and nonrenewable resources, but there are many other factors that also come into play. We can in principle build all these other complications into our equation, but this would require mathematics and analysis that go beyond the level of this book. Still, we can give a brief idea of how to incorporate these other factors.

Costs of storage. If it is costly to store the resource from one year to the next, then the price must increase at a rate fast enough to cover the cost of storage as well as the rate of interest.

Costs of extraction. We have ignored the costs of extracting oil from the ground. This cost is not constant in either the short run or the long run. In the short run (say, in a given year), the marginal cost of extracting oil increases when we pump more oil from the ground. (One implication is that the supply of oil at any given time is not perfectly elastic.) As we start to run out of oil, it is likely that the marginal cost of extracting oil from the ground will increase substantially. This, too, must be factored into the arbitrage condition. Marginal extraction costs that increase over time are an additional factor causing prices to increase.

Discovery of new oil fields. If new oil fields are discovered, then the overall supply of oil in the world increases. This increase in supply leads to a decrease in the price. As soon as there is a discovery of a new field, the price of oil jumps to a new Hotelling path. What matters for the pricing of oil is *new information* about existing resources. If the oil companies discover a new oilfield—or come up with a better technique for extracting existing reserves—there will be a decrease in the price of oil as soon as this information becomes known to the market. This is why the price of oil is so sensitive to political changes in the Middle East.

Shifts in the demand curve. The price of oil is also affected by shifts in the demand curve. For example, the increasing prosperity of countries such as China and India is causing the demand for fossil fuels to shift outward. Technological developments are one source of changes in demand. If scientists were to come up with a cheaper source of energy, the demand for oil would decrease.

Market power. The supply of oil is heavily influenced by the decisions of oil-rich countries such as Saudi Arabia and Kuwait. Producers with market power tend to restrict supply to force prices up. The effect of this is to increase prices at all times, which in

turn means that existing stocks of oil will last longer. Market power, in and of itself, does not affect the basic conclusion from the Hotelling rule unless the degree of market power changes over time.

Uncertainty. Finally, we took the stock of oil, the demand curve, and the rate of interest as known with certainty. In reality, of course, all of these are unknown. Changes in the information with regard to any of these variables will lead to changes in the price of oil.

Does the Hotelling Rule Work in Practice?

The Hotelling rule is based on a very simple arbitrage idea, so it is highly compelling. Yet it is sometimes difficult to observe the rule in operation in the data for oil or for any other nonrenewable resource. Figure 14.4.1 shows what has happened to the price of oil in the last 60 years or so. The prices are in 2008 dollars. Oil prices were reasonably steady in the 1950s and the 1960s. The 1970s are sometimes called the “oil shock decade,” and from the graph you can see why. The price of oil jumped to a level that is equivalent to over $100 a barrel in 2008 dollars. Then oil prices fell again and were relatively low in the 1980s and 1990s. The early years of the 21st century saw another big increase in oil prices, with the price reaching record levels in 2008.

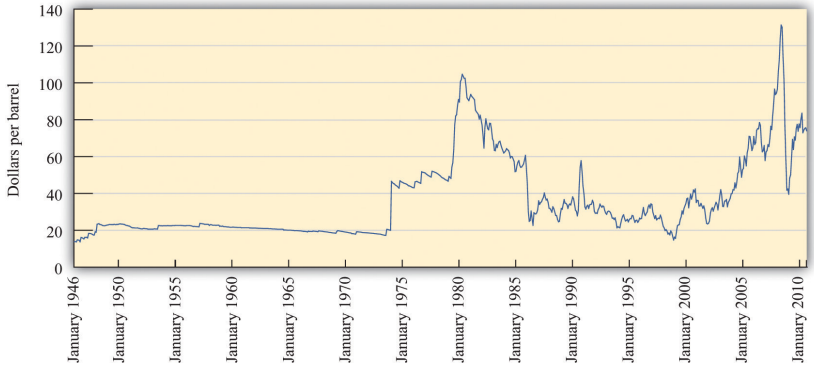


Figure 14.4.1: The Price of Oil (in 2008 Dollars). Oil prices were relatively stable in the 1950s and 1960s, rose rapidly in the 1970s, were low from the mid 1980s until about 2002, and have since increased rapidly.

Source: Federal Reserve Bank of St. Louis FRED database: Spot price of West Texas crude oil (http://research.stlouisfed.org/fred2/data/OILPRICE.txt) deflated by the Consumer Price Index (http://research.stlouisfed.org/fred2/data/CPIAUCSL.txt).

It is not easy to reconcile this figure with the Hotelling rule. We do not see the price of oil increase steadily, as the Hotelling rule seems to suggest. The problem is that the complications that we mentioned in Section "Complications" are quite significant in practice. Over the last several decades, we have seen technological improvements, the discovery of new oil fields, political instability in the Middle East and other oil producing regions of the world, price-fixing by the oil-producing countries, and so on. Most of the time, it seems as if these variables are swamping the pattern we expect from the Hotelling rule.

If we look at other nonrenewable resources, it is likewise difficult to see Hotelling effects. In large part this is because technological improvements have often made resources less valuable, even as they became scarcer. For example, copper became significantly less valuable when scientists developed techniques for transmitting information along fiber-optic cables rather than along copper wire.

For all of these reasons, you should not think of the Hotelling rule as literally describing what will happen to the price of nonrenewable resources in the real world. Instead, you should think of it as explaining one component of the price. If the exhaustion point of the resource is a long way in the future, the Hotelling rule may not play a big role in explaining the price. Other factors that shift the demand and supply curves may explain most of the price variation. Nevertheless, most economists are confident that the Hotelling rule does contribute to resource price changes, and as a resource gets closer to exhaustion, the Hotelling rule will play a bigger and bigger role. James Hamilton, an economist who is an expert on oil, put it as follows in late 2006:

My own view is that, for most of the past century … the resource exhaustion was judged to be sufficiently far off as to be ignored.…I do not infer that the next decade will necessarily be like the previous century. Certainly declining production from U.S. oil reservoirs set in long ago.…

I am not at all prepared to dismiss the hypothesis that [the Hotelling rule has] indeed started to make a contribution to oil prices over the last five years, and will become more apparent over the next five. For example, the announced intention of OPEC producers to cut back production as the price goes below $60 might be most naturally interpreted from that perspective—producers don't see it as being in their interests to sell for less, given what the oil will be worth in the future.See James Hamilton, "Is Peak Oil Irrelevant?," *Econbrowser*, October 24, 2006, accessed January 31, 2011, www.econbrowser.com/archives/2006/10/is_peak_oil_irr.html.

Excludable and Nonexcludable Resources

The Hotelling rule tells us that we expect the price of a nonrenewable resource to rise as we use it up. As its price increases, households and firms have an incentive to substitute other goods for the resource. As the price of oil increases, people switch to other forms of transportation and more fuel-efficient vehicles. More generally, an increase in the price of oil makes other forms of energy—such as wind, solar, or nuclear power—more attractive.

Nothing in this description suggests any failure of the market mechanism. Although we have explained what will happen as we start to run out of a resource, we have not given any reason to suggest that we will use up the resource too quickly. In fact, we have even pointed to one reason we might be using up oil too slowly: if oil producers have market power, they have an incentive to limit the supply to the market and increase the price. The mere fact of using up a nonrenewable resource does not mean that the market is not efficient. To understand the difference between using oil and polluting the atmosphere, we need a new distinction.

Toolkit: Section 31.19

A nonexcludable good (or resource) is one for which it is impossible to selectively deny access. In other words, it is not possible to let some people consume the good while preventing others from consuming it. An excludable good (or resource) is one to which we can selectively allow or deny access.

Go back once more to our example of a smoker in an office. The smoker actually consumes two things: she consumes cigarettes and she consumes the clean air in the office by turning it into dirty air (economists call this "joint consumption"). Clean air, like a cigarette, is a "good." But it is a good with a special property. Under most circumstances, we cannot allow some people access to clean air while denying others access to clean air. Clean air is **nonexcludable**.

Drivers in Mexico City likewise consume clean air: driving leads to less clean air and more dirty air. Like the smoker in the office building and the car drivers in Mexico City, polluting firms consume clean air as well. In this instance, we can think of clean air as being another input into the production process. *Air pollution is a process by which the nonexcludable resource that we call "clean air" is consumed by households and firms*. This is a somewhat unfamiliar way to talk about environmental pollution, but it makes sense once you think about it. At the end of the production process, there is less clean air than before because the firm has made some air dirty. The polluting firm uses clean air in its production process.

Partially Excludable Goods

Other examples of nonexcludable goods are fireworks displays, lighthouses, fish in the ocean, concerts in a public space, and broadcast television. These examples prompt the following observations.

- Whether or not something is **excludable** can change with advances in technology. Early television and radio signals were nonexcludable because they could be accessed by anyone with a suitable receiver. But cable television and Internet radio are excludable because it is technologically possible to restrict access to certain users.
- Excludability is not a hard-and-fast line. Think about roads, for example. Toll roads are excludable; it is possible to control access to toll roads because there are relatively few entrances and exits. By contrast, roads in a neighborhood or in the middle of a city are nonexcludable; it is almost impossible to imagine allowing selective access to different city blocks. The more different entrances there are to a road, the harder it is as a practical matter to make the road excludable. Similarly, fish in the ocean are neither fully excludable nor completely nonexcludable. Fish in international waters can be fished by anyone, but coastal waters are under the jurisdiction of individual nations that may restrict fishing rights.

Going back to our smoking example, we cannot make the air in a single office excludable, but we can perhaps do the next best thing: within a single building, we could certainly imagine permitting smoking in some offices and not in others. In other words, we could define property rights differently in different offices and in this way make clean air partially excludable. In some cases at least, this might be a practical way to accommodate both smokers and nonsmokers. We see something to this effect in airports and

other public spaces. Very often you will find that most of an airport is designated as nonsmoking (meaning that the property rights to clean air have been allocated to nonsmokers), but there are designated rooms or areas where smokers have the right to consume —that is, use up—the clean air.

Externalities and Nonexcludable Goods

There is a connection between nonexcludability and externalities. As our clean air example suggests, goods—like clean air—that are nonexcludable will tend to be overconsumed. When a good is nonexcludable, its marginal private cost is zero. If its marginal social cost is positive, then there is a negative externality from the consumption of a good. As these observations suggest, there are typically negative externalities associated with nonexcludable (or partially excludable) goods.

The flipside is that nonexcludable goods will be underproduced. To see why, think about whether a private firm will want to produce nonexcludable goods. For goods that are completely nonexcludable, the answer is no. There is no return from producing such goods because they cannot be sold for profit. Anyone can consume these goods without paying for them. Goods that are partially nonexcludable may be produced, but they will be produced in insufficient quantities.

Should We Worry about Running Out of Resources?

Because clean air is a nonexcludable resource, there are externalities associated with its use, and we know that this implies inefficiency. Economics provides a clear argument for why we end up overconsuming clean air. Similarly, many people take it as self-evident that we are overconsuming oil and other natural resources. The arguments we made with respect to clean air, however, do not translate to the case of oil. It is true, of course, that consumption of oil may have negative externalities because it leads to pollution. But the problem in this case has nothing to do with the fact that oil is a nonrenewable resource.

It is difficult to buy and sell nonexcludable goods—after all, why would anyone buy something that is free? This makes clean air unlike most of the other goods we study in economics. Excludable goods and services, by contrast, are those that we can prevent someone else from enjoying. Excludable goods—like oil—are easily traded in markets.

When people worry about running out of resources, they are speaking, for the most part, about excludable resources: those that are mined from the ground or grown on farms and privately owned forests. Economic analysis suggests that this worry may be largely misplaced. Unless we have specific reasons to think that markets are misallocating these resources, there is no particular cause for concern or government intervention.

There are some exceptions to this principle. Oil is sometimes drilled from pools that are not uniquely owned. Different companies may be able to access the same oil pool. In this case, the oil in the pool is partially nonexcludable—any company with access to the pool is able to drill it. It follows that there is a negative externality associated with the drilling of the oil in this case, so we expect there to be too much drilling.

Renewable Resources

Oil is a nonrenewable resource. New oil reserves are not going to be created (at least over any period of time relevant to human beings). Over time, the stock of a nonrenewable resource can only decrease, never increase. By contrast, a **renewable resource** is one that regenerates over time.

Whether the stock of a renewable resource grows, shrinks, or stays constant depends on the balance between how fast we use up the resource and how quickly it regenerates. If a resource is depleted at a rate faster than it regenerates, the stock will decrease. If a resource regenerates faster than it is depleted, the stock will increase. Examples of renewable resources are stocks of fish in the ocean and forests. Clean air and water are also examples of renewable resources: if the source of pollution disappears, then a polluted river or polluted air will tend to improve naturally over time. Of course, this process can take quite a while.

In some cases resources may be renewable only above a certain threshold. For example, we can overfish a particular species of fish to the point of extinction. So although a renewable resource will normally naturally regenerate over time, it can turn into a nonrenewable resource if the stock falls below a critical level.

As with nonrenewable resources, what matters most for efficiency is whether a resource is excludable or nonexcludable. If a renewable resource is excludable, then a modified version of the Hotelling rule applies. The arbitrage condition in this case takes into account the fact that the resource stock gets larger as time goes by. If the resource is nonexcludable, however, then all the problems we saw earlier in the chapter come into play. The marginal social cost of using the resource will typically exceed the marginal private cost, leading to overconsumption of the resource.

Accumulable Resources

Nonrenewable environmental resources get used up over time. With good stewardship, which usually requires well-established property rights, the stocks of renewable resources can stay constant or even increase. Environmental resources, however, typically cannot grow without limit. There is a limit to how many fish there can be in the ocean or how clean the air can be.

Economists also identify a third class of resources, called **accumulable resources**. Such resources can increase (more or less) without limit. The most important examples are physical capital (factories, machines, etc.), **human capital** (the skills and know-how of workers), and general human knowledge.

Even though there is no reason to think that the depletion of a nonrenewable resource is necessarily a source of economic inefficiency, it does not follow that the depletion of natural resources is without cost for an economy. For example, as oil becomes more expensive, it costs more to produce other goods and services that use energy as an input. From an economic perspective, this is a lot like a worsening of technology.

Crucially, though, we are accumulating resources such as physical capital, human capital, and knowledge at the same time that we are running down our stocks of oil, coal, and other natural resources. In the last couple of centuries, the economy of the world has grown substantially because we have been able to accumulate resources at a rate that far outpaced our depletion of the natural environment. For much of that time, we had natural resources in abundance, so their price stayed low. At the same time, we increased our technological know-how at unprecedented rates.

There is no guarantee, from economics or anywhere else, that this state of affairs will continue. We will continue to run down our stocks of nonrenewable resources. If technological advance (and other accumulation) fails to keep pace, then we might well see the prices of these goods increase. More generally, the fact that we have seen our economies growing in the past is no guarantee that they will continue to grow in the future.

Technological optimists point to the rapid growth of our knowledge and expect this to continue. If they are right, then technological advance will keep our living standards growing even though we are depleting our stocks of resources. Technological pessimists observe that the last two centuries are an anomaly if we look at the broad sweep of human history. If the rate at which we accumulate knowledge, human capital, and other accumulable resources were to decrease significantly, then the drag on the economy from declining resources would begin to seem substantial. Economics does not allow us to predict the future, and we do not know who is right.

Key Takeaways

- According to the Hotelling rule, a resource should be extracted so that the rate of price increase in the resource should be the same as the interest rate.
- A good is nonexcludable if it is impossible to deny access to it.
- An accumulable resource is one that can be increased over time with investment. Leading examples include physical capital and human capital.

Exercises

1. If the price of oil is decreasing while the interest rate is positive, is the Hotelling rule violated?
2. Is the Internet a nonexcludable good?
3. Will technological advance necessarily offset the depletion of the stocks of natural resources?

This page titled 14.4: Renewable, Nonrenewable, and Accumulable Resources is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

14.5: End-of-Chapter Material 14.5: End-of-Chapter MaterialEdit Update titlecancel

In Conclusion

Economists and environmentalists sometimes do not see eye to eye. Economists think environmentalists often focus on the wrong problems, and environmentalists think economists place too much faith in markets. Yet economics is the science that helps us understand why some environmental problems are among the most important and difficult that we face.

When economists look at excludable resources, for which property rights are well defined, they tend to be less concerned. It is certainly possible that we will run out of oil and other nonrenewable resources. But that in itself does not signal a problem. What matters is whether we are using our resources efficiently or inefficiently. Perhaps the best thing for us as a society to do is to use up our resources quickly. More important, as oil becomes scarce, we know that market prices will force users to economize on oil and look for substitutes instead.

We are assuredly not saying that an economy is better off with fewer resources. We would always like to have more of an exhaustible resource. The most important question, though, is how to best use the resources we have. Markets can sometimes provide a good answer to this question. If markets work properly, sending the correct signals to producers and consumers, then market allocation will be efficient.

Economists worry a great deal more about environmental problems where resources are nonexcludable. Pollution of the air, pollution of rivers and oceans, biodiversity loss, overfishing, and climate change are all examples of environmental problems for which we cannot rely on markets. To an economist, it is not surprising that markets fail in these cases. All of these resources are nonexcludable. When resources are nonexcludable, market allocations will not typically be efficient, and there may be a role for government to try to solve these problems.

Key Link

- James Hamilton blogs: http://www.econbrowser.com

exercises

1. In the example about second-hand smoking in the office discussed in 14.1 Section "Smokers, Nonsmokers, and the Coase Theorem", how would you use the Coase theorem to determine how much smoking should occur if there are many people in the office? What difficulties do you see in trying to apply this theorem?
2. (Advanced) Consider Table 14.2.1 "The Payoffs in a Social Dilemma Game". Suppose your payoff from taking public transportation when everyone else is taking public transportation is $3 rather than $1. Why is it harder to predict what will happen in this situation? See the toolkit. [Hint: Look at what decision you make when others take public transportation and when they don’t.]
3. Consider some situations that might arise in your college or university. Which of the following is an example of an externality?
    1. In a class that is graded on a curve, you study harder and get a better grade, so others get a worse grade.
    2. At examination time, lots of people want to study, so it is difficult to find space in the library.
    3. You find it difficult to understand your professor’s accent.
    4. Your favorite television show is on the night before you have a big test, and you can’t decide whether to watch or study.
    5. Other people ask good questions in class, which makes the class more interesting for you.
    6. Everyone is selling his or her used chemistry textbook, so you can buy one at a cheap price.
    7. Late-night parties in neighboring dorm rooms are preventing you from sleeping.
4. Give three reasons why the marginal social cost of driving is greater than the marginal private cost of driving.
5. There are many endangered species in the world, such as the white tiger and the sea otter. Why are these species endangered whereas cows and sheep are not?
6. Suppose you meet some friends at a pizza restaurant. You are all very hungry, and you know the wait for the pizza may be long. At the next table, you see some people who have just finished one pizza, and another is about to be delivered to their table. Explain how you could conceivably create a market that would make both your table and the other table better off. What is the nature of the transaction costs that prevent such trades from happening in restaurants in real life?

7. One reason people choose to buy large vehicles, such as SUVs, is because they are safer if you are involved in an accident. Explain how this could give rise to a social dilemma game.
8. Copy Figure 14.5.1. Indicate on that diagram how large a subsidy would be required to induce the socially optimal quantity.
9. If you were to graph a measure of pollution on one axis and a level of economic activity (such as the real gross domestic product per person) on the other, what type of relationship do you think you would find? How would you explain this relationship?
10. Think of an externality that arises in a college dorm. What market can you think of that would (or could) eliminate any inefficiencies from that externality?
11. Using Table 14.2.1 "The Payoffs in a Social Dilemma Game", think of another example of a social dilemma game. What are the choices of the people, the payoffs, and the outcomes?
12. Explain why tradable permits for the right to dump garbage into a river would be more efficient than telling producers how much they are allowed to dump in the river. How is the total quantity of tradable permits determined?
13. Suppose you had a system of tradable permits that allocated permits to those who needed them (that is, the firms that polluted a lot). Would that be a good system to provide an incentive for firms to undertake investment in clean technologies?

Economics Detective

1. In recent years, the European Union has discussed requiring airlines to obtain permits to emit pollution in order to operate. What is the status of this proposal? How might it work?
2. Pick a nonrenewable resource and try to find out what has happened to its price over the last few decades. Do you see any evidence of the Hotelling rule? If not, try to find out what have been the main factors affecting the price of this resource.
3. If you live in the United States, try to find the Superfund site nearest to where you live. What kind of pollution is at the site? (If you live somewhere else, pick a US city of your choice.)
4. What is a “carbon tax”? What is the economic rationale for such a tax?

This page titled 14.5: End-of-Chapter Material is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

14.6: Appendix- An Example of the Hotelling Rule in Operation 14.6: Appendix- An Example of the Hotelling Rule in OperationEdit Update titlecancel

The Hotelling rule states that *the nominal price of oil will increase at the nominal rate of interest*. This seems a little bit mysterious. After all, we also think that the price of oil is determined by demand and supply in a market. In fact, these two approaches to the price of oil are completely consistent.

To see how, imagine that the worldwide demand curve for oil each year is as follows:

$$price = 1000 - 0.000001 \times quantity,$$

where we measure the price in dollars and the quantity in barrels. Individual suppliers are willing to sell oil at the price determined by the Hotelling rule. At higher prices, they will supply a lot of extra oil to the market; at lower prices, they supply no oil to the market. In other words, the supply of oil will be perfectly elastic at the market price.

But this is not enough to determine the market price. We know the market price this year only if we know the price next year. And we know the price next year only if we know the price the year after that, and so on. Unless we know the price at some future date, we cannot calculate the price today. But here is the trick: we do know that the price will eventually increase to the **choke price** (which is $1,000 per barrel in our example).

Over time, the price increases at the rate of interest, as implied by the Hotelling rule. If the demand curve is unchanged from year to year, the quantity demanded gets smaller. This makes sense: the increasing price reflects the increasing scarcity of the resource. The price increases until it eventually reaches the choke price, where the quantity demanded decreases to zero. But when will that happen? It has to happen when the last barrel of oil is being sold. If there is oil left over at that time, the price would decrease below the choke price. And if we ran out of oil before then, the price would have to hit the choke price earlier. Anything else would cause a mismatch between demand and supply. Thus the price today needs to be just right so that when the last barrel of oil is being sold, the price increases to exactly the choke price.

Figure 14.6.1 shows an example based on this demand curve. Suppose that the total amount of oil available in the world at the end of 2010 is 29.2 billion barrels, which is owned by many competitive suppliers. How will this oil get released to the market? In 2011, 900 million barrels of oil are sold, so the market price is therefore $100. The price then increases at the rate of interest, which is 5 percent per year in our example. By the year 2044, less than 4 billion barrels remain, and the price of oil has increased to $500 per barrel. In this particular illustration, we run out of oil in slightly under half a century: the last 9 million barrels of oil are sold in the year 2058, at $991 per barrel.

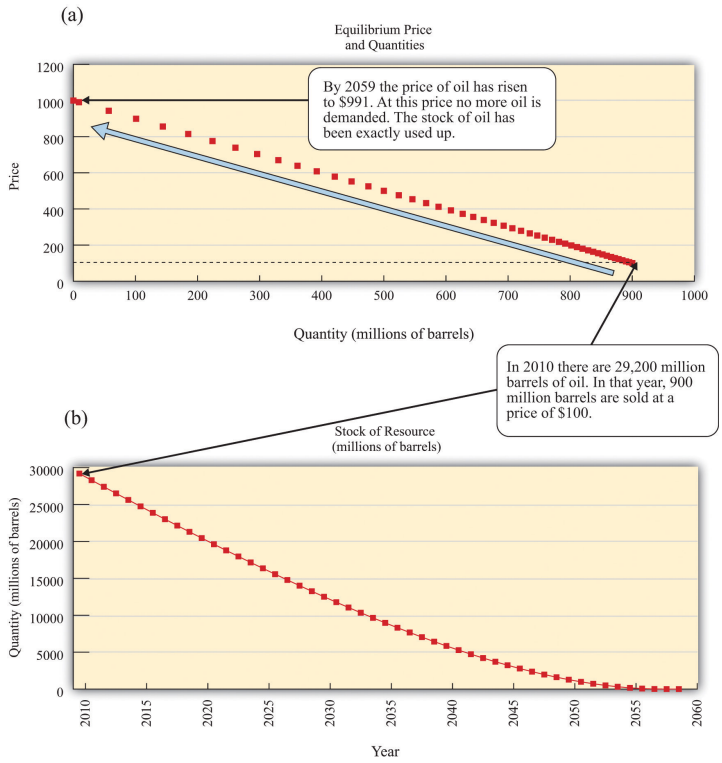


Figure 14.6.1: An Example of Oil Depletion According to the Hotelling Rule. These figures show an example of resource depletion according to the Hotelling rule. (a) The price and quantity of oil traded each year trace out the demand curve. (b) The stock of oil over time represents the supply curve. Initially, there are slightly under 30 billion barrels of oil available. In the first year, 90 million barrels of oil are sold in the market, at $10 per barrel. The price then increases at the rate of interest. In the year when the price reaches the choke price (2059), the stock of the resource is fully used up.

This page titled 14.6: Appendix- An Example of the Hotelling Rule in Operation is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

15: Busting Up Monopolies

15.1: Working for Antitrust
15.2: Market Power and Monopoly
15.3: Patents and Copyright
15.4: Markets with a Small Number of Sellers
15.5: End-of-Chapter Material

This page titled 15: Busting Up Monopolies is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

15.1: Working for Antitrust 15.1: Working for AntitrustEdit Update titlecancel

You have probably never considered working for the Department of Justice, yet here is a job opportunity open to everyone.

If you have information about a possible antitrust violation or potential anticompetitive activity, use the following questions as a guideline to describe your complaint:

- What are the names of companies, individuals, or organizations that are involved?
- How do you believe they have violated the federal antitrust laws?
- Can you give examples of the conduct that you believe violates the antitrust laws? If so, please provide as much detail as possible.
- What is the product or service affected by this conduct? Where is the product manufactured or sold, or where is the service provided?
- Who are the major competitors that sell the product or provide the service?
- What is your role in the situation in question?
- Who is harmed by the alleged violations? How are they harmed?See US Department of Justice, "Reporting Antitrust Concerns," accessed March 14, 2011, http://www.justice.gov/atr/contact/newcase.html.

You may never have heard about antitrust laws. You may have only a vague idea of what antitrust is. Yet these laws have a direct impact on your day-to-day life because they can have a significant impact on the prices you pay for goods and services. As the Department of Justice explains,

Most states have antitrust laws, and so does the federal government. Essentially, these laws prohibit business practices that unreasonably deprive consumers of the benefits of competition, resulting in higher prices for inferior products and services…

When competitors agree to fix prices, rig bids, or allocate customers, consumers lose the benefits of competition. The prices that result when competitors agree in these ways are artificially high; such prices do not accurately reflect cost and therefore distort the allocation of society's resources. The result is a loss not only to U.S. consumers and taxpayers, but also the U.S. economy.US Department of Justice, "Antitrust Enforcement and the Consumer," accessed March 14, 2011, www.justice.gov/atr/public/div_stats/211491.

Antitrust laws in the United States are principally codified in three acts of Congress—the Sherman Antitrust Act, the Clayton Act, and the Federal Trade Commission Act—and the Federal Trade Commission and the Department of Justice enforce each act. In other countries, similar government agencies perform these same tasks.

How might you become involved in pursuing antitrust violations? Without knowing the provisions of these laws, it is hard to see how you could ever detect and report violations. Fortunately, the Department of Justice provides some further guidance: look for situations where the price of a good is in excess of the marginal cost of producing that good. Remember that marginal cost is the cost of producing one additional unit of output. In a competitive market, firms set prices equal to marginal cost, but when they have market power, firms set prices in excess of marginal cost. Chapter 7 contains more information. The Department of Justice's suggestion reflects this conclusion: price in excess of marginal cost is a likely indicator of market power.

In practice, it is more complicated. For example, the cost of producing one more compact disc (CD) is the cost of the material it is made from, the cost of burning the CD, and the cost of the jewel case in which it is packed. These costs are very small: no more than a few cents. Yet, if you have bought a CD recently, you probably paid between $10 and $20 for it. Should you be reporting your local CD retailer to the Department of Justice?

In fact, the government sometimes actively protects and creates market power, despite the fact that it has entire divisions devoted to encouraging competition. It creates market power through patent and copyright laws, which prevent people from copying inventions and created works (like music or books). In the case of CDs, the government grants copyright protection because the creation of the very first copy of a CD is very costly. When a CD is produced, there are enormous music creation costs incurred by the musicians and the company recording and producing the music. These costs are much more sizable than the actual cost of producing the CD you purchased. From this perspective, the high prices are needed to make it worthwhile for the artist to incur the creation costs.

You can perhaps sense the tension. The costs of producing the very first CD are high, but the cost of producing the thousands (or millions if the artist is successful) of *copies* of the first CD are relatively small. Should the price reflect the marginal cost of that last CD or should it be higher to cover the creation costs as well? This argument was vividly illustrated in 2001 when music file

sharing first became popular. A firm called Napster (http://music.napster.com) supplied a technology that facilitated the sharing of music through the Internet. Napster essentially reduced the marginal cost and the price of a song all the way down to zero. Napster's technology was in direct competition with the record companies, and a legal battle ensued. The final ruling forced Napster to block its file sharing and effectively ended its ability to share music. Other peer-to-peer networks have replaced it, however, and a lot of music is still available for free on the Internet.

In this chapter, we investigate these varied aspects of competition policy. This might seem like an arcane topic, but it has a huge impact on the prices you pay for goods and services. It is competition policy that keeps the price of CDs high and the price of airline tickets low. In this chapter, we evaluate the impact of competition policy on the economy and answer the following questions.

- What happens in markets when there are only a few producers?
- What are the different kinds of competition policy carried out by the government?

Road Map

To make sense of competition policy, we need to first understand what firms would do if there were no antitrust laws constraining them. We therefore begin by looking at economic outcomes in the absence of government protection. We first contrast a market in which there is a single seller with a market that is serviced by many sellers—that is, a competitive market. From this comparison, we can understand the basis of antitrust laws.

Next, we look at why the government sometimes actively promotes market power through patents and copyright. Specifically, we show how such laws can encourage innovation and creation of intellectual property. Finally, we look at situations where there are only a few competing firms. We explain how the outcome is different depending on whether firms choose to set prices or to set the quantity that they produce, and we again look at the role of government policy.

This page titled 15.1: Working for Antitrust is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

15.2: Market Power and Monopoly 15.2: Market Power and MonopolyEdit Update titlecancel

Learning Objectives

1. What is a monopoly?
2. What is the outcome when there is a monopoly?
3. What are the policies taken to deal with monopolists?

When there are many buyers and sellers of a homogeneous product, we have a **competitive market** (Figure 15.2.1). Equilibrium is at the intersection of supply and demand. At the equilibrium level of output, households enjoy **buyer surplus**, given by the marked area below the demand curve and above the equilibrium price. The surplus arises from the fact that some buyers are willing to pay more than the equilibrium price for the good.

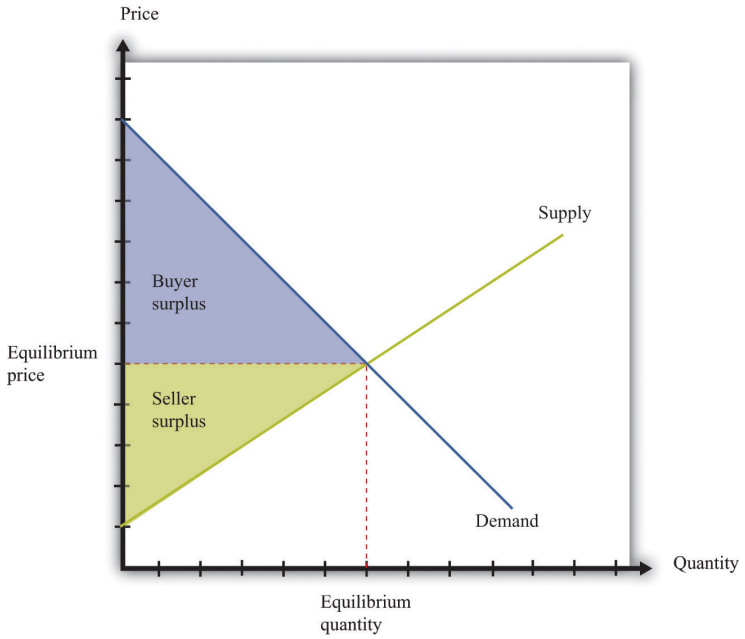


Figure 15.2.1: The Competitive Market Outcome. At the equilibrium quantity in a competitive market, all gains from trade are exhausted.

Surplus also flows to firms. Remember that a competitive firm's **individual supply curve** is equal to its marginal cost curve. In Figure 15.2.1, the supply curve slopes upward because marginal cost is increasing. Firms obtain surplus because they can produce output at a marginal cost that is less than the equilibrium price of the good. This is shown as **seller surplus** in the figure.

At the equilibrium quantity, there are no further **gains from trade**. Producing more output would not increase the total surplus. In fact, producing more output would reduce the surplus: the marginal cost of producing more output would exceed the marginal valuation of extra output. Producing less output would likewise lower total surplus because the buyers and the sellers would lose some of their surplus.

The competitive market provides a benchmark because it leads to an efficient outcome. But very few markets are truly competitive. In most markets, firms possess some market power. This means, in particular, that they are able to set a price above marginal cost without losing all of their sales. In a competitive market, the *demand curve facing a firm* is perfectly elastic at the market price, whereas when a firm has market power, its demand curve slopes downward.

Toolkit: Section 31.10

You can review the concepts of buyer surplus, seller surplus, and the gains from trade, in the toolkit.

The Definition of a Market

At the other extreme to the competitive market is the case of **monopoly**. A monopoly arises when there is a single producer in a market. The demand curve facing a firm is, in this case, the same as the **market demand curve**.

The definition of a monopoly seems easy, yet it is hard to decide exactly what we mean by "a market." Think about diamonds. It is often said that the De Beers Corporation is a monopolist in the market for diamonds because this company controls most of the world's diamond supply. Yet, depending on how broadly or narrowly we define the market, De Beers has either a lot of competitors or only a few. We could define the market very narrowly as "De Beers-branded diamonds" (De Beers is able to brand its diamonds by using certificates of authenticity). De Beers would then be a monopolist by definition. We could define the market more broadly as "all diamonds," in which case De Beers has substantial market power but does not have a total monopoly. This is perhaps the most natural definition to use, yet it misses the fact that other precious stones, such as emeralds, rubies, or opals, are also possible substitutes for diamonds. An even broader market definition is "the market for precious stones." We could go even further still and consider De Beers as part of the market for luxury goods, competing with, say, Louis Vuitton bags and Ferrari sports cars. We illustrate this in Figure 15.2.2.

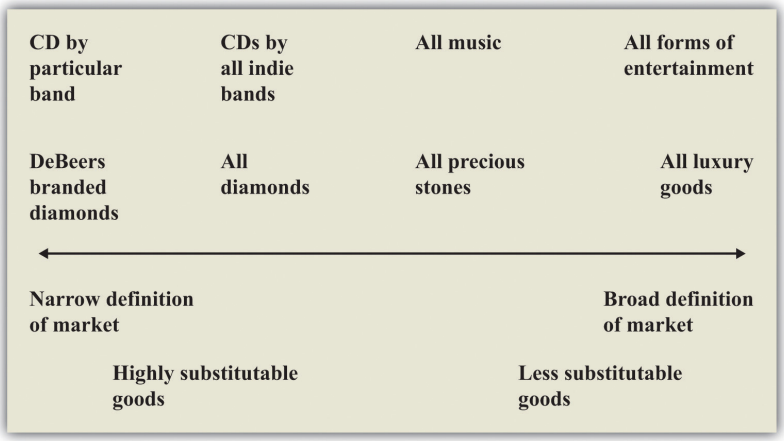


Figure 15.2.2: The Extent of Competition Depends on the Definition of the Market. There is no hard-and-fast definition of a market, but goods that are highly substitutable for each other are generally taken to be in the same market.

Figure 15.2.2 also gives an illustration for the case of music. By definition, a given indie band has monopoly power over its own music. So again, with a very narrow definition of the market, we would say that the band is a monopolist for its own songs. But that band also competes with other indie bands for consumers' dollars, so another definition of the market would be "CDs by all indie bands." Again, we could define the market still more broadly as "all music" or even "all forms of entertainment."

One of the difficult tasks of antitrust authorities is deciding what definition to use for a market. There is no single correct way to define a market, and the extent of market power depends on where we choose to draw the line. For this reason, the term *monopoly* is somewhat misleading. Almost all firms are monopolists if we adopt a sufficiently narrow definition of the market. The decision about how to set prices in the presence of market power, which is sometimes called the *monopoly pricing problem*, actually applies to nearly every firm in the economy. What matters in practice is determined by the *extent* of a firm's market power.

The extent of a firm's market power depends on two things: (1) the number of firms that potentially compete with it and (2) the extent to which those other companies produce close substitutes for a firm's product. A strip mall at the edge of town might contain several different fast-food restaurants. Each restaurant enjoys some market power because the food it has available is slightly different from that at the other establishments. But many consumers are likely to be willing to substitute reasonably freely among the different restaurants, so the market power of each restaurant is limited. Contrast that with a fancy French restaurant with a famous chef. That restaurant enjoys much more market power because there are almost certainly fewer comparable restaurants nearby and the meals at other high-end restaurants are not such close substitutes.

Market power can stem from many different sources. A firm has market power if it is selling a unique product or service. A firm can also derive market power from its superior or exceptional service. If a firm creates customer loyalty, either through exceptional service or by loyalty programs (such as frequent flyer miles), this is also a source of market power. Retail firms can derive market power from their location: your local corner store has some market power because you would prefer to walk there rather than drive to a more distant supermarket.

Firms devote a lot of resources to establishing, protecting, and increasing their market power through advertising and other forms of marketing. Many firms spend a great deal of money on developing their brand image and brand awareness. Sporting goods companies such as Nike and Adidas are classic examples. When customers are loyal to brands, firms have market power.

Pricing with Market Power and the Monopoly Outcome

The managers of a monopoly firm must pick the point on the demand curve that will maximize the firm's **profits**—the total revenues of the firm minus its costs of producing its output. We can think about the firm choosing either its level of output or its price. If a firm chooses how much to produce, then the price of its product is determined from its demand curve. (Remember that the demand curve facing the firm and the market demand curve are the same thing for a monopolist.) If a firm chooses a price for its product, then the quantity produced is determined by the demand curve. For now, we tell the monopoly story assuming it chooses the quantity of output.

When a firm is maximizing its profit, producing a little more or a little less output will not increase the firm's profit. This means that the extra revenue from producing one more unit of output is exactly equal to the cost of producing one more unit of output: marginal revenue equals marginal cost. We call this level of output the *profit-maximizing quantity* and illustrate it in Figure 15.2.3. The price the monopolist sets, termed the *profit-maximizing price*, can then be read from the demand curve. Once a monopolist determines how much to produce, the price of its output is determined by the maximum amount that consumers are willing to pay for the good.

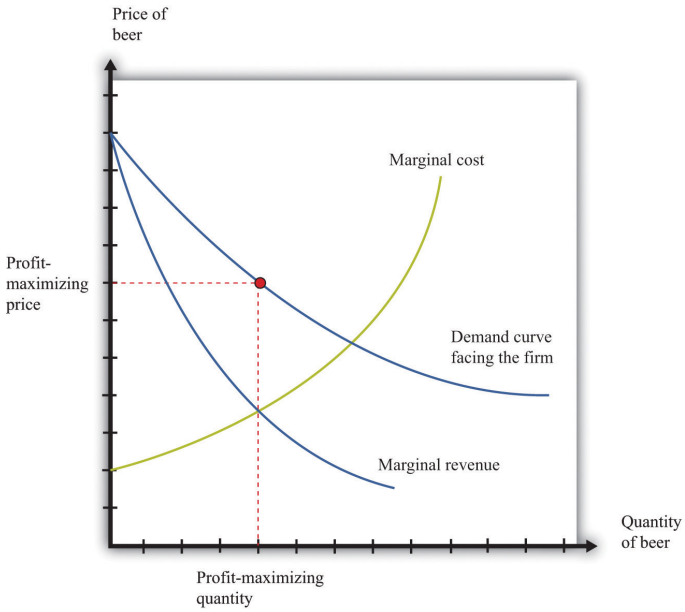


Figure 15.2.3: The Monopoly Outcome. A monopolist produces a quantity such that marginal revenue equals marginal cost. The price is determined by the demand curve.

Toolkit: Section 31.15

The principles of pricing with market power are explained in more detail in the toolkit.

Distortions Due to Market Power

We have competition policies in market economies because market power often leads to an outcome that is not efficient. To understand this distortion, compare the monopoly outcome with the competitive market outcome we saw earlier. Figure 15.2.4 is Figure 15.2.1 with some areas of the figure labeled.

- We indicate the buyer surplus—the area between the price set by the monopolist and the demand curve. Even though there is a monopolist in this market, buyers can still enjoy some surplus if their **marginal valuation** of the good exceeds the price they pay. So while the price may be too high in the monopoly case, there still remains some surplus flowing to buyers.
- There is an area labeled **variable cost**—the total area under the **marginal cost** curve.
- There is an area labeled seller surplus (monopoly profit)—the difference between revenues at the monopoly price and variable costs. The total revenue received by the monopolist is the price times the quantity sold. This is a rectangle. Monopoly profit, as

shown in the figure, equals this rectangular area minus variable costs.

- The difference between the competitive and monopoly outcomes creates a social loss—the triangular area labeled deadweight loss in Figure 15.2.4. The demand curve reflects consumers' marginal valuation of the good. Below the competitive quantity, this marginal valuation is greater than the marginal cost of producing the good. This means that there are potential gains from trade that are not being realized.

Toolkit: Section 31.1, Section 31.11 and Section 31.14

You can review in the toolkit (1) the correspondence between demand and marginal valuation, (2) the meaning of efficiency and deadweight loss, and (3) the definitions of variable cost and marginal cost (and the relationship between them).

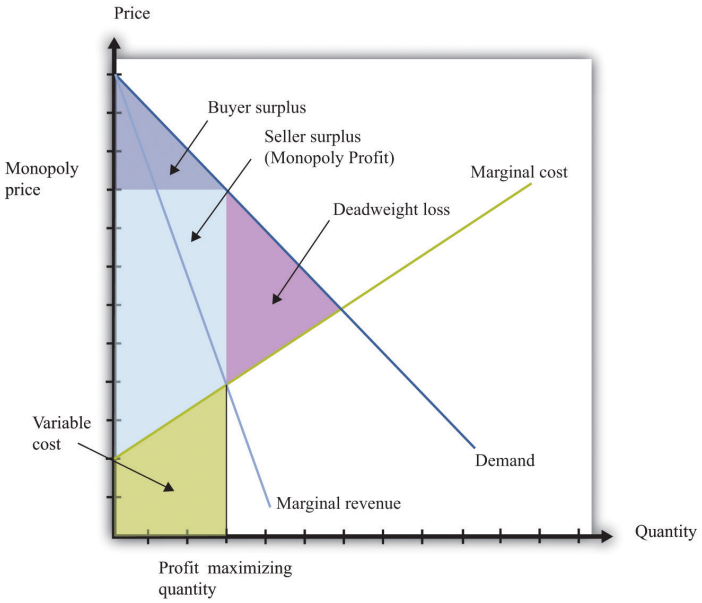


Figure 15.2.4: Distortions Due to Market Power. A monopolist produces a quantity that is less than socially optimal: there is a deadweight loss.

Comparing Figure 15.2.1 with Figure 15.2.4, we see that—relative to the competitive outcome—the monopolist charges a higher price and produces a lower quantity. Since the competitive outcome was efficient, the monopoly outcome is inefficient. Households in the economy would be better off if the monopolist produced at the competitive level of output and sold at the competitive price.

Competition Policy toward Monopolies

When a firm has market power, there is a distortion in the market: prices are too high, and output is too low. Antitrust laws are designed to address precisely this situation.

Antitrust

By virtue of the Sherman Antitrust Act of 1890, the US government can take legal action to break up a monopoly. In 1902, President Theodore Roosevelt used the Sherman Antitrust Act as a basis for trying to break up the monopolization of railway service in the United States. The resulting legal case, known as *Northern Securities Co. v. United States*, involved two key elements: restraint of trade and interstate commerce.

Multiple providers of rail service had joined together in the Northern Securities Company, resulting in a restraint of trade. The Supreme Court decisionCornell University Law School, Legal Information Institute, "Northern Securities Co. v. United States (No. 277)," accessed March 14, 2011, http://www.law.cornell.edu/supct/html/historics/USSC_CR_0193_0197_ZS.html. provides direct discussion of the anticompetitive nature of the creation of this company, noting that the company had been set up to eliminate competition among the different providers. Further, this restraint of trade had interstate implications because the railway lines themselves were not contained within any one state. This was a key point of jurisdiction: the Sherman Antitrust Act spoke directly about the effects of market power on interstate trade. Under the US Constitution, interstate commerce is part of the responsibility of the federal government.

Another famous antitrust case decided by the Supreme Court (http://supreme.justia.com/us/221/1/case.html) centered around the breakup of the Standard Oil Trust in 1911.Justia.com, US Supreme Court Center, "Standard Oil Co. of New Jersey v. United States, 221 U. S. 1 (1911)," accessed March 14, 2011, http://supreme.justia.com/us/221/1/case.html. John D. Rockefeller founded Standard Oil in 1870 soon after the discovery of oil in Ohio. By the late 1870s, the company controlled nearly 90 percent of the refineries in the United States. Shortly thereafter, the Standard Oil Trust was formed with the idea that the individual shareholders in a group of companies would come together to create a single company. This company would then jointly manage all the companies that joined the trust. In doing so, a monopoly was created. (This, by the way, is where the term *antitrust* comes from.)

Once again, the issue of jurisdiction played a role. Standard Oil was originally deemed a monopoly by the Ohio Supreme Court, which ordered it to be broken up. Instead, the company was simply reorganized as a New Jersey–based corporation. Because the company had moved to another state, the Ohio ruling became irrelevant, and the federal government had to step in to prevent unlawful interstate trade. The court ruled that the trust be dismantled. In the modern era, similar concerns arise across national borders: if two companies merge in Europe, for example, US antitrust authorities still take an interest, and vice versa.

These cases may seem like ancient history. Yet they remain very relevant today for at least two reasons. First, the legal system relies heavily on precedent, meaning that cases decided today depend on past rulings in similar cases. The actions of the Supreme Court in these two cases created a significant precedent for the present-day antitrust actions of the federal government. Second, antitrust actions continue to this day. A recent case again alleging a violation of the Sherman Antitrust Act was brought against Microsoft Corporation by the US government in 1998. Microsoft was charged with monopolizing the software market through its Windows operating system. As stated in the case, "Virtually all major PC manufacturers find it necessary to offer Microsoft operating systems on most of their PCs. Microsoft's monopoly power allows it to induce these manufacturers to enter into anticompetitive, long-term licenses under which they must pay royalties to Microsoft not only when they sell PCs containing Microsoft's operating systems, but also when they sell PCs containing non-Microsoft operating systems."A full description of this aspect of the case is contained in John Kwoka and Lawrence White, eds., *The Antitrust Revolution* (New York: Oxford University Press, 2004).

One noteworthy element of the case concerned the definition of the market. After all, there are competing operating systems to Microsoft, most notably Apple Computer's operating system. If the court viewed these products as substitutes for Microsoft's operating system, then the claim that Microsoft was a monopolist would be less clear. Interestingly, the court ruled that Apple was not in the same market as Microsoft.

Price Discrimination

Because there are people willing to pay more than the marginal cost, it seems as if the monopolist in our earlier example is leaving some money on the table. Is there a way for the monopolist to make higher profits? The answer to this question hinges on an assumption that was implicit in our analysis: we suppose that the monopolist has to sell every unit of the good at the same price. In many cases, however, firms sell different units of their product at different prices. This is known as **price discrimination**, and it can arise either because (1) a firm sells to different customers at different prices or because (2) a firm sells different units at different prices to the same customer. There are numerous examples of both kinds of price discrimination. An example of the first is discounts offered to senior citizens or students. An example of the second is quantity discounts such as "buy two, get one free."

Firms have developed many creative forms of price discrimination, and we could easily fill a whole chapter with them. Our main focus here, however, is to understand how price discrimination operates in a monopoly market. Let us think about a case where the monopolist faces a **unit demand curve**: each buyer either purchases a unit of the good or does not buy the good at all. The downward-sloping demand curve comes from the fact that different individuals have different valuations of the good.

> **Toolkit: Section 31.1**
>
> You can review unit demand in the toolkit.

Now imagine what the monopolist would do if she knew everyone's individual valuations and was also able to charge a different price to different buyers. In this case, we obtain a remarkable result. If the monopolist could sell each unit at any price she wanted, then she would charge each individual his valuation of the good, with striking consequences.

- The monopolist captures the entire buyer surplus. Compared to the competitive case, there is a redistribution of surplus from buyers to the monopolist.

- The deadweight loss from the monopoly is eliminated. The monopolist produces the same level of output as in a competitive market because it is worthwhile to sell to anyone whose valuation exceeds marginal cost.

Buyers obtain no surplus because each buyer surplus equals his valuation minus the price he pays—which we have just said is equal to his valuation. Each buyer is actually indifferent about buying or not buying the good. (If this seems odd, then you can imagine that the monopolist charges a price slightly below individuals' valuations, so buyers obtain only a tiny amount of surplus. The conclusion is the same.) So where does their surplus go? It shows up as monopoly profits. By price discriminating, the monopolist captures all the gains from trade in this market. And where do these monopoly profits eventually go? They flow to the owners of the monopoly. The buyer surplus that is taken away by the price-discriminating monopolist is eventually returned to the household sector as dividends from the firm. However, it is not the same households who gain and lose surplus. Not everyone is a shareholder in the monopoly. Monopoly power causes a redistribution from the monopolists' customers to its shareholders.

A situation like the one we have described, where a monopolist knows the valuations of all her customers, may seem a bit farfetched. Still, in situations where a monopolist negotiates individual prices with buyers, she will do her best to guess their valuations. For example, think of a used car dealer. When a prospective customer arrives on the lot, the dealer will typically engage the customer in conversation in an attempt to learn where he lives, what kind of job he has, and so on. Such information helps the dealer guess the buyer's valuation.

Other Kinds of Price Discrimination

There are many kinds of price discrimination.

- **Explicit segmentation.** The monopolist may be able to divide the market into identifiable segments and charge different prices to different segments. For example, movie theaters often offer student and senior citizen discounts. As another example, businesses often charge different prices in different countries. The idea is that the monopolist can charge higher prices to segments that are less price sensitive.
- **Self-selection.** If a firm cannot explicitly identify different segments, it can set a menu of prices and allow customers to sort themselves into different groups. When Apple introduced the iPhone, it originally charged a high price, knowing that impatient customers would buy immediately. It then dropped the price to capture the more patient, more price-sensitive customers. Discount coupons are another example. Individuals with more time to spare will clip and redeem coupons—and these individuals will typically also be more sensitive to price.

From the perspective of a firm, the biggest danger of price discrimination is the possibility of **arbitrage**. If you are selling your product more cheaply to some than to others, you don't want someone to buy at the cheap price and then resell to your higher-value customers. This is a particular issue for pharmaceutical companies that charge different prices in different countries. If you need proof of this, look at the spam in your e-mail. The chances are very good that you have recently received an e-mail offering you pharmaceuticals at Canadian prices.

Competition Policy on Price Discrimination

The Robinson-Patman Act of 1936 directly forbids certain forms of price discrimination. The following is from the act, with our emphasis added: "It shall be unlawful for any person engaged in commerce…to *discriminate in price between different purchasers of commodities of like grade and quality*, where either or any of the purchases involved in such discrimination are in commerce… and *where the effect of such discrimination may be substantially to lessen competition or tend to create a monopoly in any line of commerce.*"Cornell University Law School, Legal Information Institute, "§ 13. Discrimination in price, services, or facilities," accessed March 14, 2011, http://www.law.cornell.edu/uscode/15/13.html.

The type of price discrimination we have discussed certainly involves discrimination among different purchasers. What is less clear is whether such actions would "lessen competition or tend to create a monopoly." In our earlier example, price discrimination actually makes the market more efficient, and in general, price discrimination can increase or decrease efficiency. There are no simple guidelines, and each case must be examined on an individual basis.

Price discrimination also occurs in **business-to-business** transactions between firms. In a recent price-discrimination case involving Volvo, a car dealer charged that other dealers had obtained deeper discounts (price concessions) that permitted them to be more competitive.See Cornell University Law School, Legal Information Institute, "Volvo Trucks North America, Inc. v. Reeder-Simco GMC, Inc. (04-905) 546 U.S. 164 (2006) 374 F.3d 701, reversed and remanded," accessed March 14, 2011, http://www.law.cornell.edu/supct/html/04-905.ZS.html. This seems to fit the requirements of the Robinson-Patman Act because Volvo trucks are a homogeneous good (Volvo trucks in this case), and the practice allows one dealer to undercut another.

A jury trial led to an award of $4.1 million to the injured company. Not surprisingly, Volvo appealed the decision, partly because, so it claimed, the other dealers were not in direct competition with the dealer filing the suit. Once again, you can see the critical significance of defining the market. The Supreme Court eventually decided the case in January 2006 in favor of Volvo.

Key Takeaways

- A monopoly occurs when there is a single seller, called the monopolist, in a market.
- A monopolist produces the quantity such that marginal revenue equals marginal cost. This is a lower level of output than the competitive market outcome.
- The government has the legal authority to break up monopolies and forbids price discrimination.

Exercises

1. Think of two goods or services you have bought recently. Were close substitutes available? Do you think the producer and the retailer of those products had a lot of market power?
2. Looking at Figure 15.2.4, why do buyers have any surplus?

This page titled 15.2: Market Power and Monopoly is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

15.3: Patents and Copyright 15.3: Patents and CopyrightEdit Update titlecancel

Learning Objectives

1. What is the role of the patent and copyright systems?
2. What factors determine how long patent protection should last?
3. What is the commitment problem associated with the patent system?

In the introduction to this chapter, we mentioned the breakup of Napster, a company that facilitated the sharing of music. Napster provided file-sharing software that allowed computer users to share music files over the Internet. A music file, like any other computer file, is simply information: a collection of bits and bytes. It costs nothing to make a copy of a music file. What Napster did, in other words, was to make it easier for music files to be distributed at marginal cost. You might think that the antitrust authorities would have been delighted. But the argument in this case was that there was an infringement of the rights of the music producers. The courts held in their favor: the law came down on the side of the monopolists. We now consider *why* governments sometimes actively support and promote monopolies.

To understand the Napster case, we begin by recognizing that creating, producing, and recording a new song is a very costly process. First of all, there is the time spent by the band in writing and arranging the song. Professional recordings also need the services of a producer and an expensive recording studio. A great deal of time and resources must also be expended to create an MP3 file of the song. Economists say that the **first-copy costs** are large. As the name suggests, these are the costs involved in creating the initial version of a good. They are a particular type of entry cost. Goods that involve a large amount of research and development or other intellectual input, such as books, computer software, and pharmaceutical products, have large first-copy costs.

As we just pointed out, though, once a song is produced, it can be *reproduced* at zero cost. The fixed costs of producing a song are very large, whereas the variable costs are zero. Perhaps you can now see the problem. If songs were sold in a competitive market, their price would be zero. Producers of the music would earn no revenues. Composers would earn no money. In this world, no one would have an incentive to produce music unless they were doing it purely for their own pleasure.

Similar tensions can be found in many other industries. Nearly anything that can be stored as a computer file has high first-copy costs accompanied by low variable costs. A newspaper article takes time to research and write but can be copied at zero marginal cost. Computer software can be very expensive to develop but—once created—can be copied at no additional cost. Pharmaceutical compounds are very expensive to develop: they first involve the work of highly trained research scientists in expensive laboratories and then require years of testing on animals and humans. Once a drug has been developed, however, it is often quite cheap to produce.

If the antitrust authorities forced newspapers, software developers, and pharmaceutical companies to sell at marginal cost, these firms would not earn enough revenues to justify their initial investment. Instead, such firms are permitted to sell at above marginal cost. More than that, the government actively bestows monopoly power. It does so through legal protections for inventions and created works, known as patents and copyrights.

The Decision Problem of an Innovating Firm

To understand how patents and copyrights work, we think about a firm contemplating an *innovation*—the introduction of a new product or a new means of production into a market. The firm's decision involves several stages, as shown in Figure 15.3.1:

1. Do we innovate or not? A decision to innovate is a decision to incur certain costs: the costs of research and development and the costs of entering a market.
2. How much should we produce during the years of patent protection?
3. What do we do once the patent protection ends?

Our ultimate goal is to evaluate the innovation decision at the first stage, but to do so, we must start at the end and work backward.

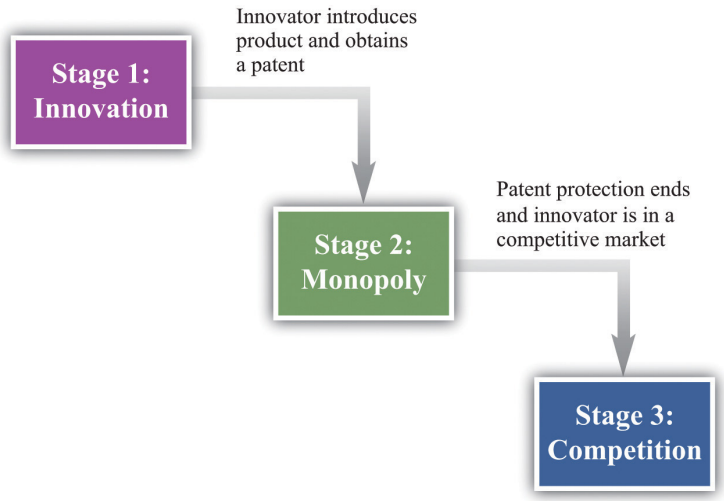


Figure 15.3.1: The Stages of Innovation

The Final Stage: Competition

Once a firm's patent expires, other firms can produce a similar or an identical product. The firm will then be operating in a competitive market and can no longer expect to gain any particular advantage from its innovation. When the patent for a pharmaceutical product expires, for example, other companies can step in and produce chemically identical copies of the product, known as *generics*.

In a competitive market, we expect the price of the product will decrease until it equals marginal cost. For this reason, the innovating firm cannot anticipate making very much profit at this stage. For simplicity, we can think of the firm making no profits. Although there may be some advantage in being the original producer of a product, any excess profits that remain after the patent expires are unlikely to be substantial. More precisely, the firm would earn no more than a "normal" level of profit—the same as it could earn in any other activity. Such normal profits would not provide any benefit to justify the initial innovation, so we can ignore them.

The Middle Stage: Patent Protection

If the innovating firm is going to make profits to justify the costs of developing its product, these profits must come in the middle stage when the firm has patent protection. During this period, the firm has monopoly power by virtue of the patent. We know how the firm behaves in this situation.

- It will produce a level of output such that marginal revenue equals marginal cost.
- It will set the price equal to the market's willingness to pay for this output level.

This is exactly what we saw earlier in Figure 15.3.4. The monopolist produces less output than is efficient and earns monopoly profits.

Notice that a firm's decision about how much to produce and about what price to set does *not* depend on the costs that it paid for researching and developing its product. After the firm gets to this second stage of its decision, those costs are in the past. They are **sunk costs**. They have no influence on the marginal cost of production and the price/output decision of the monopolist.

To calculate the total profit that a firm earns in this monopoly stage, we must do two things: (1) calculate the firm's profit in each year and (2) add these profits over the entire time that the firm has patent protection. The firm's profits in any given year are given by

profits = revenues − total cost = revenues − variable cost − fixed operating cost.

The area shown as "monopoly profit" in Figure 15.3.4 corresponds to revenues minus variable costs. (Businesspeople and accountants call this a firm's *profit contribution*.) In any given year, the monopoly will also typically incur some costs of operation in any given year that are constant, irrespective of how much output it produces. Examples include rent on a building and other long-term contracts. These are the firm's **fixed operating costs**, which also must be subtracted from the firm's revenues to calculate its profits.

Because the monopolist earns profit in each year of its patent protection, we add these profits together. We do so using the tool of **discounted present value**. This calculation takes into account that money earned in the future is less valuable than money today whenever the rate of interest is positive. Thus the proper measure of the profits at this stage is the *discounted present value of the sum of the profits made during the period of patent protection*. Factors that would increase the discounted value of a firm's profits include the following:

- Lower marginal costs
- More inelastic demand
- More years of patent protection
- Lower interest rates

> **Toolkit: Section 31.5**
>
> You can review discounted present value in the toolkit.

The First Stage: Innovation

We are finally in a position to evaluate whether or not a firm should innovate. The gains from innovation are measured by the present discounted value of the flow of profits. A firm must compare these gains to the costs of innovation to determine whether or not the innovation is worthwhile. These innovation costs are determined by the costs of the research and development (R&D) process together with any other costs of market entry. For example, a firm must pay scientists and engineers, fund research laboratories and R&D departments, and so on. Having done the hard work of analyzing stage two, the decision for stage one is straightforward. The firm should follow this rule: "innovate if the discounted present value of profits is greater than the costs of innovation." The firm should innovate as long as the monopoly profits it will earn in the second stage (appropriately discounted) are greater than the costs of innovation in the first stage.

Think again about a pharmaceutical firm. Such firms spend an enormous amount of money on the research and development phases of new pharmaceutical compounds. This occurs in stage one. Once the product goes to market, however, the costs of development are sunk and have no effect on the firm's profits during stage two. The same point applies to the production of a music CD. The costs of producing, marketing, and distributing a typical CD are estimated to be around $500,000. Of this, at least $100,000 represents the costs of production.

One thing that we have neglected in our discussion is that the payoff from research and development efforts is typically uncertain. Many promising pharmaceutical compounds turn out, on further testing, to be ineffective or have unacceptable side effects. A band recording a new song cannot know for sure if it will sell hundreds of copies, thousands of copies, or millions of copies. The decision about whether or not to innovate must be based on a firm's best estimates of the **expected value** of its profits.

An Example

Table 15.3.1 "Calculating the Discounted Present Value of Expected Profits" provides a numerical example of the innovation decision. The first year is the innovation stage: we suppose the cost of innovation is $150. The firm earns no revenues in that year and incurs no costs, so its first year profits are −$150. In the second year, the firm finds out if its innovation was successful. We suppose there is a 50 percent chance that it is successful, in which case the firm has monopoly power for the second and third years. It earns the revenues and incurs the costs listed in the third and fourth columns of the table. If the innovation is unsuccessful, it earns no profits. After the third year, the firm earns no profits. Suppose finally that the interest rate is 10 percent.

Table 15.3.1: Calculating the Discounted Present Value of Expected Profits

Year	Innovation Cost ($)	Total Revenues ($)	Total Operating Costs ($)	Profit If Successful ($)	Profit If Not Successful ($)	Expected Value of Profits ($)	Discounted Present Value of Expected Profits ($)
1	150	0	0	−150	−150	−150	−150

Year	Innovation Cost ($)	Total Revenues ($)	Total Operating Costs ($)	Profit If Successful ($)	Profit If Not Successful ($)	Expected Value of Profits ($)	Discounted Present Value of Expected Profits ($)
2	0	200	68	132	0	66	60
3	0	400	158	242	0	121	100

The profits in the second year if the innovation is successful are $132. The expected value of profits in that year is therefore given by $(0.5 \times \$132) + (0.5 \times 0) = \$66$. These must be discounted back one year using the 10 percent interest rate. In other words, we divide the second-year profits by 1.1. The discounted value of the second-year profits in the first year is therefore $60. The expected value of the third-year profits, by a similar calculation, is $121. These must be discounted back two years, all the way to the first year. To do so, we first divide by 1.1 to get the value of expected third-year profits in the second year $(\$121/1.1 = \$110)$. We then discount this back another year by dividing by 1.1 again. The discounted value of expected third-year profits in the first year is therefore $110/1.1 = \$100$.

We can now legitimately add together the numbers in the last column, and we find that the discounted present value of the firm's stream of profits is $160. This exceeds the cost of innovation ($150), so the firm should go ahead with the project. It expects to earn $10 from the project.

The Role of Patents

Imagine for a moment that we went through the same analysis in the previous section but without patent protection—that is, suppose that as soon as a firm innovates and introduces a new product into a market, it can immediately be copied and produced and marketed by any other firm. This means that the second stage is completely eliminated: the market goes straight from innovation to competition.

This has an apparent benefit. There is no longer a monopoly in the second stage. We know that monopoly causes inefficiency: the firm sets its price above marginal cost to earn monopoly profits. But there is a problem: the innovation will not occur. The firm innovates only if the discounted present value of profits exceeds the cost of innovation. If we eliminate stage two, then we eliminate the profit flows that justified the innovation in the first place. Without the profits, there will be no innovation. It is the patent protection that provides the *incentive* for innovation.

The trade-off should be clear. Patents

- provide incentives for innovation, but
- create monopoly power and hence distortions.

The gain from patent protection is that it provides the basis for the second stage of the innovation process. Without this protection, the gains from innovation would not exist. Looking back at Figure 15.3.4, this gain comes at the cost of an inefficiently low level of output and a consequent loss of gains from trade.

Lawmakers must trade off this cost and this gain. Under current US law (http://www.uspto.gov), patent protection for most products lasts for 20 years. The optimal length of patents remains an active area of research in economics and an active area of policy concern. It is very hard to maintain the right balance between incentives for innovation and eliminating market distortion.

Commitment

Perhaps the government could both encourage innovation and avoid the monopoly distortion. Suppose that, *after an innovation had been introduced*, the government removed the patent protection. As an example, many individuals in the world suffer from HIV/AIDS but cannot afford medication at current prices. This problem is especially severe in much of sub-Saharan Africa. As a matter of social policy, one would like to have the drug companies first develop HIV/AIDS treatments and then sell those products at marginal cost.

Currently, some pharmaceutical companies have developed treatments that are still protected by patents. Eliminating these patents seems like it might be good social policy, given that the lives of millions of people are at stake, but it would come with a significant cost. If the government ignores patent protection for a particular product today, innovators will suspect that the government can no longer be relied on to provide patent protection in the future, which would have a huge impact on innovative activity. Thus by failing to provide patent protection for one product, the government risks destroying its reputation for patent protection in general. This is an example of a *commitment problem*. Prior to stage one of the innovation process, the government promises patent protection to provide an incentive for innovation. After stage one is finished and the product is introduced in the market, the government's incentives change: it wants to remove the market distortions. Because the innovation stage is over, the government could potentially renege on its promise of patent protection.

Pharmaceutical companies like GlaxoSmithKline, which produce the antiretroviral drugs used for treating HIV, have in fact offered to make their products available more cheaply. However, pharmaceutical companies have strongly opposed suggestions that their patents not be honored. A better policy proposed by economists and others is a *patent buyout*. This would work as follows. The government would pay the pharmaceutical company a reasonable market price for its patent and then allow other companies to come into the market and produce generics. A related idea is that governments could offer to buy future patents if drug companies came up with treatments for particular maladies. For example, even though malaria is a major killer in the world, pharmaceutical companies have little interest in researching its treatment. The reason is one of harsh market economics: most sufferers from malaria are poor, so a malaria drug would not be very lucrative. If governments offered a substantial patent buyout for a malaria drug, however, firms might find it worth investing in this disease area after all.

International Dimension of Patents

If you have a US patent for a good that you have invented—for example, a new kind of printer—then you are protected *in the United States*. However, US patent law does not protect you if you are selling in other countries. A firm in another country could take your printer apart, analyze how it works (this is called reverse engineering), and then produce the good itself. If the firm tried to sell the good in the United States, you could take it to court, but if it is selling elsewhere in the world, you have very little protection.

Other countries (such as Japan) and regions (such as the European Union) have patent laws that are similar to those in the United States. (Indeed, one of the earliest biotech patents went to the famous biologist Louis Pasteur in 1873 for his method of producing yeast.) These laws are structured, as are those in the United States, to balance the gains from innovation against the costs of monopolization of the market. Thus producers who sell across the world will typically seek patent protection in many different countries.

Still, to the extent that the benefits of innovation flow to purchasers around the world, the innovation-monopoly trade-off is potentially altered. If we think about the stage two monopoly as "the price we pay for the benefits of innovation," then residents in countries with strong patent laws are paying in part for benefits that flow to individuals and households in other economies. In some countries—China is a leading example—patents and copyrights are not very well protected. This is good news if you want to buy some cheap DVDs in China, but if you are the owner of the rights to these pirated movies, you surely wish that copyright laws were more thoroughly enforced around the world.

Key Takeaways

- Patents and copyrights provide innovators with protection from competition so that there is a return to innovation.
- Although a patent system provides protection, it also creates market distortions by granting monopoly power. A patent system should be designed to balance the incentive to innovate against the losses from these distortions.
- After innovation has taken place, the government may be tempted to take away patent protection to avoid market distortions. This is the commitment problem faced by a government. Governments are aware that if they take away patent protection from firms that have already innovated, they will greatly damage the incentives for future innovation.

Exercises

1. Draw the diagram for a competitive market where marginal cost is zero. What does the supply curve look like? Who gets the surplus in the market? Use this diagram to explain the tension between innovation and competition.
2. The subsection on patent protection ended with a list of factors that would increase a firm's profits. Explain *why* each of these would cause the present discounted value of the firm's profit flow to increase.

This page titled 15.3: Patents and Copyright is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

15.4: Markets with a Small Number of Sellers 15.4: Markets with a Small Number of SellersEdit Update titlecancel

Learning Objectives

1. How do we predict the market outcome in a market with a few sellers?
2. How does the outcome depend on whether firms set prices or quantities?
3. What are the main tools of competition policy in markets with a few sellers?

So far we have looked at monopolies: markets with a single seller. But as we pointed out earlier, the extent to which a firm is a monopolist really comes down to how we choose to define the market. The market power of all firms is limited, to a greater or lesser degree, by the presence of firms selling competing products. In this section, we examine how the presence of competitors affects the distortions due to market power.

If there are enough competitors to give us perfect competition, then there is no distortion. But what about intermediate cases, where there are a small number of sellers? The market power of Microsoft Corporation is muted by the presence of competitors producing products that are substitutes for Microsoft software. If Microsoft were to triple the price of its Windows operating system, many buyers would switch to Macintosh computers or even start running the free operating system Linux. To the extent that substitute products are available to well-informed consumers, market power is decreased.

Once firms have to start worrying about the strategies of their competitors, decision making can become surprisingly complex. Each choice that a firm makes concerning what goods and services to produce, how to produce them, how to market them, how to price them, and so on, is now complicated by the fact that all of its competitors are making similar choices. This contrasts with a competitive market, where a single firm does not need to consider the behavior of its competitors at all; it only needs to know the market price for its output.

The task of a firm with market power is to choose the point on the demand curve that maximizes its profit. From the firm's perspective, it doesn't matter whether it (1) sets a price and then lets the quantity demanded come from the demand curve or (2) chooses a quantity and then lets the price be whatever the market will bear. Oddly enough, though, the interaction among firms is very different in a world where firms are setting prices compared to one where they choose their level of output. As we go through this section, we shall see why.

Market Outcomes When Firms Set Prices

We start by thinking about the case where firms set prices, beginning with two firms selling exactly the same product. An example is two gas stations at the same intersection, each of which has to choose what price to set for the gasoline that it sells. We first look at the choice of one of these firms and then study what happens when the two firms interact. We make a simplifying assumption, which is reasonable for gasoline retailers, that marginal cost is constant.

The Pricing Decision of a Single Firm

The first thing we need to know is the shape of the demand curve facing a firm. Let us look at the decision of one gas station (firm A). The owner of this gas station can look across the street and see the price set by its competitor (firm B) across the street. Suppose, for example, that firm B is selling gas at $2 per gallon. What does the demand curve for firm A's gas look like?

- If firm A charges a price greater than $2 per gallon, all its potential customers will go across the street to firm B. The quantity demanded will be zero.
- If firm A charges a price less than $2 per gallon, it can capture the entire market. The demand curve faced by firm A is the same as the market demand curve.
- If firm A sets a price of exactly $2 per gallon, then the two firms will divide up the market. The simplest assumption is to suppose that they will each get half the market.

We illustrate this demand curve in Figure 15.4.1.

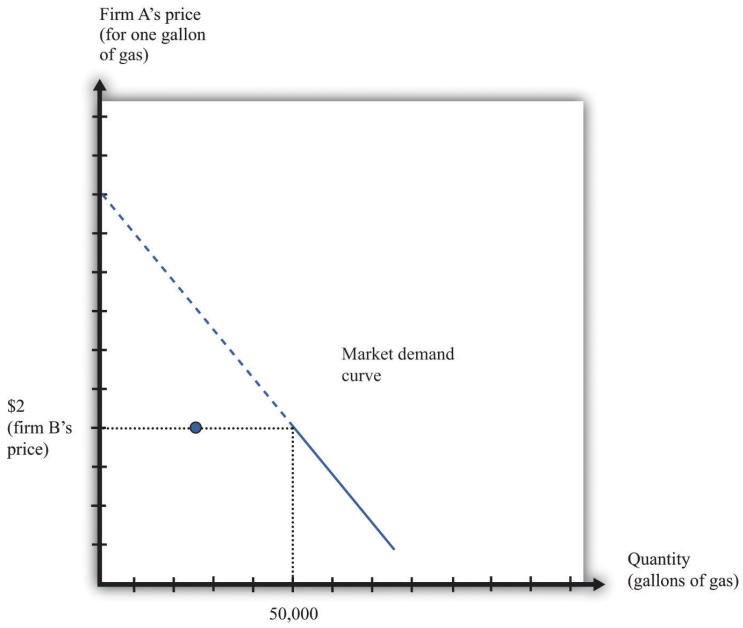


Figure 15.4.1: The Demand Curve Facing a Firm, Taking as Given the Price Set by a Competitor. This figure shows the monthly demand curve facing firm A in a market where two gas stations are setting their price per gallon.

If firm B sets a price of $2, what should firm A do? As long as $2 is greater than the marginal cost, firm A makes the most profit if it undercuts firm B a little bit. If it sets a price of $1.99, it can capture the entire market, whereas if it sets a price of $2 it gets only half the market. But exactly the same is true of firm B. For any given price that firm A sets, firm B would do better to undercut it by a penny. Competition provides a strong incentive for firms to cut their prices.

Nash Equilibrium

The discussion so far tells us how one firm will respond to the price of the other, but we don't yet know where the firms will end up. We do not yet know what the equilibrium will look like. Previously, we used the term *equilibrium* in the context of supply and demand, denoting the point where the supply and demand curves intersect. The idea of equilibrium goes beyond this, however: it denotes a situation of balance in which no one has any desire to change.

When we think about strategic situations with a small number of firms, we maintain the idea of an equilibrium where no one wants to change their decision. We use an idea of equilibrium invented by John Nash (a mathematician who won a Nobel Prize in Economics for this and other contributions). The concept of the **Nash equilibrium** also expresses a sense of balance, but it is applied to strategic situations rather than markets. The key feature of the Nash equilibrium is that no one has any desire to change what is being done.

Toolkit: Section 31.18

The Nash equilibrium is used to predict outcomes in strategic situations, often referred to as "games." In a game, a small number of players (such as firms) interact. Each player chooses an action, and each player receives a payoff (for example, profit). The payoff of a player depends on his chosen action *and the actions chosen by all the other players*. In the Nash equilibrium, two things are true:

1. Each player chooses the action that gives him or her the highest payoff, based on his or her predictions of the other players.
2. Each player's predictions of the actions of the other players are correct.

Nash Equilibrium When Firms Choose Prices (Bertrand Competition)

To see the Nash equilibrium in action, go back to our gas stations. So far, we have seen that—taking firm B's price as given—firm A will want to set a lower price. For example, if firm B sets a price of $2, firm A will set a price of $1.99. But this is not a Nash equilibrium because firm B would then like to do something different. If firm A sets a price of $1.99, firm B will want to set a price of $1.98, and so on. This process will stop only when the firms' prices equal marginal cost. Thus we can make an educated guess:

in the Nash equilibrium, each firm sets its price equal to the marginal cost of production. In this equilibrium, both firms earn no profits.

To see that this is indeed an equilibrium, we suppose firm B sets its price equal to marginal cost and then ask if firm A would like to change its price away from marginal cost. Were firm A to set its price greater than firm B, it would get no sales and no profits. This does not make firm A better off. If firm A were to set its price below marginal cost, then it would capture the entire market, but it would make a loss on each sale. It would be selling below the costs of production and thus make negative profit. This certainly does not make it better off. So firm A has no incentive to do anything different. Obviously, the same arguments apply to firm B. If both firms set the price at the marginal cost, neither can change its price and make higher profit.

The game played between two firms producing an identical product and setting prices is called **Bertrand competition**. The remarkable implication of Bertrand competition is that the predicted outcome (price) will equal marginal cost. With only two firms, in other words, we get the same outcome as with a competitive market. Even with two firms, the market can be very competitive, and there may be no need for antitrust authorities to intervene.

The Prisoners' Dilemma

Households certainly like Bertrand competition because they can purchase goods at marginal cost. The firms, however, clearly do not like the outcome. Both firms understand that they would both be better off if they could only charge more for a good. If they could meet together and collude, then they would want to behave like a monopoly and share the profit between them. But charging the monopoly price is not a Nash equilibrium. If one firm were to set the monopoly price, the other would have an incentive to undercut its rival by charging a slightly lower price.

Here is another way of looking at this problem. Let us again consider a situation where two firms are choosing their prices. To keep things simple, suppose now that the firms must choose between two prices: the (high) monopoly price and the (low) competitive price. We will no longer assume constant marginal cost, so the firms still earn some profit when the price equals marginal cost. There are thus four possibilities: (1) both firms set a high price, (2) both firms set a low price, (3) firm A sets a high price and firm B sets a low price, and (4) firm A sets a low price and firm B sets a high price.

Suppose that Figure 15.4.2 shows us the profits that firm A earns in each case. (There is a similar figure for firm B.) When both firms set high prices, firm A earns profits of $100 (think of this as monopoly profits of $200 that they share.) However, if firm B sets a high price, firm A is better off setting a low price. In this case, firm A gets profits of $120. So if firm B sets a high price, firm A does best by setting a low price. What about if firm B sets a low price? Then firm A gets nothing if it sets a high price and profits of $60 if it sets a low price. Firm A is clearly better off setting a low price in this case as well. No matter what firm B does, firm A should set a low price. The same is true for firm B, so the Nash equilibrium is for both to set a low price. If they could collude, they would both agree to set a high price and earn higher profits. But collusion is not a Nash equilibrium because both firms have an incentive to cheat. This is an example of a **prisoners' dilemma** game.

Firm A's profits

	Firm B sets high price	Firm B sets low price
Firm A sets high price	100	0
Firm A sets low price	120	60

Figure 15.4.2: The Payoffs (Profits) from Different Pricing Choices. Firm A and firm B are each choosing to sell at either a high price or a low price. No matter what firm B chooses to do, firm A is better off setting a low price.

Is there any way that the firms might be able to change the incentives so that they can collude? Each firm would like some means of punishing the other if it cheats. If they get to set their prices only once, there is no obvious punishment. But if these firms are

competing over a long period of time—as happens in the real world—then more possibilities open up.

Think again about our two gas stations. They might both agree to set a high price for their gas. Then they keep a careful eye on each other. As long as firm B keeps its prices high, firm A is content to do the same. But if firm B ever drops its prices, then firm A can punish it by retaliating. More specifically, suppose the firms adopt the following rules:

- We both set the monopoly price (cooperation).
- However, if one of us ever fails to set the monopoly price (defection), then the other firm will set the competitive price forever after.

Figure 15.4.3 shows an example of firm A's profits in this case. (This should now be understood as a discounted present value of profits because we are imagining the firms competing for months or years.)

Firm A's (Discounted Present Value of) Profits

	Firm B cooperates	Firm B defects
Firm A cooperates	500	180
Firm A defects	420	300

Figure 15.4.3: The Payoffs (Profits) from Cooperating and Defecting. Defecting means setting the low price every month. Cooperating means setting the high price as long as the other firm has set the high price in all previous months but switching to the low price every month if the other firm ever sets the low price.

Will the firms have an incentive to follow these rules? Put yourself in the shoes of firm A. If firm B is charging the monopoly price, then you can make a quick profit by undercutting firm B's price. This will generate a gain in that period. But there is a cost: in future periods, firm B will charge the competitive price, and your profit will be driven to $0. For the numbers in the table, it is better for firm A to cooperate if firm B also cooperates. Thus if two firms compete with each other over and over again, they may be able to sustain collusive high prices.

There is something else striking about Figure 15.4.3. It is a Nash equilibrium for both to cooperate, but it is still *also* a Nash equilibrium for both to defect. Look at the profits of firm A if firm B defects. Firm A earns $300 if it defects but only $180 if it cooperates. If you thought the other firm was going to renege on the agreement, then you would want to do the same thing. Economists say that this game has *multiple equilibria*. Expectations are critical: if each firm expects the other firm to cooperate, then they will indeed both cooperate; if each firm expects the other firm to defect, then they will indeed both defect. This is called a **coordination game**.

Toolkit: Section 31.18

The prisoners' dilemma game, the coordination game, and other games are discussed in more detail in the toolkit.

One final note: we are showing how firms can (and often do) sustain high prices even in the face of competitive pressures. We are not suggesting that this is what *you* should do if you are ever responsible for setting prices! Conspiring to set high prices is very often a violation of antitrust laws.

Price Competition with Imperfect Substitutes

Up to this point, we supposed that the two firms were producing an identical product. If we think of two firms producing goods that are close—but not perfect—substitutes, we still reach very similar conclusions. For example, suppose there are two pizza restaurants on the same street. If one restaurant undercuts the other's price, it would no longer expect to immediately capture the entire market, but it would still expect to gain a lot of business. In this situation, the arguments that we have just made still apply.

Each restaurant would have an incentive in the short run to undercut the other's price. If they compete repeatedly, however, they may be able to sustain high prices.

Market Outcomes When Firms Set Quantities

Look again at Figure 15.4.1. The demand curve in that figure is based on the idea that if firm A sets its price a little below firm B's price, firm A will capture the entire market. This presumes, though, that firm A can produce enough to supply the entire market. Suppose instead that firm A is unable to supply more than 25,000 gallons of gas per month. Now, if firm B has set a price of $2, then the best that firm A can do is to also set a price of $2. Firm A no longer sees any benefit from cutting its price because it cannot supply any more gas to the market. Similarly, think of the two pizza restaurants. If both restaurants are typically full most evenings, then neither would see a benefit from cutting its price. There is no point in trying to attract your competitors' customers if you cannot then supply them with the goods or services that they want.

This suggests another way in which firms can keep prices high. They can deliberately limit their capacity to change *their own incentives* about price-setting. In effect, this is another way of "changing the game" of Figure 15.4.2. To analyze this kind of behavior by firms, we suppose that they choose their level of output rather than the price they set.

The Capacity Decision of a Firm

We again consider a situation where two firms are competing in the same market. As before, the first step is to determine the demand curve facing an individual firm. Figure 15.4.4 shows our gas station example again. In contrast to our previous analysis, firm B is choosing how much to produce (that is, how much gas to sell) rather than what price to set. As firm B increases its output, the demand curve faced by firm A shifts to the left. If firm B produces nothing, firm A faces the entire market demand curve. If firm B produces (sells) 30,000 gallons, firm A's demand curve is shifted to the left by that amount. You can see that the demand curve faced by firm A has a familiar shape, unlike the odd demand curve in Figure 15.4.1.

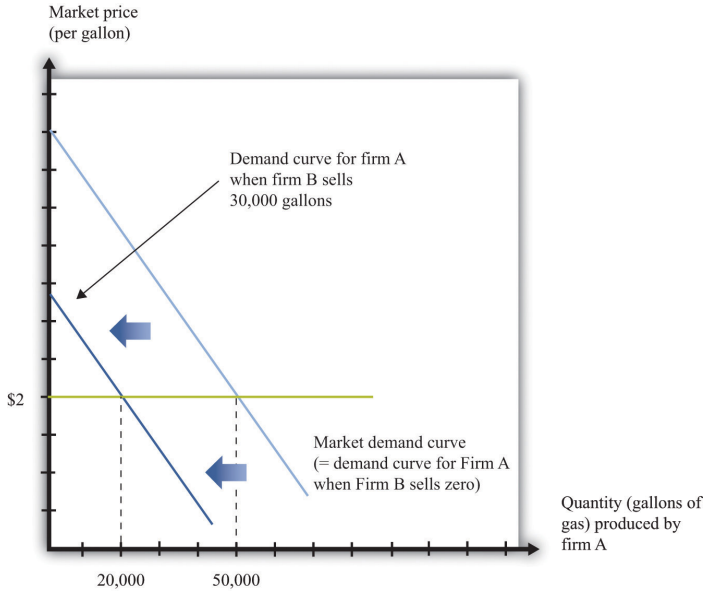


Figure 15.4.4: The Demand Curve Facing One Firm Shifts to the Left as the Other Firm Increases Its Output. As firm B produces more output, the demand curve faced by firm A shifts to the left.

When its demand curve shifts to the left, firm A's marginal revenue curve also shifts to the left. Figure 15.4.5 shows what happens. Note that the downward-sloping curves here are now marginal revenue curves, not demand curves. We omitted the demand curves to keep the diagram from being too cluttered.

Part (a) of Figure 15.4.5 shows two marginal revenue curves for firm A associated with different levels of output for firm B. An increase in firm B's output causes the marginal revenue curve facing firm A to shift to the left. How will firm A respond? As always, we know it will produce a level of output such that marginal revenue equals marginal cost. So as the marginal revenue curve shifts inward, firm A will produce less output. If firm B produces more output, firm A will produce less. This response of firm A to firm B is shown in part (b) of Figure 15.4.5. Here the output of firm B is on the horizontal axis, and the output of firm A is on the vertical axis. The downward sloping curve, sometimes called a **reaction curve**, shows us the output of firm A for every level of output of firm B.

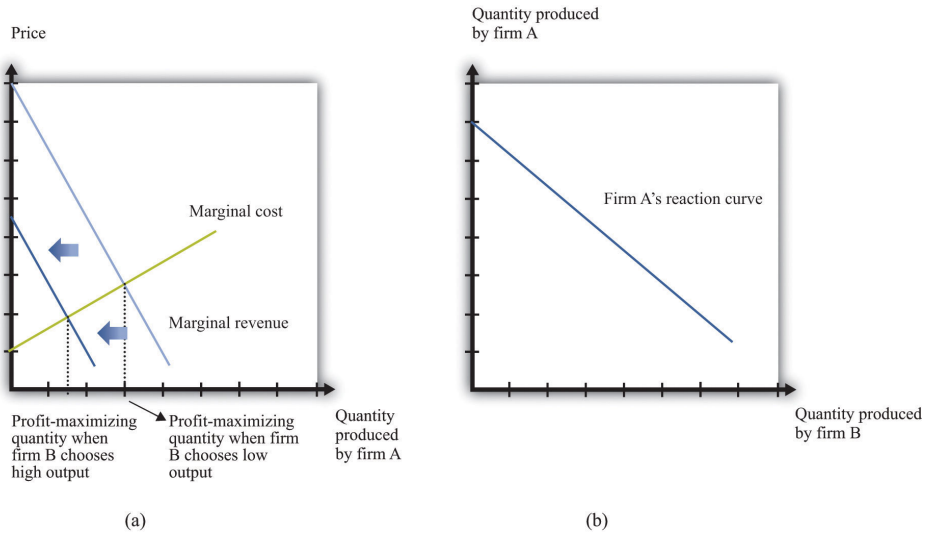


Figure 15.4.5: Firm A's Profit-Maximizing Choice of Output as Firm B Changes Its Level of Output. As firm B produces more output, firm A's marginal revenue curve shifts to the left (a), and firm A responds by producing less output (b).

Toolkit: Section 31.18

A reaction curve is used to help find the equilibrium in a strategic situation. It shows what happens to one player's best strategy when the other player's (or players') strategy changes.

Nash Equilibrium Revisited

We can now predict what will happen in this market. To simplify matters, we assume that the two firms are identical. This will make it easier to find a Nash equilibrium. In a Nash equilibrium, the following things are true.

- Firm A is choosing the level of output that maximizes its profits, which is based on its prediction of how much output firm B is producing.
- Firm A's prediction about firm B's level of output is correct.
- Firm B is choosing the level of output that maximizes its profits, which is based on its prediction of how much output firm A is producing.
- Firm B's prediction about firm A's level of output is correct.

If the two firms are identical, they will produce the same levels of output in the Nash equilibrium. Then, as shown in Figure 15.4.6, the equilibrium level of output corresponds to the intersection of the reaction curve and the 45-degree line. It is at this point, and only at this point, that all four conditions that we have listed hold. To understand this, put yourself in the position of firm A. You make a forecast about how much firm B will produce. Suppose you correctly forecast firm B's profit-maximizing quantity. Then you will respond with your own profit-maximizing quantity. This is the point labeled as the Nash equilibrium in the figure. But why should you predict that quantity for firm B? That quantity is in fact its profit-maximizing choice, given what you are doing. The beliefs that each firm has about the other's actions are consistent, and indeed they are *self-enforcing*.

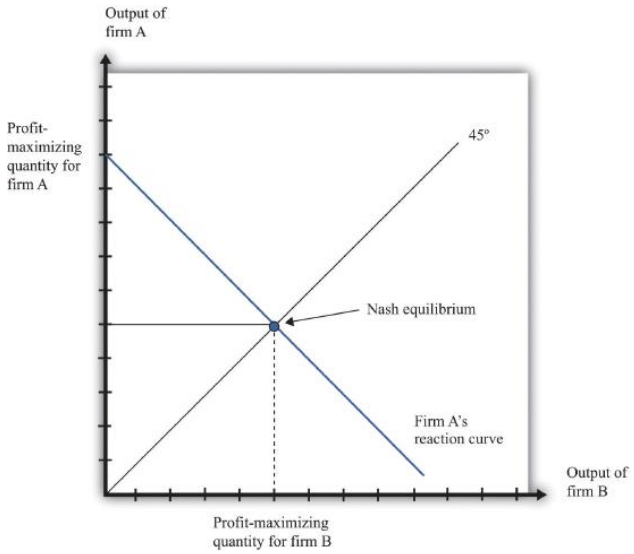


Figure 15.4.6: Nash Equilibrium for Quantity Game. The Nash equilibrium when both firms are identical occurs at the level of output where the reaction curve crosses the 45-degree line.

Determining Prices

Now that we know how firms choose capacity (quantity), how are prices determined? The answer is easy: prices come from the demand curve. If the two firms are producing identical products, the price comes from the market demand curve, given the *total* output of the two firms. This is similar to what we did in the monopoly case: given the output level, we turned to the demand curve to find the price. If the two firms are producing products that are not perfect substitutes, then the analysis is similar. However, there is not one demand curve in this case; there are two. Each firm faces a demand curve that depends on the output of the other firm, as shown in Figure 15.4.7.

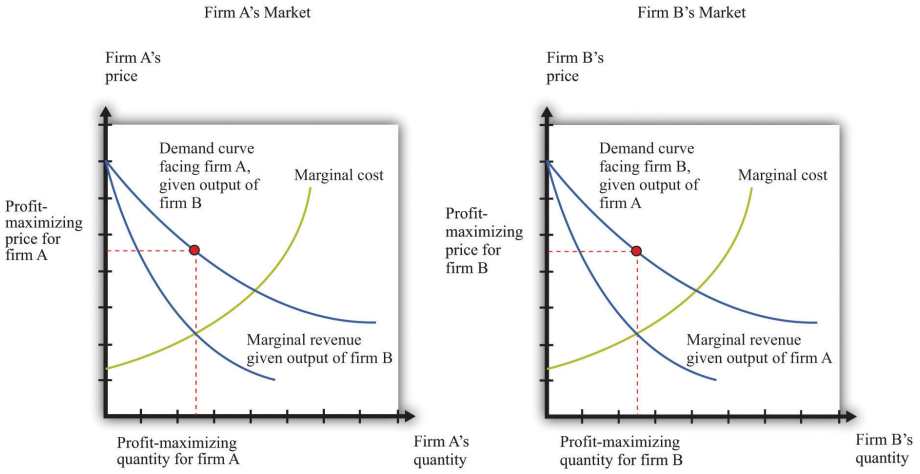


Figure 15.4.7: The Markets for Both Firms. Firm A correctly predicts firm B's profit-maximizing level of output, and firm B correctly predicts firm A's profit-maximizing level of output.

Inefficiency

When firms are choosing the price to set, it is possible for competition between two firms to drive prices all the way down to marginal cost and eliminate all monopoly inefficiency. This is what we see with Bertrand competition. When firms choose quantity rather than price, the effects of competition are much weaker. Look again at Figure 15.4.7. You can see that both firms are setting price in excess of marginal cost: there is still a distortion due to market power.

Competition between the firms does matter, however. Increases in firm B's output, for example, shift firm A's demand curve to the left. As a consequence, firm A ends up choosing a lower price than it would otherwise. Competition from firm B helps keep firm A's prices low. By similar reasoning, competition from firm A helps keep firm B's prices low.

In an ideal world (from their point of view), these firms would both limit their output further to get closer to the monopoly outcome. Exactly the same kind of strategies that we discussed earlier could come into play: firms that compete repeatedly over a long period of time might tacitly agree to reduce output further, punishing any defection by increasing output and cutting prices.

Competition Policy with a Small Number of Firms

In the United States, there are two aspects of policy when there are a small number of firms. First, a small number of firms in a market may be able to collude to set high prices. Antitrust laws can be used to punish such collusive behavior. Second, if there are a small number of firms, they may want to merge and become one large firm. Such mergers must be approved by the US government.

Collusive Behavior

The Sherman Antitrust Act is not only used against existing monopolies but is also more generally applied to groups of firms that act to jointly monopolize a market. If multiple firms act collusively to exert market power, they may fall foul of the act. Indeed, in the Standard Oil case, the key issue was the way the Standard Oil Trust brought multiple firms together into a single decision-making unit.

One form of collusive behavior occurs when firms come together to jointly decide on output levels and/or the price of the goods and services they sell. Occasionally, managers of firms are foolish enough to get caught on record making such proposals. In a famous incident a couple of decades ago, the CEO of American Airlines, Robert Crandall, proposed a 20 percent price increase to his counterpart at Braniff Airlines. Unfortunately for Crandall, the conversation was taped. (Interestingly, Crandall was not actually guilty of a violation of the Sherman Antitrust Act because no actual price-fixing took place.) More often, such agreements are likely to be *tacit*. One firm may try raising its prices, to see if others will follow. It can be very hard for the antitrust authorities to determine if price-fixing is actually occurring.

Other countries also have laws and agencies that seek to prevent collusion by firms. For example, in April 2008, the Office of Fair Trading in Britain charged two tobacco companies of colluding with supermarkets to set high prices for cigarettes. Specifically, the Office of Fair Trading said that the companies had set up arrangements "linking the retail price of a manufacturer's brand to the retail price of a competing brand of another manufacturer."Quoted in G. Wearden, "OFT Accuses Tobacco Firms and Retailers of Cigarette Price Fixing," *Guardian*, April 25, 2008, accessed January 31, 2011, http://www.guardian.co.uk/business/2008/apr/25/regulators.retail.

Another form of collusive behavior occurs under the heading of "bid rigging." Suppose only a few dairies provide milk to all schools in a region of the country. The schools set up auctions to decide which dairies will supply milk to different school districts. There is one auction for each district, and several dairies compete to provide milk to the different districts. The situation seems at first glance to be very competitive. We might expect the dairies to compete with each other in all the auctions, with this competition driving down milk prices for the schools. But imagine instead that the dairies agree ahead of time to divide up the districts. So for example, if there are three districts and three dairies, the dairies might agree that dairy 1 would win the auction in one district, dairy 2 in another, and dairy 3 in the third. They do this by each putting in very high bids (meaning they would charge a lot for a service), in the districts where they do *not* want to win the auction, allowing one dairy to win the auction and still charge a high price. Exactly such a scheme occurred in Texas, and an individual was charged with a felony.

Mergers and Acquisitions

Suppose you heard that Apple Computer and Microsoft Corporation were proposing a merger. You can perhaps imagine Steve Jobs and Bill Gates telling us how much we as consumers would benefit from this merger. They would say that Apple and Microsoft could combine the best features of their operating systems. The two companies could avoid costly duplication of research, so they would be able to provide goods more cheaply. They would no longer need to spend so much on advertising, again providing savings that could be passed onto the consumer, and so on. In a nutshell, Jobs and Gates might claim, the merger would bring new exciting products with lower costs of distribution and marketing.

You can be sure, however, that the Department of Justice, European antitrust authorities, and other similar bodies throughout the world would look on a proposed Microsoft-Apple merger with a highly skeptical eye. They would carry out their own studies of the costs and benefits of the merger. Even if the merger were to bring all the advertised benefits, it would also make the computer operating system market much less competitive. If there are originally two firms in a market and they then merge, they become a monopoly. The analysis in this chapter allows us to predict that output would decrease and prices would increase. Because a market

with only two firms may still be very competitive, the loss in buyer surplus going from a competitive market to a monopoly can be sizable.

Stemming from the Clayton Act, the Department of Justice and the Federal Trade Commission must approve mergers and acquisitions of larger companies. The guidelines used for those decisionsUS Department of Justice, "Horizontal Merger Guidelines," accessed March 14, 2011, http://www.usdoj.gov/atr/public/guidelines/hmg.htm. emphasize two general points: (1) the effect of the merger on efficiency and (2) the effect of the merger on market power. Efficiency here refers to *cost efficiency*—the extent to which a merged company will be able to reduce its costs of production. The Department of Justice puts it as follows in its merger guidelines: "Efficiencies generated through merger can enhance the merged firm's ability and incentive to compete, which may result in lower prices, improved quality, enhanced service, or new products."US Department of Justice and Federal Trade Commission, *Revised Section 4 Horizontal Merger Guidelines*, April 8, 1997, accessed March 1, 2011, http://www.ftc.gov/bc/docs/horizmer.shtm.

Efficiency may also relate to the *quality* of the good being produced. If a merger enables a better product to be produced at the same price as before, then the merger increases market efficiency. Put differently, if the merger increases the surplus of buyers, either because the product is improved or because the reduced costs of production lead to a lower price, then the merger has had a positive impact on efficiency. For example, one benefit from airline mergers might be the more efficient use of the information systems that handle travel reservations. Another might be more efficient use of airport landing rights.

With these guidelines in mind, how would the Department of Justice respond to a proposed merger of Microsoft and Apple? They would study the proposed merger with the goal of determining if the merger would create a more efficient market or if it would increase the market power of the sellers. Department of Justice economists would use frameworks like the ones we have presented in this chapter to help them predict the outcome of the merger.

The Department of Justice would almost certainly decide not to approve an Apple-Microsoft merger. However, it does not block most proposed mergers. There is debate among economists about whether the antitrust authorities are too lenient or too stringent. A recent study by economists Orley Ashenfelter and Daniel Hosken looked at five mergers that they suspected of being anticompetitive.Orley Ashenfelter and Daniel Hosken, "The Effect of Mergers on Consumer Prices: Evidence from Five Selected Case Studies" (National Bureau of Economic Research Working Paper 13859), February 2008. They compared prices of the goods produced by the merged firms with prices of goods that were not close substitutes yet had similar costs. They concluded that, in four of the five cases, prices increased between 3 percent and 7 percent, and "given the large amount of commerce in these industries, the implied transfer from consumers to manufacturers is substantial." The evidence from this study thus suggests that the antitrust authorities are too permissive in allowing mergers.

The European Union also watches mergers closely, likewise balancing cost-efficiency and competitiveness considerations. Here is a statement of its policy on mergers: "If the annual turnover of the combined businesses exceeds specified thresholds in terms of global and European sales, the proposed merger must be notified to the European Commission.… These rules apply to all mergers no matter where in the world the merging companies have their registered office, headquarters, activities or production facilities. This is so because even mergers between companies based outside the European Union may affect markets in the EU. "See European Commission, "Mergers: Overview," accessed March 14, 2011, ec.europa.eu/comm/competition/mergers/overview_en.html.

Key Takeaways

- The market outcome with a few sellers is the Nash equilibrium of the game they play. In the Nash equilibrium, none of the firms has an incentive to change what is being done.
- The market outcome depends on the strategy variable of the firms. If each firm is choosing the price of its output, then the outcome with many firms is the competitive outcome. If each firm is choosing the quantity of its output, then there is a distortion in the output market as price exceeds marginal cost.
- Governments act to regulate markets with a small number of sellers by making sure that firms do not make decisions jointly and evaluating the efficiency gains and market distortions from proposed mergers.

Exercises

1. Show that one firm setting its price at marginal cost and the other one setting a price above marginal cost is not the Nash equilibrium.
2. Explain how the interest rate will influence the choice of a firm to cooperate with another one in setting the monopoly price.

This page titled 15.4: Markets with a Small Number of Sellers is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

15.5: End-of-Chapter Material 15.5: End-of-Chapter MaterialEdit Update titlecancel

In Conclusion

There are very few examples of truly competitive markets. Most firms in the economy possess a certain amount of market power because their product, service, or location is distinctive. This means that most prices in the economy are in excess of marginal cost.

That said, the degree of market power of most firms is relatively small. Canon has some market power for its cameras because Canon cameras are not identical to Nikon, Olympus, or Sony cameras. But the presence of these other manufacturers severely limits Canon's ability to charge high prices. Your local Thai restaurant has some market power because its food is different from that of other restaurants in the neighborhood. Again, though, this does not mean it can charge very high prices because customers can easily eat at other restaurants instead.

Occasionally, however, firms are so large relative to their markets that they have substantial market power. This distorts prices and output in the economy. Firms with such market power can make a lot of money by restricting their output and charging very high prices. This is where the antitrust authorities come into play. Their task is to identify firms that are abusing their market power in this way. In effect, their job is to try to bring the economy closer to the economists' ideal world, where markets are competitive, there are no distortions, and all possible gains from trade are realized.

In some cases, though, governments have reasons to create and support market power through patents and copyrights. They do so because the benefits from innovation outweigh the distortions associated with monopoly. Policy in this area is highly contentious because the right balance between encouraging innovation and fostering competition is unclear. Economists and policymakers continue to struggle with this and are likely to do so for years to come.

Key Links

- US patent law: http://www.uspto.gov
- Japanese patent law: http://www.jpo.go.jp
- European Union patent law: www.epo.org/patents.html
- Department of Justice, Antitrust Division: http://www.usdoj.gov/atr
- Federal Trade Commission: http://www.ftc.gov/bc/index.shtml
- Department of Justice, Microsoft case: http://www.usdoj.gov/atr/cases/ms_index.htm, http://www.usdoj.gov/atr/cases/f0000/0046.htm
- Supreme Court decision in Standard Oil case: http://www.law.cornell.edu/supct/html/historics/USSC_CR_0221_0001_ZS.html
- Texas milk contracts: www.usdoj.gov/atr/public/press_releases/1992/211110.htm
- Clayton Act: http://www.usdoj.gov/atr/foia/divisionmanual/ch2.htm
- Guidelines for mergers and acquisitions: http://www.usdoj.gov/atr/public/guidelines/hmg.htm

exercises

1. Suppose you have two types of beverages: a cola and a beer. Are these products in the same market?
2. The table in Question 3 shows data for a monopolist who sells a good to four households, each of which buys at most one unit and each of which has a different valuation for the good. The monopolist can produce the good at a marginal cost of $4. The monopolist can discriminate perfectly in its pricing, charging each household its valuation. Fill in the missing elements in the table. How many units should the monopolist produce? How does your answer change if marginal cost is $6?
3. (Advanced) Looking again at the following table (with marginal cost equal to $4), calculate the marginal revenue. What is its relationship to price? Explain your findings.

Household	Quantity	Household Valuation	Price	Total Revenue	Marginal Cost	Total Cost	Profit
A	1	12	12	12	4	4	8
B	2	6	6	18	4	8	10

Household	Quantity	Household Valuation	Price	Total Revenue	Marginal Cost	Total Cost	Profit
C	3	4	4		4		
D	4	3	3		4		

TABLE 15.5.1: PRICE DISCRIMINATION BY A MONOPOLIST

1. Write an explanation of the monopoly pricing problem assuming the monopolist sets the price rather than chooses quantity. Why is the outcome the same either way?
2. Looking at the table in Question 3, if the interest rate increased to 15 percent, would the firm still have an incentive to innovate?
3. Explain why there is a greater incentive to innovate if the final stage of competition is with a small number of quantity-setting firms rather than price-setting firms.
4. Why might a merger lead to a price reduction? Why might a merger lead to a price increase?
5. Suppose that a firm (the incumbent) produces with constant marginal cost at $10 and has a constant (minus) elasticity of demand of 2. What is its profit-maximizing price? Now suppose that a new firm enters the market. The demand curve facing the incumbent firm shifts inward, but suppose that the elasticity of demand does not change. Should the incumbent firm change its price? What happens to the quantity that it sells? Draw a diagram to illustrate this market.
6. Imagine there is a motorcycle dealer in your neighborhood. You know both the price of the motorcycle set by the dealer and the amount of money the dealer paid for that motorcycle. It turns out that your valuation of the motorcycle is less than the posted price but greater than the cost of the motorcycle to the dealer. Are there gains to trade? Do you think you could convince the dealer to sell the motorcycle to you? If so, is there a deadweight loss? Why might the dealer be unwilling to sell the motorcycle to you?
7. Plane tickets are often sold at different prices to different people. Is this a form of price discrimination?
8. Writers of textbooks sometimes make their products available at a price of near zero. Does this mean they are altruists, or are they earning revenue some other way?
9. If interest rates increase, what needs to happen to patent lengths to maintain incentives for innovation?
10. In 15.4 Section "Market Outcomes When Firms Set Quantities", we looked at the situation when two firms chose quantity simultaneously. Describe the game and the outcome if one firm chose its quantity first and the other one followed. Would the outcome be the same as that discussed?
11. If you were a judge looking at a prospective merger between Coke and Pepsi, would you be more inclined to support the merger on efficiency grounds or argue against the merger as being anticompetitive?

Economics Detective

1. If a company invents, patents, and produces a product in the United States and sells the product in China, what type of protection does the company have in China?
2. If a US company operates in Europe, is it subject to European competition policy?
3. What legal authority does the European Union have over US firms?

Spreadsheet Exercise

1. (Advanced) Build a version of Table 15.3.1 "Calculating the Discounted Present Value of Expected Profits" starting with entries on demand and costs. To do so, use the examples in Chapter 7 to create demand, revenue, and then marginal revenue. Also use the examples there to create variable cost and marginal cost. Then find the profit-maximizing quantity and price. Using this information, calculate the profit for each year and then calculate the discounted present value of these profits.

This page titled 15.5: End-of-Chapter Material is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

16: A Healthy Economy

16.1: The Cost of Health Care
16.2: Supply and Demand in Health-Care Markets
16.3: Health Insurance
16.4: Government Policy
16.5: End-of-Chapter Material

This page titled 16: A Healthy Economy is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

16.1: The Cost of Health Care 16.1: The Cost of Health CareEdit Update titlecancel

What do you do when you are ill? You might first go to a drugstore, browse the shelves a bit, and find an over-the-counter medication that you think will make you feel better. Your choice of product could be influenced by many things, including past experience, the advice of friends, or perhaps an advertisement you saw on television.

If the trip to the drugstore doesn't solve the problem, a visit to a doctor usually comes next. The first doctor you visit is likely to be a general practitioner, or GP for short. Even if insurance is picking up some of the cost, a trip to the doctor is often not cheap. Nor is it usually fun: it may involve long waits and unpleasant tests. We go to the doctor not because we enjoy the experience in itself but because of a deeper demand—a desire to be healthy.

A trip to the doctor typically ends with a bill, a prescription, and perhaps a smile along with a "see you again soon." (That last bit, of course, is not quite what you want to hear.) Then you go to the pharmacy to fill the prescription. If you look at the piece of paper the doctor gave you, you might notice a couple of things. First, the doctor's handwriting is often illegible; penmanship is evidently not high on the list of topics taught at a medical school. Second, even if you can read what is written, it probably means nothing to you. The chances are that it probably names some medication you have never heard of—and even if you have heard of it, you probably have no idea what the medication does or how it works.

In other words, though you are the purchaser and the patient, your treatment is largely out of your hands. Health-care purchases do not directly reflect individual choices the way most other spending decisions do. You did not choose to be sick, and you do not choose your treatment either.

Occasionally, your GP might recommend that you visit another doctor, called a specialist. Your GP might try to explain the basis of this recommendation, but you probably lack the expertise and the knowledge to evaluate the decision. Once again, you must trust your doctor to make a good decision for you: the decision to visit the specialist is largely your doctor's rather than your own. You typically follow the doctor's advice for two reasons: (1) you trust the doctor to make decisions in your best interest, and (2) if you have medical insurance, you do not have to pay most of the costs.

We have described this as though you have no control at all over your own health care and treatment. This is an exaggeration. If you are somewhat informed or knowledgeable about what is wrong with you, then you can discuss different treatment options with your doctors. You can become at least somewhat informed by reading articles on the Internet. You can seek out second and third opinions if you do not trust your doctor's diagnosis. If you are having serious treatments, such as a surgical procedure, you will have to sign forms consenting to the treatment. There is a trend these days for people to become more involved and empowered about decisions involving their own health. Yet, unless you have medical training yourself, you will have to rely to some degree, and probably a very large degree, on the advice of your doctors.

If you are seriously ill, you may have to go to the hospital. There you have access to many more specialists as well as a lot of specialized equipment. Whatever autonomy you had about your treatment largely disappears once you enter the hospital. At this point—at least if you are living in the United States—you certainly should hope you have insurance coverage. Hospital costs can be astounding.

If you look back in history, health care was not always provided the way it is today. One difference is that doctors used to visit patients at home. They would traditionally arrive with a small black bag containing their basic tools. (This type of service is still provided in some communities and in some countries, but it is now rare in the United States.) For the most part, that was where your medical treatment ended.

In part, this reflected the state of medical knowledge at the time. It is hard to comprehend how much medical science has advanced in the last century. One hundred years ago, our knowledge about the workings of the human body was rudimentary. There were few treatments available. Antibiotics had not yet been discovered, which meant that the simplest injury—even a scratch—could become infected and be fatal. If you had appendicitis, it would very likely kill you. There were few means of diagnosis and no treatments for cancer.

Today, the story is very different. We visit specialists who have highly advanced (and expensive) training. We have access to advanced diagnostic tools, such as magnetic resonance imaging (MRI) scans, blood tests that identify markers for cancer, and genetic testing. We also have access to expensive treatments, such as kidney dialysis and radiation therapy. Perhaps most strikingly, we have access to a range of pharmaceutical products that have been developed—sometimes at great expense—by scientists and researchers. These products can treat medical conditions from asthma to apnea to acne.

With all these visits to doctors and all these medications, we spend a great deal on health care. Spending as a fraction of gross domestic product (GDP; a measure of the total output of the economy) has been increasing since 1960 (Figure 16.1.1). Figure 16.1.2 shows total spending on health services per person around the world.World Health Organization, “Total Expenditure on Health per Capita, 2007 (in US$),” accessed March 1, 2011, http://www.who.int/nha/use/the_pc_2007.png. The shaded areas indicate the level of spending on health-care services. The United States spends the most on health care per person, with Norway and Switzerland also being high-spending countries. Other rich countries, such as Japan, Australia, New Zealand, South Korea, and countries in Western Europe, likewise spend relatively large amounts on health care. The poorer countries in the world, not surprisingly, spend much less per person on health care. Across countries in the world, as within a country, health-care purchases are related to income.

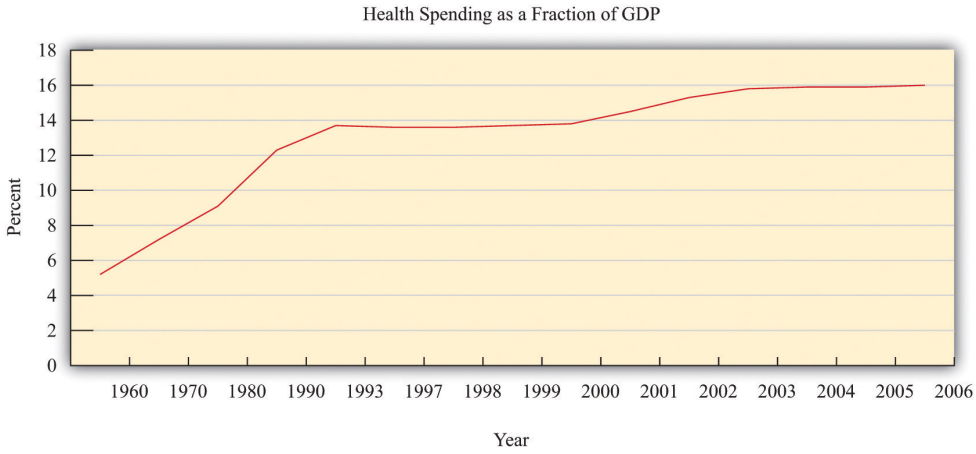


Figure 16.1.1: US Health-Care Expenses as a Percentage of GDP

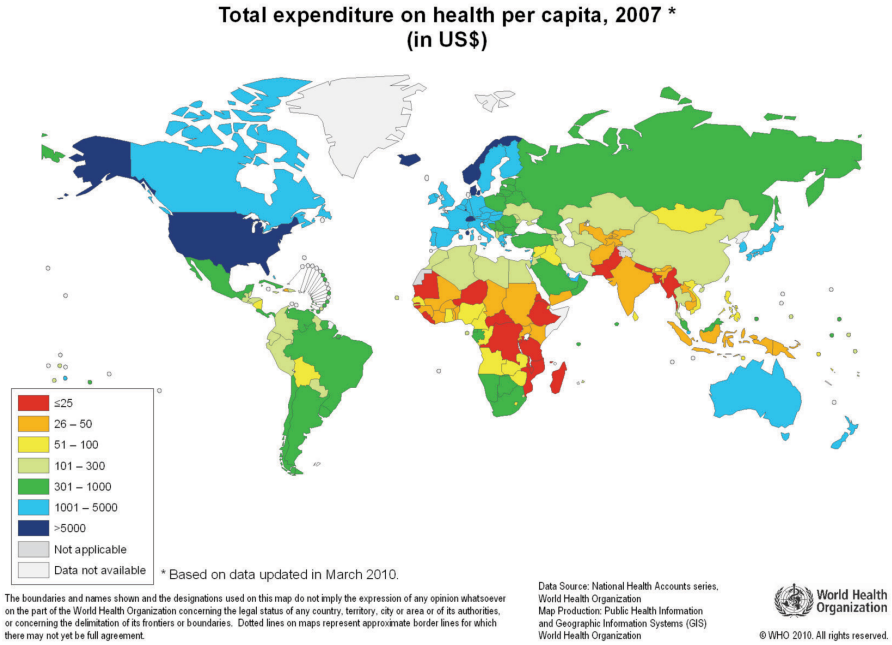


Figure 16.1.2: Global Health-Care Spending

One reason that we spend so much on health care in the United States is that high-quality care, such as is available in rich countries, is at least in part a luxury good—that is, something that we spend relatively more on as our income increases. Yet even across relatively affluent countries, health care takes very different forms.A comparison of programs is provided by Ed Cooper and Liz Taylor, “Comparing Health Care Systems: What Makes Sense for the US?” *In Context*, accessed March 14, 2011, www.context.org/ICLIB/IC39/CoopTalr.htm. Compare, for example, the United States and Canada. Canada has a system in which the government pays for health care. The program is financed by the payment of taxes to the government. Doctors’ fees are set by

the government, which limits competition within the health industry. Furthermore, other developed countries spend much less on health care than the United States but have health outcomes that are as good or even better.

Differences in both the quality and cost of health care mean that, perhaps surprisingly, health care is traded across national boundaries. In some cases, people travel across the globe to obtain care in other countries. Sometimes, people travel to obtain treatments that are unavailable in their home countries. For example, US residents sometimes travel to other countries to obtain stem-cell treatments that are banned in the United States. Or people may seek health care in other countries simply because it is cheaper: people from around the world travel to Thailand, for example, to obtain cheap and reliable dentistry services. There are even tour operators that arrange such "health tourism" trips. National Public Radio had a March 18, 2008, story of a husband and wife going to China for get stem-cell treatment for their seven-month-old daughter (http://www.npr.org/templates/story/story.php?storyId=88123868). There is also a company that organizes trips to Canada (http://www.findprivateclinics.ca/resources/general/medical-tourism.php). Given all these differences in health care costs around the world, we address the following question in this chapter:

What determines the cost of health care?

Road Map

Because we want to talk about the price of health care, supply and demand is a natural starting point. As we use this framework, though, it will rapidly become clear that there are many things that are unique about the market for health care. One indication of this is that there is a whole subfield of economics called "health economics." There is no subfield called "chocolate bar economics," "tax advice economics," or "lightbulb economics." Evidently, there is something different about health care. Another indication is the fact that governments around the world pay an enormous amount of attention to this market. Governments intervene extensively in this market through taxes, through subsidies, and sometimes by being the direct provider of health-care services.

We first study the demand for health care by households. Then we look at the supply of health care, after which we turn to the determination of prices. As we proceed, we will see that health care includes all sorts of different products and services. We will also see that there are many reasons why it is difficult to analyze health care with a simple supply-and-demand framework.

One key reason why health care is such a complicated topic has to do with the fact that, frequently, we do not pay for health care ourselves. Rather, we (or our employers) purchase health insurance, and then the insurance company pays the health-care providers. We therefore discuss health insurance in some detail. The chapter ends with a discussion of the government's role in the health sector, in which we talk about market failures and a variety of proposed government solutions.

This page titled 16.1: The Cost of Health Care is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

16.2: Supply and Demand in Health-Care Markets 16.2: Supply and Demand in Health-Care MarketsEdit Update titlecancel

Learning Objectives

1. What factors determine the price and quantity of health care?
2. In what sense is spending on health an investment?
3. What factors determine the demand for health-care services?
4. What is the production function for health?

Suppose we want to explain why health care is more expensive in the United States than in Europe. Then supply and demand seems like a natural starting point. If we imagine a market for health care drawn in the usual way, with the quantity demanded on the horizontal axis and the price on the vertical axis, then the question becomes, “Why is the price of health care higher in the United States than in Europe?”

Supply and demand offers two possible answers (Figure 16.2.1). The prices can be high because demand is high. For example, if the demand curve is further to the right in the United States compared to Europe (part [a] of Figure 16.2.1), this implies—all else being equal—higher prices in the United States. The other reason for high prices is because supply is limited. If the supply curve in the United States lies further to the left than the supply curve in Europe (part [b] of Figure 16.2.1), then this also would imply—all else being equal—higher prices for health care in the United States. Neither argument seems that compelling, which naturally leads us to wonder if the supply-and-demand framework is really the best framework for analyzing health care. In fact, there are good reasons to think that the supply-and-demand framework is not the best approach to this market.

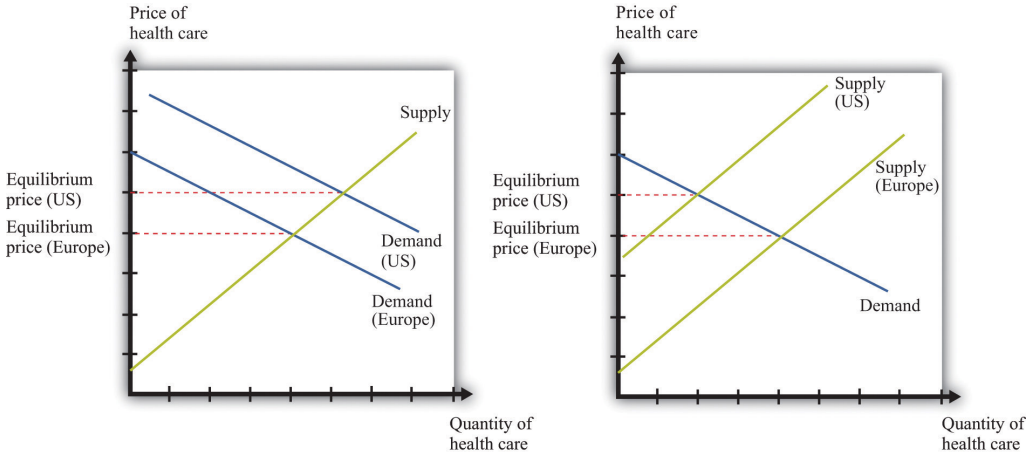


Figure 16.2.1: Two Explanations for Why Health Care in the United States Is More Expensive Than in Europe. Supply and demand offers two possible explanations of high health-care costs in the United States: demand in the United States is high (a), or supply in the United States is limited (b). Neither is a very compelling explanation.

Let us think about the demand side first. Our standard approach to demand is based on the idea that each individual will consume a good or a service up to the point where the **marginal valuation** from one more unit equals the price of that additional unit. Chapter 4 explains this idea in more detail. Unfortunately, the health-care consumer often has very little idea of the value—let alone the marginal valuation—of the particular treatment being received. The consumer is very often not paying the full price for that treatment because the cost is frequently covered, at least in part, by insurance. Together, these mean that our traditional approach to demand does not work very well for health-care services.

The supply side is also problematic. First of all, some health-care suppliers have significant **market power**. This does not mean that we can get no insights from supply-and-demand reasoning. But it is trickier to compare the price of health care across countries because we have to consider differences in market power as well. A bigger problem is that some health-care suppliers, such as hospitals, are either government-controlled or not-for-profit institutions. The standard economic approach presumes that firms seek to make as much profit as possible, but government or not-for-profit hospitals may not have profit maximization as their goal.

In addition, health-care prices are not necessarily determined by supply and demand. Again, the government has a significant influence on prices: for example, the governments in some countries set prices for pharmaceutical products. Even if they are not set

by the government, prices may be determined by bargaining between, say, hospitals and drug companies rather than by supply and demand. Furthermore, if people need health-care services, then their demand is likely to be very inelastic (the quantity demanded does not respond much to price changes). Inelastic demand is not, in and of itself, a problem for a competitive market. It just means that the equilibrium price could be very high. But if we couple inelastic demand with consumers who lack information and add in some market power by suppliers, then matters become more complicated. Perhaps you already have a sense of why: we have a large group of consumers with very inelastic demand who are relatively uninformed. This sounds like a gold mine for the supplier.

We have so far ignored the issue of what exactly is being traded in this market. “Health-care services” can mean many very different things:

- Labor time of various trained professionals, such as GPs, specialists, nurses, medical technicians, pharmacists, and many others
- Procedures and testing, such as magnetic resonance imaging (MRI) scans and laboratory analyses of blood samples
- Hospital and nursing care services
- Emergency services such as ambulances
- Pharmaceutical products (which itself covers a huge range, from bandages to chemotherapy drugs)

You can probably think of other components as well. So it is more than a little misleading to treat health care as something homogeneous that is bought and sold in a single market.

We hope that by now we have completely muddled your view of the health-care market. Our main point is that the simple framework of supply and demand is not sufficient for understanding health care. There are too many different markets, each with its own peculiarities and unusual features. And those features mean that there are several reasons why we might expect inefficiency. One, as we have already noted, is the presence of market power. Another is the various information problems we have mentioned. A third is that some aspects of health care have the characteristics of a **public good**.

The Demand for Health Care

Now let us dig a little deeper into the demand side of health care.

Response to Price

The law of demand applies to health care as in other markets: as the price of health care increases, you demand less of it. But we must be careful. What matters is the price of health care to *you*. If you have health insurance, this price may be much lower than the actual cost of providing you with care. Under most health-insurance contracts, the *marginal private cost* of care to a household is less than the **marginal social cost** of providing that care. The household has an incentive to purchase a lot of health-care services because its purchases are, in effect, being subsidized by insurance companies. We take up the topic of health insurance later in this chapter.

> **Toolkit: Section 31.19**
>
> You can review the distinction between marginal private cost and marginal social cost in the toolkit.

Another key characteristic of health care is that demand is relatively inelastic. If you are sick and require care, you will purchase health-care services at almost any price. Of course, your ability to purchase health care is ultimately limited by your income, but you are likely to trade off spending on many other products to purchase the medical care you need. This is why we often read stories about people without insurance being bankrupted by medical expenses.

Health as Investment

Everyone prefers being healthy to being sick. The demand for health care is in part an expression of this preference. One thing that makes health care different from most other goods and services, though, is that it is simultaneously an investment. Money you spend on being healthy today will also benefit you in the future. There are several different ways in which spending on health care represents an investment.

Mortality. One clear impact of our health-care choices can be seen in terms of mortality rates. Mortality rates measure how likely we are to die at different ages. In 2004, the mortality rate in the United States for people ages 15–24 was about 80 out of 100,000, or 0.08 percent. In contrast, the mortality rate for those over the age of 85 was 13,823 out of 100,000, or 13.8 percent.Mortality rates can be found at the National Center for Health Statistics, National Vital Statistics System, “Mortality Tables,” accessed March

14, 2011, http://www.cdc.gov/nchs/nvss/mortality_tables.htm. In other words, the typical young person has about a 1 in 10,000 chance of dying in a given year, whereas the typical old person has more than a 1 in 10 chance of dying.

It is not surprising that the mortality rate increases with age—that is, that young people have a lower probability of dying than older people. (Infants are an exception: a 6-month-old child is more likely to die than an 18-month-old child because very young children are particularly susceptible to certain diseases.) But these average mortality rates disguise a lot of variation, much of which is under our control. There are many behaviors that have predictable effects on our likelihood of dying. Smokers have a higher probability of dying than nonsmokers. Those who are obese have a higher probability of dying than those who are not. Diet, exercise, and risky behaviors (which includes everything from unprotected sex to skydiving) affect mortality rates as well.

Cigarette smoking is linked to lung cancer and thus to mortality. If you compare two similar individuals of the same age, one who is a smoker and the other a nonsmoker, then the mortality rate is significantly higher for the smoker. This does not mean that the smoker will necessarily die before the nonsmoker. It means that all else being the same, smoking increases the **probability** of death. Refraining from smoking is a type of investment in your future.

Our diet also affects our probability of becoming ill and of dying. As with cigarettes, there are often trade-offs between eating and drinking things we enjoy and the effects of such consumption on our long-term health. Making these types of choices is an economic decision. Each of us makes different choices because we value the taste of particular foods differently, and we value our overall health differently as well. If a thirty-year-old discovers he has elevated cholesterol levels that pose a long-term risk of heart disease, he may decide to adjust his diet, perhaps consuming less red meat. If an eighty-year-old learns the same news, he may not think the long-term benefit is worth giving up his steaks for.

Productivity. Being healthy also means that you can work and earn wages. One of the costs of poor health is lost days at work. This is a cost not only to the individual but also to society as a whole: the economy's population is producing less output. If you are in poor health, then you risk losing wages for the days when you cannot come to work. Many employers provide insurance for these lost wages through the provision of sick days: if you are sick, you are not expected to work but you will still be compensated up to a contracted number of days per year. In addition, there is disability insurance as part of the social security system in the United States.The government program is summarized at http://www.ssa.gov/disability. Private employers sometimes also offer disability insurance as part of their compensation packages, and you can also purchase insurance directly from an insurance company.

Information Problems

Health care is an example of a good for which the typical individual is unable to determine the quality of what is being purchased. You can think of other examples, such as legal services and used cars. In such situations, how can we make good decisions? Generally we do so by relying on the advice of experts. In the case of health, these are the doctors, dentists, and other health professionals who are trained to analyze our health situation and make suggestions to us. We listen, try to understand, and, using their advice, make an informed choice.

Suppose you get a phone call from someone telling you they know of a stock, trading on Wall Street, that will double in price the following day. You might be very skeptical, suspecting that they have other reasons for wanting you to buy. Compare this to a conversation with a medical expert. Generally you are going to believe that the expert is acting in your best interests. Although you might get a second (or third) opinion, you do so because health problems are complex and the first expert may have missed something, not because you are afraid the doctor is misleading you in order to profit from your visit. But why do we trust medical experts so much more than the provider of stock tips? We generally do so because we trust that their incentives are aligned with our goals; that is, we hope that they are motivated to act in our best interests.

The Supply of Health Care

We now turn to the supply side of health care. Economists often talk of output being produced using a **production function** that uses labor, capital, and intermediate inputs. What is the production function of a hospital?

- The labor in a hospital includes doctors, surgeons, orderlies, technicians, nurses, administrative staff, janitors, and many others.
- The hospital buildings are part of the hospital's capital stock. In addition, hospitals contain an immense quantity of other capital goods, such as hospital beds and diagnostic tools—everything from stethoscopes to x-ray machines.
- Intermediate inputs in a hospital include dressings for wounds, and pharmaceutical products, such as anesthetics used for operations.

Other sectors of the health-care industry likewise employ labor, capital, and intermediate inputs.

> **Toolkit: Section 31.17**
>
> You can review the meaning and definition of a production function in the toolkit.

Doctors

If you look at the wall in your doctor's office, you will typically see a large number of framed degrees and other qualifications. To become a doctor, you must first succeed as an undergraduate and then go through multiple years in medical school. After this comes an internship and then you finally graduate and can practice on your own. In most countries, you must have a license to practice medicine. This makes sense: you would not want anyone to advertise as a doctor regardless of their skill level. Most of us would be unable to tell whether a particular individual was a qualified professional or a quack. When buyers cannot easily evaluate the quality of the good or the service they are purchasing, it is useful to have external validations of quality.

Licensing provides more than a guarantee of quality, however. It also limits entry into the profession. Suppose you learned that a small group of lobbyists in your hometown wanted gas station owners to be licensed in the same way as physicians. You would quite rightly suspect that their goal was not to guarantee high-quality gasoline. More likely, they would be trying to limit the number of gas stations to increase their market power. Your suspicions would not be allayed if these lobbyists argued that gas was potentially a very harmful commodity, so by licensing the sellers of gas, they were protecting the community. In the case of doctors, the underlying reason for licensing is not so nefarious. But it still creates a **barrier to entry** that limits competition and increases market power, just as it would with gas stations.

Doctors differ from gas station owners in many other ways. Typically, we suppose that gas stations and other firms in an economy have profit maximization as a goal. It is this presumption that allows us to develop our theory of supply. Doctors not only think about profits but also take an oath of office, called the Hippocratic Oath, which is as follows:

I swear by Apollo, the healer, Asclepius, Hygieia, and Panacea, and I take to witness all the gods, all the goddesses, to keep according to my ability and my judgment, the following Oath and agreement:

…

I will prescribe regimens for the good of my patients according to my ability and my judgment and never do harm to anyone…

I will not give a lethal drug to anyone if I am asked,…*Wikipedia*, s.v. "Hippocratic_Oath," accessed March 14, 2011, http://en.Wikipedia.org/wiki/Hippocratic_Oath.

This oath is administered to nearly everyone obtaining a medical degree.

Other Health-Care Workers

In addition to doctors and specialists, there are many other kinds of workers in the health care industry, including nurses, dental hygienists, administrative staff, technicians, staff in care facilities such as hospices and nursing homes, and many others. The health-care industry employs almost 10 percent of all civilian workers in the United States.

Table 16.2.1 "Employment in Health Services (in Millions)" shows the breakdown of employment by health-service site for three years: 2000, 2003, and 2006. We list some (not all) of the types of health sites. From the table, we see that the number of workers in this industry has increased from 12.2 million in 2000 to about 14.4 million in 2006. This increase is typical of many service industries and contrasts with manufacturing, where the number of workers employed is declining.

Table 16.2.1: Employment in Health Services (in Millions)

Type of Health Service Site	2000	2003	2006
Physicians (offices and clinics)	1.4	1.7	1.8
Hospitals	5.2	5.6	5.7
Nursing care facilities	1.6	1.9	1.8
Total	12.2	13.6	14.4

Source: National Center for Health Statistics, *Health, United States, 2007 with Chartbook on Trends in the Health of Americans* (Hyattsville, MD: US Department of Health and Human Services, 2007), table 105, accessed February 1, 2011, http://www.cdc.gov/nchs/data/hus/hus07.pdf.

Hospitals are the most important type of employment site for health-care workers. In 2006, 40 percent of health-care workers were employed in hospitals. About three-fourths of the workers in the health-care sector are women. Women are particularly prevalent in nursing care facilities: of the 1.6 million workers in nursing care facilities in 2000, about 1.4 million were women.

There is a wide variety of occupations within health care: managers, professionals (doctors, dentists, pharmacists, etc.), service occupations (assistants, cooks, cleaners, etc.), and office workers. Professional groups account for about 44 percent of all workers, while about 32 percent of the jobs are in service occupations. There will typically be considerable variation of wages within a sector because of the different occupations of workers in that sector. For example, individuals working in diagnostic laboratories earn, on average, close to twice the wage of workers in nursing homes.

Health-Care Capital

When we look at an industry such as health care, one way of describing it is by counting the number of doctors' offices, clinics, and so on. There are many different kinds of establishments that provide health services. Hospitals are only one example; others include doctors' offices, clinics, nursing homes, and so on. According to the Bureau of Labor Statistics (www.bls.gov/oco/cg/cgs035.htm), in 2008, there were 595,800 establishments in the health-care sector in the United States. Of these, doctors' offices are 36 percent. Hospitals are only 1 percent.US Department of Labor, Bureau of Labor Statistics, "Career Guide to Industries, 2010–11 Edition: Healthcare," accessed March 14, 2011, www.bls.gov/oco/cg/cgs035.htm. Another way to describe the industry is by detailing the number of workers employed in different activities, as in Table 16.2.1 "Employment in Health Services (in Millions)".

Taken together, these statistics paint an interesting picture. Hospitals are a small fraction of the total health-care establishments but employ 35 percent of the workers. This tells us that there are relatively few hospitals (compared to doctors' offices), but they are big. About 70 percent of hospitals employ more than 1,000 workers. Such a pattern is not peculiar to the health-care sector. In US manufacturing, the majority of establishments are small, and a few large establishments employ the majority of workers.This is described in Steven J. Davis, John C. Haltiwanger, and Scott Schuh, *Job Creation and Destruction* (Boston, MA: MIT Press, 1998). We discuss this phenomenon in more detail in Chapter 9.

The ownership of hospitals is also complicated. Some are private, while others are public, meaning that federal, state, or even county governments run them. In addition, not all private hospitals are in business for profit; some are classified as not-for-profit institutions. Table 16.2.2 "Hospital Activity, 2005" provides a breakdown of hospitals by type. From this table, you can see that most admissions are in not-for-profit hospitals that are not federally run. Their goal is largely to provide a public service. These hospitals have a length of patient stay averaging about 5.5 days.

Table 16.2.2: Hospital Activity, 2005

Type of Hospital	Admissions (Thousands)	Average Length of Stay (Days)
All	37,006	6.5
Federal	952	11.6
Nonfederal, not-for-profit	25,881	5.5
Nonfederal, for-profit	4,618	5.3

Source: National Center for Health Statistics, *Health, United States, 2007 with Chartbook on Trends in the Health of Americans* (Hyattsville, MD: US Department of Health and Human Services, 2007), table 105, accessed February 1, 2011, http://www.cdc.gov/nchs/data/hus/hus07.pdf.

There are other capital goods that enter the production function for the health sector. For example, pharmaceutical production facilities are part of this capital stock. So too is the capital stock of companies that produce the machines, such as MRIs, used in doctors' offices and hospitals.

Technological Progress

Technological advances in health care are truly staggering. Technological progress in this sector, as in other sectors, comprises both product and process innovations. By product innovations, we mean increases in the types of goods and services available to households and doctors. A leading example is the vast array of drugs now available on the market, which is the outgrowth both of research and development at pharmaceutical companies and of publically funded research. Another example is the advanced machinery used in modern health-care facilities. A modern dentist's office is filled with high-speed drills, x-ray machines, and other pieces of technology that would have been unthinkable in your grandparents' day. MRI machines are another example: these are a significant advance over previous imaging techniques such as X-rays, but they are expensive—a new MRI machine will typically cost a hospital more than $1 million.An article describes this technique at NOVA, "The Picture Becomes Clear for Magnetic Resonance Imaging," accessed March 15, 2011, www.science.org.au/nova/062/062key.htm.

Process innovations refer to how techniques are implemented. For example, surgeons today can perform operations that surgeons of previous generations could not even imagine. The knowledge for these procedures was created by a few people and then taught to others in medical school and other training programs.

Process and product innovations come together when you compare how certain procedures are performed now relative to years ago. Consider surgery to repair a hernia. The first hernia (hiatal) surgery took place around 1919, and the procedure was risky and painful. Even relatively recently, a procedure like this involved hospital stays, many days lost from work, and a significant risk of medical complications.For the history of hiatal hernia surgery, see Nicholas Stylopoulos and David W. Rattner, "The History of Hiatal Hernia Surgery: From Bowditch to Laparoscopy," *Annals of Surgery* 241, no. 1 (2005): 185–193, accessed March 14, 2011, http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1356862. Today, the leading method for surgery uses a piece of capital called a laparoscope—a tube with light that allows a surgeon to see inside a patient's abdominal cavity. Then, using another instrument, again inserted through a small incision, the surgeon can repair the hernia. Remarkably, this is an outpatient procedure. The patient emerges from the hospital with a few small wounds and can return to work and normal life within a few days.

Price Determination

If this chapter were like most others in this book, we would now turn to a discussion of how supply and demand interact in a competitive market to determine the price. Or, recognizing that firms with market power set prices, we might use the condition that marginal revenue equals marginal cost to talk about price determination. Unfortunately, when it comes to understanding the market for health care, these tools are not as useful.

To understand why, imagine you want to book a hotel room in New York City. You can call up any hotel and find the price of a room. Or you can go on the Internet and check prices either at the hotel's website or at any number of other sites that provide booking services. You can find information about the hotel online, read reviews from previous guests, and talk to hotel staff members on the telephone if you need more information. If you are in the city, you can also walk into a hotel and find out the price and the hotel's amenities.

Now compare this to a hospital. It is much harder to get information about prices, and you cannot simply walk in off the street and purchase an operation. You can in fact find out prices for hospital procedures if you look hard enough. For example, there is a website that allows you to find charges for different procedures in Wisconsin (http://www.wipricepoint.org). Here you can "shop" for, say, different types of knee surgery. But these charges do not necessarily reveal the true price to you as a consumer because they may not include all the costs of doctors and other inputs. If you have insurance coverage, meanwhile, you need to find out what portion of any bill will be covered by your insurance. Figuring out the price of a procedure is quite complicated.

How, then, are prices determined? And, importantly, what price are we talking about: the price you pay or the money received by the hospital? Many of the most important prices are determined by the interaction of a few big players, including the government, insurance companies, and pharmaceutical companies. Figure 16.2.2 gives a sense of the sources of income for hospitals and doctors. Hospitals and doctors get paid by insurance companies, households, and the government.

- Medicare (http://www.medicare.gov/default.aspx) is a federal program intended to provide health services to elderly (over 65) and disabled people. It covers nearly 40 million people. Under this program, the government sets fees for services provided by physicians. A listing of those fees is available through the Health and Human Services website.See Centers for Medicare and Medicaid Services, "Overview," accessed March 14, 2011, www.cms.gov/apps/physician-fee-schedule/overview.aspx. A search engine at that site allows you to look for specific fees. These are the fees the government will pay physicians and hospitals for these services.

- *Medicaid* (www.cms.gov/MedicaidGenInfo) is run by the US government in conjunction with state governments.The states design the programs subject to approval by the federal government. Thus there are differences across states. The federal government reimburses states according to a rule that depends on the average income per person in that state. This program provides health care to low-income households through payments made directly to a health-care service provider, such as a hospital.

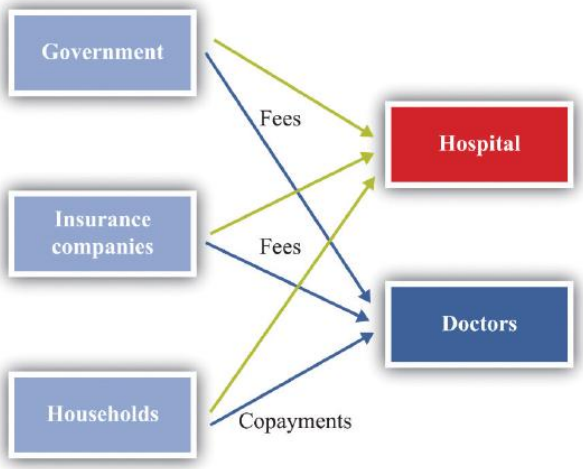


Figure 16.2.2: Payments to Hospitals and Doctors

Because of these programs, the government is a big player in the health-care market. Government decisions determine the demand for health-care services. Governments do not take prices as given. In some cases, the government sets rates for certain procedures, and health-care providers respond. In other cases, the government is involved in negotiations—with pharmaceutical companies, for example.

Insurance companies provide additional sources of revenues to the hospital and a doctor. If you are a policyholder and are admitted to a hospital, your insurance company will reimburse the hospital for part of the cost of your care. It also reimburses your doctor directly. How much of that cost is reimbursed depends on your insurance policy. If you enter a hospital, say, for an operation, the amount of money the insurance company will pay the hospital is set by an existing agreement. As a result, hospital administrators face a complex set of repayment schedules. Reimbursement rates for a given service depend on who is buying the service, as the following quotation illustrates: “Medicaid pays 80 percent of what Medicare pays and about 50 percent of what a commercial insurance carrier like Blue Cross/Blue Shield pays. For example, if Medicaid reimbursed $500 for a gall bladder removal, Medicare would pay $625 and the commercial carrier $1,000 for the same procedure.”This quote comes from a hospital in Topeka, Kansas. See Jan Biles, “Medicare: Care versus Cost, *cjonline.com*, October 23, 2007, accessed February 1, 2011, www.cjonline.com/stories/102307/sta_211162036.shtml. Because of these differences in reimbursement rates, doctors and hospitals may sometimes decide not to provide services to certain patients. The same article notes that doctors sometimes turn down Medicaid patients because of these low rates.

And what will you pay if you walk into a hospital without health insurance? An April 28, 2008, article in the *Wall Street Journal* describes the plight of a cancer patient without adequate insurance. The patient was looking for treatment at a not-for-profit hospital in Texas. Her treatment required a payment of $105,000 *in advance*. This practice of requiring prepayment is part of a trend in the industry.

Hospitals are adopting a policy to improve their finances: making medical care contingent on upfront payments. Typically, hospitals have billed people after they receive care. But now, pointing to their burgeoning bad-debt and charity-care costs, hospitals are asking patients for money before they get treated.

Hospitals say they have turned to the practice because of a spike in patients who don’t pay their bills. Uncompensated care cost the hospital industry $31.2 billion in 2006, up 44 percent from $21.6 billion in 2000, according to the American Hospital Association.Barbara Martinez, “Cash before Chemo: Hospitals Get Tough. *Wall Street Journal*, April 28, 2008, A1.

Key Takeaways

- Due to informational problems for households, market power by suppliers, and government intervention, the market for health care cannot be analyzed by using standard supply-and-demand curves.
- Spending on health care today has an effect on your health status in the future. In that sense, this spending is an investment.
- The demand for health services, like other goods, depends on your income and the price of the services. Unlike your demand for many other goods, your demand for health services is influenced by the costs of health insurance. Also, unlike the case for many other goods, consumers who demand health services are relatively uninformed about the service they are buying.
- The production function for health takes inputs, such as doctors, nurses, and machines, and produces health-care services.

Exercises

1. List three reasons why the conventional supply-and-demand model may not fit the market for health services well.
2. How is the demand for health services influenced by age?
3. Give an example of two intermediate inputs into the provision of health-care services.

This page titled 16.2: Supply and Demand in Health-Care Markets is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

16.3: Health Insurance 16.3: Health InsuranceEdit Update titlecancel

Learning Objectives

1. What are the incentive issues associated with the demand for health insurance?
2. Why is health insurance linked to employment in the United States?
3. How does the law of demand apply to the demand for health services when there is health insurance?

Insurance is something that human beings have developed to help us deal with the risks we face in life. Here are some examples of risks that you might confront.

- Your car or other property will be stolen.
- You will lose some of your possessions due to a fire, flood, storm, or other natural disaster.
- Your car will be damaged in an accident.
- You will lose your job.
- You will be injured in an accident—for example, while working, driving, or playing sports.
- You will become ill.

You can easily add to this list. We always have to worry about bad things happening. One consolation is that, for all the risks listed, you can obtain insurance. This means that we pay a fee (the premium) to an insurer; in return, we receive payment from the insurer if the bad thing happens.

Insurance is based on the idea of the **diversification** of risk. Chapter 5 goes into much more detail about insurance and diversification. As an illustration, suppose you face a 1 in 5,000 chance of breaking your leg in a given year. If this happens, it will be very costly to you: between hospital bills and lost earnings, perhaps you would lose $10,000. If you are like most people, you are **risk-averse**, meaning that you don't like facing this risk. Suppose, however, you can get together in a group of 5,000 people and agree that if any one of you breaks a leg, you will all share in the bill. The most likely outcome is that only one person will suffer a broken leg, and your share of the costs will be $2. There is still a bit of uncertainty: maybe no one will break a leg; maybe two, three, or four people will. But the likelihood that you will have to pay out more than a few dollars is very small.

Insurance companies are firms that carry out such diversification of risk by bringing together large groups of people. Insurance companies set a premium equal to the **expected value** of the loss (in the example, × $10,000 = $2), plus a fee to ensure the insurance company also profits from the deal.

Toolkit: Section 31.7

You can review the calculation of expected value in the toolkit.

Insurance, like other services, is traded in a market. You can choose to buy from a variety of sellers at a price that reflects the risk of the type of insurance you purchase. The gains from trade come from the fact that an insurance company is capable of pooling risk. The insurance company assumes your risk at a price you are willing to pay. Because people differ in terms of their attitudes toward risk, some people buy insurance against certain events, while others do not. If you are very cautious (more precisely, very risk-averse), then you are more likely to buy insurance.

What Makes Health Insurance Different?

Health insurance has the same basic structure as any other insurance: you pay a premium to an insurance company that then pays your medical bills if the need arises. Like other types of insurance, there are gains from the sharing of risk. However, health insurance differs from other kinds of insurance in a couple of ways:Melissa Thomasson, "The Importance of Group Coverage: How Tax Policy Shaped U.S. Health Insurance," *American Economic Review* 93, no. 4 (2003): 1373–1384. See also the related discussion of the history of health insurance: Melissa Thomasson, "Health Insurance in the United States," *EH.net*, February 1, 2010, accessed February 1, 2011, eh.net/encyclopedia/article/thomasson.insurance.health.us. (1) health insurance is largely provided by employers, and (2) informational problems are particularly acute.

Who Pays for Health Insurance?

In most European countries, health insurance is largely provided by the government. In some cases, the government is also a provider of health services. In the United States, the government provides some health insurance—to the very poor, the old, and military veterans. But for the most part, the provision of health insurance in the United States is very different. Table 16.3.1 "Sources of Health Insurance in the United States" shows the types of health insurance that households can obtain in the United States. The “Total” column indicates the fraction of households with insurance. Since 1999, this has averaged about 85 percent but has been falling somewhat. On the bright side, this tells us that most people are covered by insurance. It also tells us that about 50 million people in the United States have no health insurance. The table reveals in addition that by far the most important source of health insurance is through employers: about 60 percent of all individuals have insurance provided through a firm. The other forms of insurance are through the government (about 30 percent) and direct purchase (about 9 percent). These numbers add to more than 85 percent because many individuals have insurance from more than one source.

Table 16.3.1: Sources of Health Insurance in the United States

Year	Total (%)	Government (%)	Employment (%)	Direct (%)
2009	83.3	30.6	55.8	8.9
2007	84.7	27.8	59.3	8.9
2005	84.7	27.3	60.2	9.2
2003	84.9	26.6	61.0	9.3
2001	85.9	25.3	63.2	9.3
1999	86.0	24.5	63.9	10.0

Source: US Census Bureau, “Health Insurance Historical Tables,” accessed March 14, 2011, www.census.gov/hhes/www/hlthins/data/historical/files/hihistt1.xls.

It might seem odd that your health insurance is likely to be linked to your job. After all, your employer doesn’t pay for your car insurance or for insuring your bank deposits. Historically, this phenomenon has its roots in the Stabilization Act of 1942, which was signed into law by President Franklin Roosevelt. The idea of the legislation was to stabilize wages and prices during World War II. Although President Harry Truman repealed most of the provisions of the act in 1946, some of the effects of that act remain today.The end of the act by President Harry Truman is documented at the American Presidency Project, “Executive Order 9801: Removing Wage and Salary Controls Adopted Pursuant to the Stabilization Act of 1942,” accessed March 14, 2011, http://www.presidency.ucsb.edu/ws/index.php?pid=60709.

A key provision of the act established wage and price controls. This meant that wages were no longer determined by market forces but were instead set (in part) by the government. But when the government places restrictions on the way people trade, they will often try to find ways around those restrictions.This idea is at the heart of Chapter 12. The loophole in the Stabilization Act was that it exempted pensions and insurance from the calculation of wages. This meant that firms could vary the overall compensation they offered workers through the provision of pensions and health insurance. Even though wage and price controls are no longer in place, the practice of offering health insurance as part of a compensation package persisted.

An employment-based health insurance system was furthered by tax actions, such as the 1954 Internal Revenue Code, which made employer contributions to employee health insurance nontaxable. Individuals were also allowed to deduct medical expenses from taxable income. So if you are paid $1,000 in extra income by your firm and you use these funds to buy health insurance, you are taxed on the $1,000 of income. But if the firm buys the insurance for you, then you do not pay tax on the $1,000 worth of benefits.

Being employed also changes the price you pay for insurance. If you contact an insurance company directly, the rates you will be quoted for health insurance are much higher than the rate (for you and your employer combined) if you buy a health policy through your job. One explanation for this is that it is cheaper to write an insurance policy for many people together than individually. A second explanation, which we explain in more detail later, is that, on average, employed people are likely to be a lower risk than those not working. A third factor is that a group of employees is already partly diversified, so the group is less risky than a single individual.

We saw in Table 16.2.1 "Sources of Health Insurance in the United States" that about 15 percent of individuals in the United States do not possess health insurance. But who are these individuals, and why are they uninsured? The “who” is easier to answer than the “why” because we have statistics on the uninsured. Table 16.2.2 "The Uninsured (in Millions)" reveals the following:This table comes from US Census Bureau, *Income, Poverty and Health Insurance Coverage in the United States: 2009*, table 8, page 23, accessed March 14, 2011, http://www.census.gov/prod/2010pubs/p60-238.pdf.

- Many of the uninsured are poor. Of the nearly 45 million uninsured in 2005, about 14.5 million had incomes less than $25,000. (As a benchmark, according to the US Census Bureau, the 2006 poverty level for a family of four was $19,350.) Only about 17.6 percent of the uninsured had incomes in excess of $75,000. In 2009, of the 50.7 million uninsured, about 15.5 million had income less than $25,000.
  - Many of the uninsured are young. In 2005, there were 8.0 million uninsured people under the age of 18. This number was 7.5 million in 2009.
  - Many of the uninsured are young *and* poor. About 2.5 million of the 8.5 million have family incomes below the poverty line. According to the US Census Bureau, 65.5 percent of the children in poverty were covered by the government Medicaid program.Medicaid (www.cms.hhs.gov/MedicaidGenInfo) is a joint federal-state program providing funding for health care to qualifying low-income households (defined relative to the poverty level).
- Many of the uninsured are working. There were over 20 million individuals in 2005 who were working full time and yet did not have health insurance. In 2009, the number of full-time workers without health insurance was lower—14.6 million—while the number of part-time workers without health insurance was higher. We do not know from these data whether they were offered health insurance at work and declined or had jobs that did not offer this benefit.

Table 16.3.2: The Uninsured (in Millions)

	2005	2009
Number of uninsured	44.8	50.7
Age		
Under 18	8.0	7.5
18–24	8.2	8.9
Household income		
Under $25,000	14.4	15.5
$75,000 or more	7.9	10.6
Work status		
Full time	20.8	14.6
Part time	5.5	14.7

Source: US Census Bureau, *Income, Poverty and Health Insurance Coverage in the United States: 2009*, table 8, page 23, accessed March 14, 2011, http://www.census.gov/prod/2010pubs/p60-238.pdf.

Adverse Selection

One complication of health insurance markets is that those who demand insurance are the ones who are more likely to need insurance. This in itself might not be a problem, except that individuals also know more about their own health than do the companies that are insuring them.

Suppose that half the population carries a gene that gives them a 1 percent risk each year of contracting a particular kind of cancer. The other half does not carry this gene and has only a 0.1 percent risk of this cancer. If an individual becomes sick, suppose that the cost of treatment plus lost work time is $100,000. Table 16.3.3 "Probabilities and Expected Losses" summarizes the situation. Group A has a 0.1 percent chance of contracting the cancer, so the expected loss for them is $100 (= 0.001 × $100,000). Group B has a 1 percent chance of contracting the cancer, so their expected loss is $1,000 (= 0.01 × $100,000).

Table 16.3.3: Probabilities and Expected Losses

Group	Probability of Cancer (%)	Expected Loss ($)
A	0.1	100
B	1	1,000

Now let us think about an insurance company that wants to make money by selling insurance policies against this loss. Suppose these policies completely cover all losses in the event that the individual contracts the disease. If the insurance company were to set the price of a policy very high (say, $5,000), then only very risk-averse people would buy the policy. If it were to set the price very low—say, $50—then everyone in the population would want this policy, but the insurance company would make a loss on every individual.

However, suppose that the insurance company were to offer a policy for $550. It might reason as follows: if everyone buys this policy, then we will lose $450 on average from the group B individuals, but we will gain $450 on average from the group A individuals. Because there are equal numbers of both groups in the population, we should expect to make no profits on average. For example, if there are 2,000 typical individuals (1,000 of each type), then on average one group A person will become sick (because their chance is 1 in 1,000) and 10 group B individuals will become sick. In total (reasons the firm), we expect to have to pay out for 11 people, implying payments of $1.1 million (= 11 × $100,000). We will get revenues of $1.1 million (= 2,000 × $550). At this price, our expected revenues and costs are the same. If we were to charge a slightly higher premium, then we could make profits from this contract.

As long as the individuals in the population do not know their own type, this works fine. Risk-averse individuals would find it worthwhile to buy this contract. The problem comes if individuals can make a good guess as to which group they are in—perhaps because they know the history of this cancer in their own families. Then the insurance company might be in for a shock. Group B people would definitely want to buy the contract. The premium is less than their expected loss. But group A people might reason that they are very unlikely to get this disease and might decide that an insurance policy that costs $550 is much too expensive, given that their expected loss is only $100. This means that group A people—unless they are very risk-averse—choose not to buy the contract. Now the insurance company will only sell 1,000 contracts, bringing in revenue of $550,000, but it will have to pay out $1 million (= 10 × $100,000).

If the insurance company could distinguish members of group A from members of group B, then it could offer insurance at different rates to the two groups. It could offer insurance to group A at a premium of $100. They would find it worthwhile to buy this insurance. Likewise, the insurance company could offer insurance to members of group B with a premium of $1,000, and they would also find it worthwhile to buy insurance. In practice, insurance companies often cannot classify people into such precise risk groups nor offer such targeted policies. In this case, the only kind of contract that is profitable for the insurance company is one that is aimed at the group B people only, with a premium of $1,000. Group A people are left with a choice of buying no insurance at all or buying a policy that is vastly overpriced given their actual risk of contracting the disease.

This is an example of what economists call **adverse selection**: a situation in which individuals of different risk types decide whether or not to buy insurance (this is the selection). Lower-risk individuals opt out of the insurance market, leaving only high-risk individuals in the market (this is the sense in which the selection is adverse). Adverse selection is an information problem that is a source of market failure: low-risk individuals also want insurance, but it is unavailable to them at a reasonable price.

How do insurance companies deal with their informational disadvantage? One thing they can do is look for other sources of information. For example, firms presumably want to hire healthy, responsible individuals and put some time and effort into making good hires. Insurance companies can use the fact that you work for a firm as a (highly imperfect) signal of your health risk. This is one of the reasons why it is usually cheaper to get insurance through your firm than directly from an insurance company.

A second form of information about you comes from your history. If you have a car accident, your car insurance premium will increase. After an accident, your car insurance company revises its view of your riskiness and resets the price of your insurance. The analogous situation in health care is called preexisting conditions, meaning some disease or disability you already possess when you apply for insurance. For example, someone who has previously suffered a heart attack will find that insurance coverage is more expensive because the insurance company knows that this person is at greater risk of another attack.

If you apply for insurance and have a preexisting condition, then the terms of the insurance will reflect the chance that the condition will recur. This is reasonable enough: insurance is meant to provide protection against things that might happen to you in the future, not those that happened in the past. But it raises a problem with employer-provided health insurance. Suppose an

individual has health-care coverage on the job and then suffers a heart attack. His current policy covers him because the heart attack was not a preexisting condition when he obtained insurance. But if he wishes to change jobs, his heart attack becomes a preexisting condition for his new insurer. This can make it very costly to change jobs—in turn making the economy function less efficiently.

Moral Hazard

Another complicating element for insurance is the **moral hazard**: the idea that, after purchasing insurance, individuals may behave in riskier ways. For example, think about your likelihood of being in a car accident. The probability that you will have an accident depends on many things: road conditions, the actions of other drivers, luck, and many others. It also depends on the actions you take as a driver of the car. There are many things we do that influence our likelihood of having an accident, including (but not limited to) the following:

- Properly maintaining the car
- Paying attention when driving
- Driving when tired
- Driving after consuming alcohol

These items are influenced by decisions that we make. The link back to insurance is that, if we are insured, we may make different choices about the condition of our car, the way we drive, and our physical state when we drive. The analogous idea with health insurance is that we may choose to live a less healthy lifestyle or engage in riskier behavior if we know that we have health insurance to cover our expenses if we become sick or injured.

Insurance companies understand very well that their policies influence the choices that people make. Their response is to design insurance contracts that provide insurance without affecting individuals' incentives too much. In the case of automobile insurance, you will not receive full coverage for your loss in case of an accident. Instead, insurance contracts typically include the following: (1) a deductible, which is the amount of a loss you have to cover before any insurance payment occurs, and (2) a copayment, which is the share of the loss for which you are responsible. The same applies to medical insurance. In the event you are ill, health insurance will typically cover a wide variety of medical costs, but there will usually be a deductible and often a copayment as well. As with property or automobile insurance, the deductible provides an incentive for you to take actions that make you less likely to claim against the policy.

There are two main moral hazard issues with health care. First, health care is an individual investment. Although no one wants to get sick, the more you pay for your own treatment, the more likely you will invest in your own health. Choices pertaining to exercise, diet, and preventive care can all depend on the insurance payments we anticipate if we need health care. The more insurance we have, the less incentive we have to take care of ourselves. And the less we take care, the more likely we are to present the insurance company with a sizable health bill.

Second, the size of the health bill also depends on your choices about treatment. When you are ill, you will meet with your doctor to jointly decide on treatments. Although your doctor will probably talk to you about various treatment options, their price will not be the focus of the discussion. Eventually you will meet with someone else in the office to discuss how your treatment will be paid for and, in particular, how much will be covered by your insurance. In the end, you have a menu of treatments and a menu of prices that you have to pay. You will then make a choice from this menu that is in your best interest.

The insurance company pays some of your bill, so the amount you pay is lower than the actual price of treatment. By the law of demand, you purchase more than you would if you had to pay the full price. For example, you might be much more inclined to get second and third opinions if you don't have to pay the full price for these.

Even if you are not ill but are instead going to see your doctor for a checkup, incentives still come into play. Many insurance policies include funding for an annual checkup with a small copayment. We respond to those incentives by going for the annual checkups covered under the policy. We don't go for checkups every month because such visits are not covered by most policies. The insurance company deliberately designs the incentives so you are likely to find it worthwhile to engage in basic preventive care.

Health Insurance and the Law of Demand

Now that we have a better understanding of health insurance contracts, we can say more about the demand for health care. We start with the cost of health care to us as households. We have just seen that if you have health insurance, the cost to you of a trip to the

doctor is determined by your health insurance contract. Many of these contracts have a copayment provision—for example, you must pay $20 for an office visit. Of course, the doctor charges the insurance company much more for the visit, but you don't pay that cost. To you, a trip to the doctor costs $20.

The economic approach to individual choice still applies. Your demand for visits to the doctor comes from comparing the marginal valuation of these visits against this cost of $20. The law of demand works in the usual way: if your insurance company increased its copayment to, say, $50, you would make fewer visits to the doctor. The extent to which the quantity demanded responds to the price depends, of course, on what exactly is wrong with you. If you are seriously ill, your demand is likely to be inelastic. If you have only a sore throat, you might wait a few days to see if you really think you need medical care.

There is another element of the health insurance contract that has a direct effect on your demand. Consider a dental contract. These contracts often provide insurance up to an annual limit. If you need dental work, your dentist may design a treatment plan spread out over several years so that you can obtain maximum insurance coverage for the plan of work. In this case, you and your dentist are responding to the incentives of the dental contract.

Key Takeaways

- Incentive problems of adverse selection (the health insurance provider not knowing your risk class) and moral hazard (actions you take to influence your probability of needing health care) are pervasive in the provision of health insurance. These incentive problems are present when the insurance is provided by private companies and the government.
- Health insurance in the United States is linked to your job as a consequence of legislation in 1942 that exempted the provision of insurance from controls on wages.
- When you have health insurance, your demand for health services will reflect the marginal cost to you of the service. This is usually through the copayment.

Exercises

1. Suppose the provision of health insurance at your firm induces you to stop exercising. Is this an example of a moral hazard or adverse selection?
2. If the copayment increases on your dental insurance, what will this do to the frequency of your visits to the dentist and the time you devote to taking care of your teeth?
3. What is the difference between a copayment and a deductible?

This page titled 16.3: Health Insurance is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

16.4: Government Policy 16.4: Government PolicyEdit Update titlecancel

Learning Objectives

1. What is the basis for government intervention in the market for health-care services?
2. What forms does this intervention take?

In the United States and around the world, governments are involved in the provision of health care. Some form of national health insurance is commonplace in Europe and Canada. In the United States, bills promoting national health care and universal health insurance have been debated for many years.

Here is a quote from President Dwight Eisenhower's 1954 State of Union Address: "I am flatly opposed to the socialization of medicine. The great need for hospital and medical services can best be met by the initiative of private plans. But it is unfortunately a fact that medical costs are rising and already impose severe hardships on many families. The Federal Government can do many helpful things and still carefully avoid the socialization of medicine."AMDOCS: Documents for the Study of American History, "Dwight Eisenhower, 'State of the Union, 1954,'" accessed March 14, 2011, http://www.vlib.us/amdocs/texts/dde1954.htm.

As we noted earlier, the federal tax code was modified in 1954 to provide incentives for employer-provided health insurance. We structure our discussion of government health policy by answering two questions: (1) why do governments intervene? and (2) how do they intervene?

Why Do Governments Intervene?

As usual, we analyze government involvement in the economy through the lens of market failure. When it comes to health care, there are several market failures to consider.

Externalities. One argument for public involvement in health care is the presence of **externalities**. If one individual is sick, then the likelihood that others around that person get sick increases. Individuals typically make decisions about their health care without thinking much about the effects of their decisions on the welfare of others. You may decide to go to work even though you are suffering from the flu because you need the money, and you may not think very much about your likelihood of infecting others. This is a classic example of an externality.

Commitment. Hospitals are often unwilling or unable to turn away individuals needing care but lack the resources to pay for that care. Through legislation passed in 1986, hospitals are required to treat patients in emergency situations, no matter what their insurance coverage.The regulation is called EMTALA (http://www.emtala.com/history.htm). It literally applies only to those hospitals that accept Medicare, but this is almost universal. In many cases, this is an inefficient way to treat people. For example, one consequence is that the uninsured have an incentive to seek normal care in hospital emergency rooms, even though this is an expensive place to provide care.

If hospitals could commit not to serve people unless they had health insurance, then some of the uninsured might be induced to purchase health insurance, instead of relying on emergency wards. But hospitals are not able to make such a commitment; although this might be more efficient, it is also unacceptably callous and runs counter to the Hippocratic Oath.

Adverse selection and moral hazard. We explained earlier that when insurance companies are unable to observe the probability of illnesses, some individuals will obtain insurance while others do not. Although these choices may be optimal from the standpoint of an individual, the market outcome is not efficient.

Drug quality. The health-care market is filled with gaps in information. Patients and even their doctors cannot fully assess the safety and efficacy of pharmaceutical products. Although the drug companies test their own products, the government has a role in assessing this information and determining the safety and effectiveness of medications.

Doctor quality. Another informational problem in the health-care market is the inability of a patient to properly evaluate the quality of a doctor. You as a patient can look at some indications of your doctor's ability, such as years of practice, school of graduation, and number of people in the waiting room. But it is not possible to make a fully informed judgment about the quality of your doctor. Again, the government plays a role by requiring that doctors obtain specialized training and pass a licensing examination before they are allowed to practice.

Patents. The research and development needed to create a new drug is substantial.We also discuss this in Chapter 15. For firms to earn a return on this investment, the patent system exists to provide them protection from other firms producing the same product

and selling it at a lower price. Although this type of competition may be valued *given that a product exists*, it destroys the initial incentives that a firm has to undertake research and development. Governments provide patent protection to induce firms to undertake the necessary research and development.

Market power. Market outcomes are not efficient when there are relatively few sellers of a product. This may occur in various health-care markets because there may be relatively few doctors and few hospitals in a given location. Furthermore, pharmaceutical companies have market power based on exclusive knowledge of their specific product, as protected through patents. Finally, there are relatively few health insurance providers, and some are very large.

Equity and fairness. Even if health-care markets were efficient (and we have explained many ways in which they are not), they may not be equitable. One argument for government involvement is to provide for a more equitable allocation of goods and services. From this perspective, the fact that many Americans lack health insurance and adequate health care is also a basis for government involvement.

Article 25 of the Universal Declaration of Human Rights includes the right to health care: “Everyone has the right to a standard of living adequate for the health and well-being of himself and of his family, including food, clothing, housing and medical care and necessary social services, and the right to security in the event of unemployment, sickness, disability, widowhood, old age or other lack of livelihood in circumstances beyond his control” (http://www.un.org/en/documents/udhr/index.shtml).United Nations, “Universal Declaration of Human Rights,” accessed March 14, 2011, http://www.un.org/en/documents/udhr/index.shtml. To the extent that basic health care is viewed as a basic human right, then the government ought to guarantee access to at least a minimal level of care.

Indeed, if the aim is to move toward equality of well-being, there is an even stronger equity argument for health care. Imagine for a moment that people could decide how to allocate health care and other resources before they knew anything about their own health or well-being. (Thought experiments of this kind are associated with the philosopher John Rawls. In Chapter 13, we discuss this kind of thought experiment in more detail.) People might well agree that those who became sick or disabled should be given *extra resources* to compensate them for their ill health.

How Do Governments Intervene?

Now that we have some understanding of the sources of market failure in the health-care market, we turn to a discussion of government policy.

Taxes and Subsidies

We have already mentioned one of the key ways in which the government subsidizes health care—by allowing employees tax-free health insurance benefits provided by an employer. In this way, the government reduces the cost of firm-provided health care. It is now common for employment contracts in the United States to include provision for health care.

One of the main issues surrounding employer-provided health insurance is the possibility of losing insurance when you change jobs (sometimes called the “portability problem”). In our economy, shifts in demand for goods and services and changes in productivity naturally lead to the creation of new jobs by some firms and the destruction of jobs by other (perhaps less profitable) firms. The efficient working of an economy therefore requires that workers leave old jobs for new ones. Unfortunately, insurance can get in the way of worker mobility. If you have a job with health insurance, then quitting your job to look for another may be costly for several reasons. First, you may lose insurance coverage during the period of job search. Second, an ailment that was covered by insurance by your existing firm could be viewed as a preexisting condition when you acquire insurance at a new firm. This can have an adverse effect on your insurance rates and the type of coverage you can obtain. In some cases, people choose not to change jobs purely because of the implications for health insurance.

Health care is also subsidized through income taxes. If you look carefully at your income tax forms, you will see that you can deduct medical expenses. If you itemize deductions on your tax form, and if your medical expenses are substantial enough, you can offset those payments against your taxes.

Regulation

Government regulations are common in the health industry. These regulations influence both demand and supply in this market.

On the demand side, households are required to obtain certain medical services. Some vaccinations are mandatory, for example. At the college level, there is ongoing concern about the spread of meningitis.See “Meningitis on Campus,” *American College Health*

Association, April 27, 2005, accessed February 1, 2011, www.acha.org/projects_programs/meningitis/index.cfm. With concerns like this in mind, it is also common for colleges to require some vaccinations prior to admission. The argument for such interventions is that there are externalities from your health to the health of others.

The government licenses many of the actors on the supply side of the health-care market. This is another form of quality control. Doctors who practice in a state must pass exams called medical boards. Hospitals are certified for the types of activities they offer. Often the certification occurs at the state level.For details about accreditation in Texas, for example, see Texas Department of State Health Service, "General Hospitals—Health Facility Program," accessed March 14, 2011, http://www.dshs.state.tx.us/HFP/hospital.shtm. Other providers of health care are also licensed. For example, a nursing home must be certified as a Medicare provider to receive reimbursements. The rationale for such interventions stems from the extensive information problems in the health-care market. As consumers are unable to accurately assess the quality of care provided by doctors and hospitals, the government provides a service to us all by regulating health-care providers.

Provision of Insurance

The government, through its Medicaid and Medicare programs, provides insurance to both low-income and elderly households. There is continuing debate about expanding the availability of health insurance to the general population. On March 23, 2010, President Obama signed a health-care reform bill.Ample discussion by the White House along with the final bill is available at "Health Reform Puts American Families and Small Business Owners in Control of Their Own Health Care," The White House, accessed March 14, 2011, www.whitehouse.gov/health-care-meeting/proposal. The main goal of the bill is to reduce the number of individuals without health insurance in the United States. This bill seeks to achieve this by requiring that everyone purchase health insurance, either through an employer or individually. The legislation provides opportunities for households to obtain insurance on their own through subsidies and a "marketplace for insurance."

The bill regulates insurance policies in several ways. For example, it places limits on the ability of insurance companies to exclude people from coverage due to preexisting health conditions or other health risks. It also restricts the ability of insurance companies to set and change rates on insurance policies.

The new policy will not be fully in force until 2014. It is extremely complex (the bill itself is almost 1,000 pages long), and its impact on health-care outcomes, health-care costs, and the deficit remains an open question. (Even the short summary of the actLibrary of Congress, "Bill Summary & Status: 111th Congress (2009–2010), H.R.3590, CRS Summary," accessed March 14, 2011, thomas.loc.gov/cgi-bin/bdquery/z?d111:HR03590:@@@D&summ2=m&. does not make easy reading.) By the time of its implementation, details on the new law may be clearer. In particular, exactly how insurance markets will be organized and regulated will be made more precise. Further, when the bill was passed, estimates were made of the cost savings from the measure. Over time, we will be better able to forecast the spending and taxation implications of the bill once household and firm responses are observed. Then we can see if this legislation improves the efficiency of the health-care market.

A more fundamental question is whether the government should even be in the business of providing health insurance. One set of arguments for government involvement rests on the various market failures that we have identified in this chapter. Health care is complicated, and there are many ways in which health-care markets depart from the competitive ideal. It is sometimes argued that spending on health services in the United States is very high because the market is very inefficient. From that perspective, having the government in charge of this sector of the economy might reduce inefficiencies. Second, government involvement can be justified on the grounds of equity and fairness.

Provision of Information

One of the primary roles of the government is to provide information to the public about health matters. This comes in a variety of forms. In January 1966, the following warning first appeared on cigarette packs: "Warning: Cigarette Smoking May be Hazardous to Your Health." This initial warning from the Surgeon General's office of the United States was followed by many others concerning the consumption of cigarettes and other potentially harmful products. Such warnings are a good example of government provision of information. Each consumer of these products wants to know the impact on health. Gathering such information is a public good because the information is available to everyone and can be "consumed" by everyone simultaneously.

Another form of information is through drug testing. The US Food and Drug Administration (FDA; http://www.fda.gov) is responsible for testing drugs before they appear on the market. The FDA also supplies public information about a wide range of food items.

Key Takeaways

- Government intervention in this market reflects inefficiencies in the market as well as concerns over equity.
- Government intervention takes many forms around the world, including the provision of health insurance, the direct provision of health-care services, and the regulation of drug companies.

Exercises

1. What is the commitment problem of a hospital?
2. Recent concerns about the H1N1 virus led the governments in many countries to intervene. How would you explain the basis for this intervention using the list of market failures provided in this section?
3. Give a recent example where the government provided a health warning.

This page titled 16.4: Government Policy is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

16.5: End-of-Chapter Material 16.5: End-of-Chapter MaterialEdit Update titlecancel

In Conclusion

The debates that we have introduced in this chapter are far from settled. The issue of health-care policy is one that is not likely to be quickly resolved in the United States or elsewhere in the world. At the moment, we see many different ways in which health care is provided in different countries.

There is enormous pressure within the United States to deal with the perceived problem of uninsured households. As we have seen, this is a key element of the health-care bill that was signed into law in 2010. In addition, there are unsolved problems associated with Medicare and Medicaid. These government programs are in need of reform to deal with the escalating costs of medical care. According to the General Accounting Office, the current level of Medicare and Medicaid spending is about 4 percent of the gross domestic product (GDP). If there are no changes in current programs, this fraction is expected to increase to about 11 percent by 2050.This draws the discussion of the US budget through 2075: "A 125-Year Picture of the Federal Government's Share of the Economy, 1950 to 2075," *Congressional Budget Office*, July 3, 2002, accessed February 1, 2011, http://www.cbo.gov/ftpdoc.cfm?index=3521. Also see David M. Walker, "U.S. Financial Condition and Fiscal Future Briefing," US Government Accountability Office, January 2, 2008, accessed February 1, 2011, http://www.gao.gov/cghome/d08395cg.pdf. To put this in perspective, government outlays as a fraction of GDP have typically been about 20 percent of GDP over the past 40 or so years. This number is forecast to increase to nearly 27 percent by 2050, with a significant fraction of this driven by Medicare and Medicaid programs.

There are two overarching lessons to be drawn from this chapter.

1. **Incentives matter.** Whatever programs are deemed to be desirable for dealing with health-care problems, we need to be sure to take into account the incentives that these policies will create.
2. **People are different.** Health-care providers, insurance companies, and the government cannot observe many of these differences. Policies must take into account all these differences and recognize the importance of adverse selection. To the extent that markets fail because of adverse selection, government policy must address this source of the underlying problem of health care.

The nature of adverse selection is heavily influenced by technology. Recently, the US Congress has been considering legislation, HR 493,Govtrack.us, "Text of H.R. 493 [110th]: Genetic Information Nondiscrimination Act of 2008," accessed March 14, 2011, http://www.govtrack.us/congress/billtext.xpd?bill=h110-493&show-changes=0&page-command=print. which limits the ability of insurance companies to use genetic information. If insurance companies have more information about individual health risks, then they can design more targeted insurance contracts. From the perspective of efficiency, this might seem to be a good thing because it eliminates some adverse selection problems. Imagine that technology were to reach a point where all your major health risks could be identified from your genetic code at birth. Almost all the uncertainty over your health would disappear, and there would be almost no role for health insurance. (There could, of course, still be insurance for accident risk.) Although this world might be more efficient, it would also be much more unequal. People with genetic predispositions to certain illnesses would face steep medical bills, while healthy people would not.

In this world people would want insurance before they were born, while there was still uncertainty about their genetic makeup, which takes us back to the Rawlsian thought experiment that we discussed earlier. It might be that, paradoxically, one of the most compelling arguments for government-provided universal health care will turn out to be the gradual *elimination* of market failures from adverse selection.

Key Links

- World Health Organization: http://www.who.int/nha/en
- US Department of Health and Human Services: http://www.hhs.gov
- US Census Bureau, "Income, Poverty and Health Insurance Coverage in the United States: 2009" http://www.census.gov/prod/2010pubs/p60-238.pdf
- Bureau of Labor Statistics, Consumer Price Index in medical care: www.bls.gov/cpi/cpifact4.htm

exercises

1. Is health insurance a complement or a substitute for the demand for health-care services?
2. If doctors no longer needed a license to practice so that there was free entry into the provision of that service, what would happen to the price and quantity of health-care services? What would happen to the quality?
3. Give an example of technological innovation in the health-care industry.
4. If adverse selection is a problem, does allowing an insurance company to know your entire health history improve market efficiency?
5. What type of inefficiencies does the commitment problem of a hospital create? Why don't restaurants have this same problem?
6. One government intervention in health care is compulsory vaccinations for children against various infectious diseases. Can you explain why governments might enact such policies? (Hint: are there any externalities involved?)
7. In some countries and some regions, there are shortages of doctors. Why is this problem not quickly resolved by the normal workings of supply and demand in the labor market?
8. As the United States has become richer, an increasing proportion of GDP is spent on health care. Does this fact, in and of itself, indicate inefficiencies in health care? (Hint: to what extent are aspects of health care luxury goods?)
9. One complicated part of the demand for health care is that consumers are not quite sure of the quality of the product they are buying. Can you think of other goods or services that have this same property? Are there measures to protect consumers? Why don't sellers sell only low-quality goods to consumers who are not able to judge quality?
10. In the health-care market, private and public hospitals coexist. Can you think of another market in which both public and private providers exist? What are the differences in that market among the public and private firms?
11. What are the incentives for parents to provide health care for their children?
12. The next time you visit a doctor, ask for a price list. Discuss what happens.
13. The government provides deposit insurance so that funds deposited at a bank are insured even if the bank goes out of business. What are the moral hazard implications of providing deposit insurance?
14. What does the link of health-care coverage to employment do to the incentives of someone to quit one job and look for another?

Economics Detective

1. Table 16.2.2 "The Uninsured (in Millions)" came from a 2006 and 2010 census. What is the current number of uninsured? What fraction is under 18 and in the 18–24 age group?
2. Find out about the health-care system in France, Sweden, or Canada. How does it compare with the US system? How do health outcomes in the country you have chosen compare with those in the United States?
3. The election results in November 2010 reflected, in part, concerns over the health-care bill signed into law in March 2010. What were the main concerns discussed in the election campaign with regard to the bill?

This page titled 16.5: End-of-Chapter Material is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

17: Cars

17.1: Looking into Cars
17.2: The Demand for Automobiles
17.3: Supply of Cars
17.4: Market Outcomes in the Automobile Industry
17.5: Policy Issues
17.6: End-of-Chapter Material

This page titled 17: Cars is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

17.1: Looking into Cars 17.1: Looking into CarsEdit Update titlecancel

Cars are so common that we rarely give them a second thought. If you live in the United States, then you either already have a car or will most likely have one in a few years' time; over 90 percent of households in the United States either own or lease a car. Although other countries do not have quite the same levels of car ownership as the United States, there are more than half a billion automobiles in the world.

The familiarity of the car is so great that it is easy to forget how the automobile transformed the world. The automobile made modern cities and suburbs possible because people were no longer obliged to live close to where they worked. The automobile made it easier to transport goods from place to place, dramatically altering patterns of trade in the global economy. At the same time, automobile emissions have degraded the air we breathe to the point where they sometimes seriously damage people's health. Indeed, because emissions also contribute to the accumulation of greenhouse gases in the atmosphere, the automobile may be changing the very climate of the planet.

Although you may own a car, it is likely that you are unfamiliar with how it works. Half a century ago, car owners were typically very knowledgeable about how their vehicles operated. They needed to be because cars broke down frequently. People knew how to adjust spark plugs, clean distributor caps, and so on. But the modern automobile is a remarkably sophisticated and complex piece of engineering. Today, it is unlikely that an owner of a modern-day car knows how to do much more than very basic maintenance. Even car mechanics rely on computer diagnostics to perform repairs.

Just as the product itself has become increasingly complex, so too has its method of manufacture. In the early years of the 20th century, cars were produced in small numbers and largely by hand. In 1913, however, Henry Ford introduced mass production of cars at Ford's Highland Park plant.The Library of Congress has an extensive discussion of Henry Ford, including photos of production in 1923 at http://memory.loc.gov/ammem/today/jul30.html. By the middle of the 20th century, cars were typically produced on assembly lines. In contrast to the early years of car production, there were far fewer workers at this stage of the production process.You can find more details at How Stuff Works, "1957–1959 Ford Fairlane 500 Skyliner," accessed March 14, 2011, http://auto.howstuffworks.com/1957-1959-ford-fairlane-skyliner3.htm.

If you were to visit the production line at a modern automobile manufacturing plant, you would hardly see any people at all. Modern production uses a great deal of capital and relatively little labor. Computerized robots perform manufacturing roles. Yet despite the relative absence of workers on the production line, over one million workers are employed in this sector of the economy in the United States.Basic information about firms and workers in this sector comes from the Bureau of Labor Statistics, "Career Guide to Industries, 2010–11 Edition," www.bls.gov/oco/cg/cgs012.htm. In 2006, about 360,000 jobs were associated with the production of automobile parts alone.This number is from a Bureau of Labor Statistics study of employment in the automobile parts industry: Benjamin Collins, Thomas McDonald, and Jay A. Mousa, "The Rise and Decline of Auto Parts Manufacturing in the Midwest," *Monthly Labor Review Online* 130, no. 10 (2007): 14–20, accessed March 14, 2011, http://www.bls.gov/opub/mlr/2007/10/art2full.pdf. In the 21st century, though, there have been significant job losses in this part of the economy.

Road Map

This chapter is different from others in this book because it is a capstone discussion. We use the automobile industry to illustrate the different ideas we have explained in the book. We also use this industry to provide further examples of how to use the different tools we introduced in previous chapters. Whereas other chapters were largely self-contained, here we will repeatedly remind you of ideas that we have already studied.

We begin our look at cars in a familiar way, using the supply-and-demand framework. All of us who own cars reside on the demand side of the market. We make choices about the type of car we want, whether to buy a new car or a used one, and when to replace it. We also make decisions about related products like gasoline and insurance. The supply side of the car market illustrates technological progress, enormous growth in product diversity, the impact of trade on domestic markets, and the social costs of automobiles. We examine some decisions of automobile producers, including where to locate their operations, why they introduce new models, and what price they should set.

We then study the equilibrium of the car industry. The US car industry began around the start of the 20th century, survived the Great Depression of the early 1930s, and has been transformed by international competition. Understanding these dynamics provides a perspective on other industries. After understanding industry equilibrium, we look at the variety of government policies that impact this industry, including trade and environmental policies.

This page titled 17.1: Looking into Cars is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

17.2: The Demand for Automobiles 17.2: The Demand for AutomobilesEdit Update titlecancel

Learning Objectives

1. What determines a household's demand for a car?
2. In what sense is a car an asset?
3. What are some of the complementary goods and services for cars?

People don't buy cars only because they want to look at a piece of fine engineering and enjoy a luxurious ride (although this sometimes plays a role). They buy cars because they want to be able to travel from one place to another. The demand for automobiles is a piece of a larger market: the demand for transportation in general. As the price of a particular car increases, the law of demand tells us that the quantity demanded of that car will decrease. There are three kinds of substitution at work here. In response to a price increase, households can

- delay the purchase of a new car,
- choose to purchase another type of car, or
- choose not to buy a car and use another mode of transportation instead.

Suppose you are thinking of buying a car, but the price of your favorite model increases. One possible response would be to delay your purchase until a later time. With this form of substitution, you decide not to buy a new car right now. This does not mean you will never buy a new car. Instead, you are keeping your options open for the future: you will drive your old car for perhaps another year and then search again next year for a replacement. Next year, of course, you might decide to defer your purchase still further. A second possible response to a price increase in your preferred model is to purchase another type of car. There is a substitution effect at work again, but now it applies across cars rather than over time. Perhaps you are indifferent between buying a Ferrari racing car and a Mini Cooper. If the price of the Mini Cooper increases, you would be induced to buy the Ferrari. Finally, if the price of your preferred car increases, you might substitute another form of transportation for your car, such as a bicycle or the bus. From this perspective, your demand for a car is really a demand for transportation.

Household Demand for Cars

The decision to buy a car is best understood as an example of *unit demand*. Most households—even if they own more than one car—do not buy a large number of cars at a time. Instead, they buy a single car. The decision about whether or not to purchase a car thus involves comparing the **valuation** a household places on the car with the price of the car. One way to illustrate this is to look at the household's **budget line** when it does or does not purchase a car.

Toolkit: Section 31.1

You can review unit demand, valuation, and budget lines in the toolkit.

The household's choice is shown in Figure 17.2.1. (This draws on our presentation of unit demand in Chapter 4.) The household can spend income on three items: chocolate bars, downloads, and a new car. If the household chooses not to buy a car, then it consumes the combination of downloads and chocolate bars indicated by point A in the graph. This is the household's most preferred point in the budget line, given that it does not buy a car. If the household buys a car, then the combination of downloads and chocolate bars it consumes is given by point B. The budget line is shifted inward by the amount of income the household spends on the car—that is, by an amount equal to the cost of the car. The household's decision about whether to purchase the car involves comparing bundle A to (bundle B plus a car).

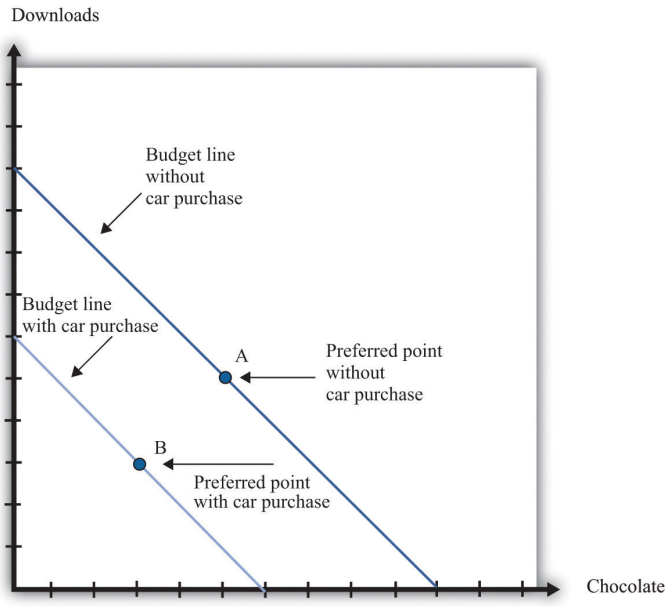


Figure 17.2.1: The Household's Budget Line When It Does or Does Not Buy a Car. Here we illustrate a household's choice about whether or not to purchase a car.

Remembering the idea of **buyer surplus**, this is the same as saying that the household would buy the car if the purchase gave it some surplus. In other words, the household's decision rule is

$$purchase\ car\ if\ valuation\ of\ car\ -\ price\ of\ car\ =\ buyer\ surplus\ >\ 0.$$

If the price of the car is greater than the household's valuation, the household prefers point A to point B and does not buy the car. If the price of the car is less than the household's valuation, it prefers point B to point A and buys the car.

Toolkit: Section 31.10

Buyer surplus was introduced in Chapter 6. You can also review the various kinds of surplus in the toolkit.

There is another way of looking at the same decision that gives us a way to measure the household's valuation of the car. Remember that the household's valuation of the car is the *maximum* amount that it would be willing to pay for it. Look again at Figure 17.2.1 and begin at point A. Now move the budget line inward until we find that the household is just as happy at point A or point B. We have now found the point where the household is *indifferent* between the combination of chocolate bars and downloads it buys without a car and the bundle it buys along with the car. The amount by which we have moved the budget line is the household's valuation of the car.

If there were only a single model of car for the household to choose from, we could stop here. The household would compare the valuation of the car against the price and buy the car as long as the valuation is greater than the price. Today, however, cars differ in numerous ways. Like many goods, a car consists of many different features all bundled together. These include the car's performance features, styling, color, and sound system; whether it has leather seats, a sunroof, and air conditioning; and hundreds of other attributes that we could list. The household's valuation of a car embodies a valuation of each attribute of a car.

This complexity makes the decision to buy a car a challenge. How can this decision be made? For every car available on the market, the household can calculate the buyer surplus attainable from that car. After considering all these alternatives, the household should then buy the car that gives the most surplus. Of course, households do not literally sit down with a list of cars and try to calculate the exact surplus from each one. But this is a useful, if stylized, representation of how such choices are made. In effect, the household is making a unit demand decision—buy or not buy—about every single car. For almost all cars, the household chooses to purchase zero.

(There is a subtlety you may be wondering about here. Hundreds of different cars might yield positive surplus, but obviously the household does not buy hundreds of cars. The trick is that, once the household has bought the car that gives the highest surplus, the valuation of all other cars it might consider buying decreases substantially. If you don't own a car, then a Ford Focus might be very valuable to you. If you already own a Mazda 5, then the value of a Ford Focus would be much smaller.)

Deciding what car to buy is only one part of the household's decision. As we already noted, the household must decide *when* it wants to buy a car. A car is a **durable good**; it lasts for several years. If a household already has a car, it can decide to defer purchase of a new car until later. A household is likely to do this if (1) there is substantial uncertainty about future income (perhaps members of the household fear losing their jobs) or (2) cars are likely to be relatively cheaper in the future. To understand this choice, we turn to some of the tools introduced in Chapter 5, and Chapter 10.

A car is an asset: it yields a flow of services. As a consequence, buying a car is both an act of consumption and an act of saving. This means that the decision to buy a car is an example of decision making over time. The household looks at both current and expected future income when deciding about the purchase, and it knows that the car will yield benefits for many years. Furthermore, because a car is an asset, its valuation today depends on its value in the future. You might buy a car this year and then discover your transportation needs have changed. In that event, you can sell your car in the used car market. The more you expect to get for your car whenever you might sell it, the more you are willing to pay for it today.

The demand for a particular car also depends on factors other than the price of the car itself. The prices of other goods—most importantly, other cars and other forms of transportation—also matter. Household income, both now and in future years, is another determinant of demand. Finally, because you often purchase a car with a car loan, the interest rates charged on loans may matter for your car purchase. A decrease in the interest rate for car loans will increase the demand for cars.

In sum, buying a car is a very complex decision. There are rich substitution possibilities involving the choice of different models, the timing of the purchase, and the possibility of using public transportation rather than owning a car at all. The law of demand applies to cars, just as it does to other goods and services. But as we move along the demand curve in response to a change in the price of cars, the substitution possibilities are complex. Understanding these substitution possibilities is critical when firms are choosing the prices to set for the cars that they produce.

Complementary Products

There are several products that are complementary to the purchase of a car. Here we look at three: gasoline, insurance, and infrastructure. A complementary product is one for which the **cross-price elasticity of demand** is negative. In other words, we expect that if the price of gasoline increases, the quantity of cars purchased will decrease.

Toolkit: Section 31.2

The cross-price elasticity of demand measures the response of the quantity demanded of a good to a change in the price of another good. Formally, it is the percentage change in the quantity demanded of one good divided by the percentage change in the price of another good:

- If the cross-price elasticity of demand is a negative number, then the quantity demanded of one good decreases when the price of the other good increases. In this case, we say that the two goods are **complements**.
- If the cross-price elasticity of demand is a positive number, then the quantity demanded of one good increases when the price of the other good increases. In this case, we say that the two goods are **substitutes**.

Gasoline

During late spring of 2008, the price of oil and gasoline rose considerably. By the end of May, the price of unleaded gas averaged over $4.00 per gallon in major US cities. Figure 17.2.2 shows the price of regular gasoline in the United States. The first series is just the price of a gallon of gas in pennies. The second shows prices after *correcting for inflation*; all prices are quoted relative to the price level in the base year of 1990. You can see from this figure that the dollar price of gas has increased steadily since August 1990, but once we correct for changes in the overall price level, the real price of gas was actually decreasing until around 1998. Since then, it has increased to about twice its 1990 level.

Toolkit: Section 31.8

You can review how to correct for inflation in the toolkit.

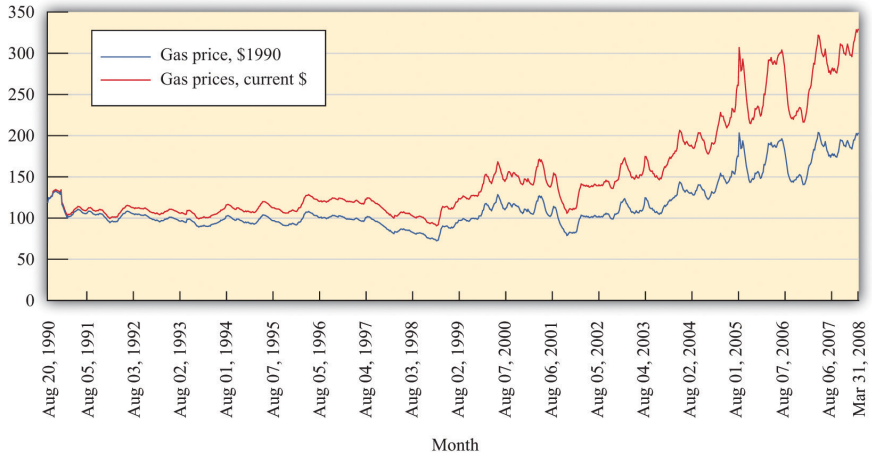


Figure 17.2.2: Gasoline Prices: Pennies per Gallon

What is the impact of an increase in the price of gasoline on the demand for cars?

- As the price of gas increases, buying and operating a car becomes more expensive. Thus we expect the demand for cars to decrease.
- An increase in the price of gasoline induces a move away from cars with low mileage per gallon to cars with higher mileage per gallon. This type of substitution is often seen after rapid increases in the price of gasoline.
- As the price of gasoline increases, individuals substitute competing means of transportation for cars. For example, an increase in the price of gasoline might induce commuters to use bicycles, take buses or trains, or walk to work.

All of these channels were in the news in 2008 as consumers responded to higher gasoline prices by driving less and buying more fuel-efficient vehicles.

If we think of cars in general, then the cross-price elasticity of demand with the price of gasoline is negative. But the second channel reminds us that, for fuel-efficient models, the cross-price elasticity of demand might be positive. Higher gas prices mean that fewer cars will be purchased, but households that do purchase favor cars that are more fuel-efficient. Remembering that cars typically last for several years, households think about not only the current price of gas but also what they expect gas prices to do in the next few years. If gas prices increase but consumers think that this increase is likely to be temporary, then people will drive less, but the demand for cars will be little affected. Conversely, if gas prices increase and consumers expect them to stay high for some years, we see a much bigger effect on the demand for automobiles.

The gas price increase in 2008 coincided with an intense battle between Senator Hilary Clinton and Senator Barack Obama for the Democratic presidential nomination. By this time, Senator John McCain had effectively won the Republican nomination. Senator Clinton and Senator McCain proposed a temporary reduction in the tax on gasoline in an effort to shield households from the high price of gasoline during the summer season. To see the effects of such a tax, we can use the tool of **tax incidence**, which we introduced in Chapter 12.

Toolkit: Section 31.11

You can review tax incidence in the toolkit.

Figure 17.2.3 shows the effect of a tax on the price paid by a buyer and the price received by a seller. In part (a) of Figure 17.2.3, the demand for the product is very inelastic. As a consequence, when a tax is imposed, the price paid by the buyer increases a lot compared to the price in the absence of a tax. This means that when demand is inelastic, buyers bear the tax burden. Consequently, if a tax is removed, even temporarily, the price buyers pay will decrease considerably. As shown in part (b) of Figure 17.2.3, we reach the opposite conclusion if supply rather than demand is inelastic. When supply is inelastic, the price received by the seller varies with the tax, while the price paid by the buyer is almost independent of the tax.

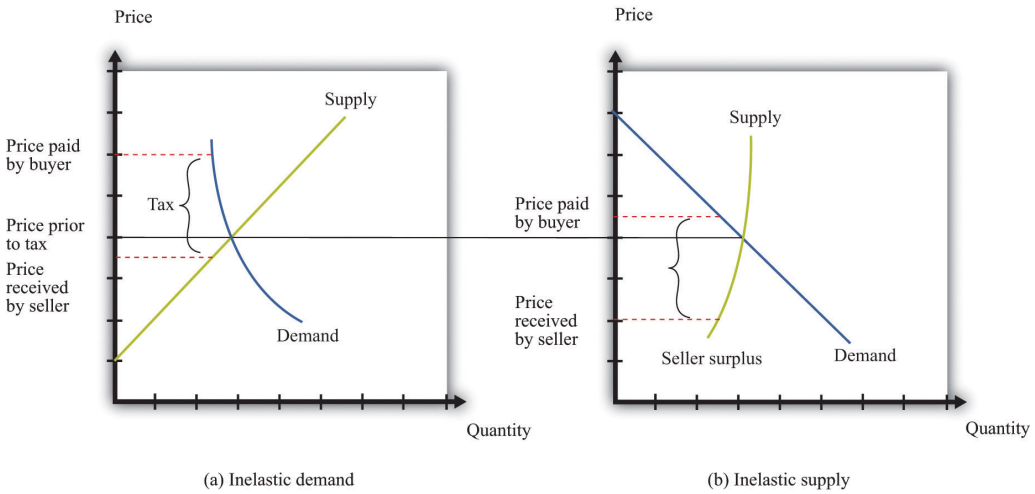


Figure 17.2.3: The Incidence of a Tax on Gasoline. When demand is inelastic (a), the price paid by buyers increases a lot, and the price received by sellers decreases only a little bit, so most of the tax burden is borne by buyers. The opposite is true when supply is inelastic (b).

Whether a temporary repeal of the gas tax would reduce the price of gas depends on the elasticities of supply and demand in this market. The more inelastic demand is relative to supply, the more the tax reduction will lower the price paid by households. In fact, the supply of gas tends to be very inelastic in the short run because refining capacity is limited. Part (a) of Figure 17.2.3 is the one relevant for a temporary change in the gas tax.

Insurance

A second complementary product is car insurance. In most countries, having car insurance is mandatory. Typically, drivers must at the very least purchase some liability coverage, meaning that your insurance company will pay out if you are responsible for an accident that injures or kills another person or causes damage to a car or property. You may also choose to buy collision coverage to cover damage to your own car. You can also purchase “uninsured motorist coverage,” which protects you in the event you are in an accident with someone who is uninsured. Exactly what type of coverage you are able (or required) to purchase varies from country to country and from state to state in the United States.

You may be aware that your insurance rates depend on your age, gender, and driving record. Insurance companies work very hard to determine the probability you will have an accident and make a claim for funds from them. But they do not know exactly what that probability is for each person individually. Thus they rely on information about us, such as our age and sex, and also look at driving records for indicators of the likelihood that we will file a claim. If you have an accident and file a claim, the insurance company will often revise its assessment of how safe a driver you are and increase your premium. Such adjustments can be so severe that sometimes people prefer to fix their own car after a minor incident than have the insurance company handle a claim. It is even possible to obtain insurance against the costs of obtaining a traffic ticket. Getting a ticket can be expensive, both directly in terms of a fine and then through an increased insurance premium. Recognizing the desire of a household to shed this risk, insurance policies that compensate ticketed drivers for these costs are available.One such program can be found at the TraffiCare International home page, accessed March 14, 2011, trafficare.net.

Government restrictions on trades often lead people to avoid these restrictions. This was one of our themes in Chapter 12. In many states, insurance is required by law, yet drivers sometimes flout this law, choosing to drive without insurance instead. The law of demand tells us that the higher the cost of insurance, the more likely people are to drive without insurance.

If high insurance rates lead to a large number of uninsured motorists, then it is more likely that, if you have an accident, the other driver will be uninsured. In this case, your insurance company will be obliged to cover your damages, even if the other driver was at fault. Think for a moment about what this means from the perspective of an individual insurance company. If there are more uninsured motorists, then the insurance company expects to make higher payouts per accident, on average. This means that when there are more uninsured motorists, insurance premiums will be higher.

Economists Eric Smith and Randy Wright noticed that insurance premiums are very different in different places. In a paper titled “Why Is Automobile Insurance in Philadelphia So Damn Expensive?” they speculated that this was because of a **coordination game**. They argued that the decision to purchase insurance could lead to payoffs like those in Table 17.2.1 "A Coordination Game

for Automobile Insurance". The rows show your decision about whether to buy insurance or not. The columns show other people's decisions. The numbers in the table refer to your payoff from every combination of what you choose to do and what everyone else chooses. (To keep things simple, we state what your payoffs are when everyone else does the same thing. We also suppose that everyone else faces the same payoffs you do.)

Toolkit: Section 31.18

You can review the coordination game in the toolkit.

Table 17.2.1: A Coordination Game for Automobile Insurance

	Everyone Else Buys Insurance	Everyone Else Does Not Buy Insurance
You Buy Insurance	10	2
You Do Not Buy Insurance	4	8

Look at the first column. This says that if everyone else purchases insurance, then you have an incentive to do so as well. We see this in the table because 10 is greater than 4. Because everyone else buys insurance, the price of insurance will be relatively low, and you will be induced to buy insurance as well. Thus one Nash equilibrium of this game is for everyone to buy insurance. Now look at the second column. Here, everyone else is an uninsured motorist. If no one else buys insurance, your insurance will be very costly. This will induce you not to buy insurance either (8 is greater than 2). Thus there is a second Nash equilibrium of this game in which no one buys insurance.

If, as Smith and Wright suggested, Philadelphia was an example of a city in which no one bought insurance, you can see from Table 17.2.1 "A Coordination Game for Automobile Insurance" that everyone there *wished* that other people bought insurance. Everybody is better off in the Nash equilibrium where everyone buys insurance. But starting from the equilibrium in which no one buys insurance, no single individual, acting alone, can coordinate everyone else's choices to reach the preferred outcome.

Infrastructure

Gasoline and insurance are products that are complementary to automobiles. There is another significant complementary product—the roads on which you can drive your car. Without roads, cars have limited value. The same argument applies to bridges and highways and even to the police who enforce the laws of the road. These various kinds of infrastructure serve to increase the value of a car.

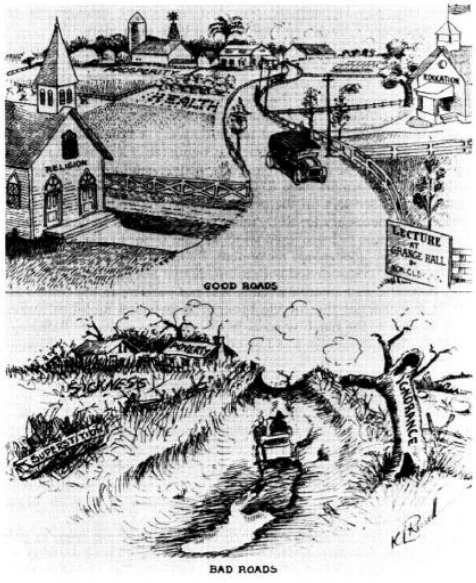


Fig. 1—Idealized effect of good roads on the American rural landscape. *Source: Better Roads and Streets*, Vol. 5, December, 1915, p. 9.

Figure 17.2.4: Idealized Effect of Good Roads on the American Rural Landscape. Source: *Better Roads and Streets, Vol. 5, December 1915, p. 9.*

This cartoon, which is taken from an article on the history of automobiles in the early 20th century, illustrates the link between car demand and roads.The cartoon comes from Peter Hugill, “Good Roads and the Automobile in the United States 1880–1929,” *Geographical Review*, July 1982, 327–49. The value of a car is much higher in the setting labeled “good roads” compared to that labeled “bad roads.” In the developed countries of the world, we now take good roads as a given, but that was not the case at the start of the automobile industry. The evolution of roads was directly linked to the spread of automobiles as a form of transportation. There was a seasonal aspect to this as well. After a long hard winter, the roads were not ready for use, and additional maintenance was needed to put them back into shape for drivers.

The infrastructure of an economy is a special type of good, called a **public good**. A public good has two characteristics:

1. **Nonrivalry.** Public goods are **nonrival**. A good is nonrival if one person’s consumption of that good does not prevent others from consuming the good as well. In other words, if you supply a nonrival good to one person, you can supply it to everyone.
2. **Nonexcludability.** Public goods are nonexcludable. We discussed this idea in Chapter 14. A good is nonexcludable if it is not possible to selectively exclude people from access to the good. In other words, if you supply a nonexcludable good to one person, you *must* supply it to everyone.

There are many examples of public goods, such as roads, bridges, highways, police services, national defense, and lighthouses. Because public goods are nonexcludable, it is difficult for a private firm to produce them. After all, it is hard to expect someone to pay a positive price for a good if she can always get it for free. Instead, governments generally provide these goods. In the United States, local or state governments may provide roads and bridges, while the federal government is the provider of the highway system. Because these goods are not privately produced and traded in the economy, we cannot rely on supply and demand to determine the quantity of these goods produced in an economy. The quantity of public goods produced is an outcome of a political process.

Suppose the government is thinking about building a bridge. The cost of building the bridge depends on the design and the cost of materials and labor. We take this cost as given. Because everyone in the community can use the bridge, the benefits flow to everyone, not to any particular individual. Suppose the government knows how much each citizen in the community values the bridge: call these the individual valuations. Consider then the following procedure.

1. The government adds together the individual valuations of the bridge. Call the total the *social benefit* from the bridge.
2. The government builds the bridge if the social benefit of the bridge exceeds the cost of the bridge and does not build it otherwise.
3. The government uses its power of taxation to raise revenues to build the bridge.

This is a rule that determines which public goods should be provided. It is a rule that leads to the efficient provision of public goods: those with a large enough social benefit are provided, those with a lower benefit are not. More precisely, by following this rule, it is possible for the government to make sure that public goods are provided only when their provision makes everybody better off. Whenever the rule justifies the building of the bridge, the government can tax each individual an amount that is less than that person’s individual valuation of the bridge and still raise enough money to finance the building of it.

There are two major problems with this scenario. First, we supposed the government knew everybody’s valuation. Obviously, this will not be true in practice. The government could ask people to provide their valuations, but the problem here is that people have no incentive to tell the truth. In particular, if people thought that the amount they would be taxed was related to their valuation, then they would have an incentive to understate their valuations. Second, even if people truthfully revealed their valuations, the tax scheme might be perceived as very unfair because different people would be taxed different amounts.

To see in more detail how this incentive problem arises, imagine a different rule. Individuals in the community individually decide how much to contribute to the construction of the bridge.Marco Haan and Peter Kooreman, “Free Riding and the Provision of Candy Bars,” *Journal of Public Economics*, February 2002, 277–92. The bridge is built if the sum of everyone’s contributions exceeds the cost. Because the bridge is a public good, each resident enjoys the benefit of the total bridge, not only the segment built by their individual contribution. You benefit from the contributions of others, and they benefit from your contribution. If everyone independently decided how much to contribute to the construction of the road, they would be unlikely to contribute at all. If others are not contributing, then there is no reason to contribute because the bridge will not be built anyway. And if others are contributing enough to finance the bridge, then you can benefit without having to pay.To be more precise, you would contribute only in the unlikely event that the amount you are willing to pay would make the difference between the bridge being built or not. We expect that the contributions will, from a social perspective, be small or zero because each member of the community ignores the benefit of his or her contribution to others.

More generally, there will be underprovision of public goods because individuals do not take into account the effects of their contributions on others' well-being. This is sometimes called a "free-rider problem." The term comes from the fact that if everyone else pays for the good, you can travel for free.

Key Takeaways

- The demand for cars is an example of unit demand.
- A car is an asset because it is a durable good that can be resold.
- Gasoline and insurance are two important complementary products to cars. Infrastructure, such as roads and bridges, is also a complementary product to cars. Because such infrastructure often has the characteristics of a public good, the government often provides it.

Exercises

1. Suppose there is a temporary shortfall in refining capacity, so the supply of gasoline decreases for a three-month period. Use a supply-and-demand diagram to show what happens to the price of gasoline and to the quantity traded. What implications will this have for (a) how much people drive and (b) the demand for automobiles?
2. If everyone is required by law to have car insurance, what will happen to the demand for cars?

This page titled 17.2: The Demand for Automobiles is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

17.3: Supply of Cars 17.3: Supply of CarsEdit Update titlecancel

Learning Objectives

1. What factors determine the price of a car?
2. How do car manufacturers compete beyond their choice of price?
3. What factors influence the choices that automobile producers make about the location of their production?

If you walk around the streets of your town, you could conduct a survey of the cars you see. For each car, you could make your best guess as to the answers to the following questions:

- What was the sticker price of the car when it was first sold?
- How old is the car?
- What is the car's estimated value now?
- What are the car's most important features?
- Which company produced the car?
- Is the car manufacturer a US company or a foreign company?
- In which country was the car assembled?
- At which manufacturing plant was the car assembled?

We start with the price of a car. We then look at other aspects of the production decision, such as the key attributes of the car and the choice of production location.

Car Prices

The basic rule for pricing is as follows: set the price so that

$$marginal\ cost = marginal\ revenue.$$

This rule was explained and developed in Chapter 7. **Marginal cost** is the extra cost incurred by producing an additional unit, and **marginal revenue** is the extra revenue earned by producing an additional unit. To understand how this rule applies to cars, we need to look more carefully at both the costs of production and the demand for cars.

Marginal Cost

Cars are produced in automobile assembly plants using a variety of inputs, such as steel, rubber, glass, and labor. Lying behind the assembly of the car is an organization that engineered the car and designed the production process. At one level, there is nothing special about the cost structure for car production. We can decompose costs into three components: **entry costs**, **fixed operating costs**, and **variable costs**. We explained these notions of cost in Chapter 9.

- Entry costs are incurred prior to production. For cars, the most significant entry costs are design costs and the costs of establishing the production line. Many of these expenses are incurred in the research and development stage of the product. Once the car has been designed, the production line must be organized to manufacture the car efficiently. Finally, specialized machinery must be ordered or custom built for the production line. All of these expenses are incurred before a single car can be produced.
- Fixed operating costs include the costs of managing the automobile plant. Think of these as the costs of the various divisions of the plant not directly engaged in production: the operations department, the human resources department, and so on. These costs are the same no matter how many automobiles are being produced.
- Variable costs depend on the number of cars being produced. To build more cars, the firm must hire more labor. If the firm wants to produce fewer cars, it needs to buy less steel. These and other variable costs fluctuate according to the number of cars rolling off the production line.

By definition, entry costs and fixed operating costs are the same no matter how many cars are produced. The only costs that matter for the pricing decision are the firm's variable costs. Managers in auto plants must do their best to determine how much these variable costs change when they produce one extra vehicle. In other words, they need an estimate of the marginal cost of production.

> **Toolkit: Section 31.14, and Section 31.15**
>
> You can review different cost definitions and the definition of marginal revenue in the toolkit.

The history of automobile manufacture reveals that costs of production change over time. Technological progress is visible as we compare production processes at different dates. Ford's move to mass production was key to its success in the early 1900s because this new production method reduced costs substantially. Meanwhile, modern, highly automated, capital-intensive production facilities make those Ford production techniques seem primitive.

Even today, however, the labor input into the production process differs across producers. A recent report compared the labor hours required to produce a car at different manufacturing facilities.See Gary S. Vasilash, "Assembly Plants: How They Compare," *Automotive Manufacturing & Production*, August 1997, accessed March 14, 2011, http://findarticles.com/p/articles/mi_m0FWH/is_n8_v109/ai_20855370/pg_2. For 2006, a Nissan plant in Smyrna, Tennessee, required 28.32 labor hours to produce one vehicle. A Toyota plant was next at 29.54 hours. In contrast, a General Motors (GM) car required 44.59 hours of labor input. Thus GM is using a much more labor-intensive method of production than Toyota or Nissan, whose facilities are more automated. These are not exactly measures of marginal cost because they measure average labor hours rather than the labor hours required to produce one extra car. Still, it is very likely that the GM plant has a higher marginal cost than the Nissan plant.

Marginal Revenue

We can take marginal revenue = marginal cost and rewrite it as a markup pricing formula:

$$price = (1 + markup) \times marginal\ cost.$$

For example, if the marginal cost of producing the last unit is $30,000 and the markup is 0.50 (50 percent), then the firm sets a price of $45,000. For a given value of marginal cost, a higher markup translates into a higher price. And for a given markup, higher marginal cost translates into a higher price. The **markup** depends on the **own-price elasticity of demand**.

Suppose a firm has a lot of market power. This means it can increase its price with relatively small changes in the quantity demanded: that is, demand is inelastic so −(elasticity of demand) is small. In this case, a firm will choose a large markup. If demand is more elastic, a firm will choose a smaller markup.

> **Toolkit: Section 31.2, and Section 31.15**
>
> You can review the definition and measurement of own-price elasticity of demand and markup pricing in the toolkit.

The markup pricing equations seem easy to implement, at least in principle. For an automobile producer, pricing is actually quite complex. There are several reasons for this:

- With something as complicated as an automobile, the calculation of marginal cost is not straightforward. The operations department must determine how much additional labor and raw materials are needed to produce exactly one more car.
- Extensive market research and a certain amount of careful experimentation may be required to find the elasticity of demand. Remember that we saw that a household's decision about whether to buy a car depends on many different factors, including income, interest rates, the current and expected price of gasoline, the price of public transportation, and so on.
- The price elasticity of demand depends on the decisions of other producers. We discussed this idea in Chapter 15. When there is a relatively small number of suppliers, firms have to keep a close eye on the strategies of their competitors. For example, if other producers increase the prices of their cars, then you can expect households to substitute toward purchasing your car. The demand curve for your car will shift to the right. This is good news. But to determine whether to change your own price in response, you have to determine the elasticity of the demand curve for your product, given the new prices set by the other producers. You also have to worry about whether changes in your price will in turn lead other producers to change their prices again.
- Automobile producers manufacture many different models of vehicles. In effect, they compete with themselves as well as with other producers. If Ford cuts the price of, say, a Thunderbird, then the demand curve for other Ford vehicles will shift to the left as at least some potential buyers now choose the Thunderbird ahead of other Ford products. Ford must take into account how its various pricing decisions interact.

- As we explained earlier, an automobile can be thought of as a “bundle of attributes,” such as performance, style, color, and so on. The valuation that a potential buyer places on a car depends on the buyer’s valuations of these various attributes. Thus when manufacturers want to assess how much a car is worth to potential buyers, they really need to determine how much each attribute might be worth. Sophisticated statistical techniques are used to develop these numbers, and this information is used in both the pricing of vehicles and the decisions about which attributes to include in new models, which to exclude, and which to have as available options.

Pricing is only one of many decisions made by car producers. They make other key choices as well. Two of the most significant are design changes when they introduce new models and the decision about where to locate their production facilities. We turn to these next.

Model Introductions

A century has passed since Henry Ford introduced one of the most famous automobiles ever: the Ford Model T.For details on the history of the Model T, see The Henry Ford Museum website, “The Model T,” accessed March 2011, www.thehenryford.org/exhibits/showroom/1908/model.t.html. This car remained in production for almost two decades, with 15 million automobiles produced. There were two versions of the Model T: a car and a truck. Otherwise, there were very few changes made to the vehicle design throughout its years of production. Famously, Henry Ford is claimed to have said, “You can paint it any color, so long as it’s black.”

In July 2008, 59 different vehicles were listed on the Ford website, including an entire family of brands: Ford, Land Rover, Lincoln, Mercury, Mazda, and Jaguar. In other words, Ford produced an immense variety of vehicles—available in more than one color. The same is true of other automobile producers. And, of course, such product variety means more than just a large number of models: any particular model may be available with all sorts of different styling, performance, and features. Interestingly, a visit to the Ford website in 2011 yields a different picture. There are Ford vehicles available, but the other brands are gone. Both Jaguar and Land Rover were sold by Ford in 2008, partly in response to the financial crisis. Over time, companies decide both to introduce and to remove models from their range of offerings.

Cars are not the only products that display such diversity. You can buy many different kinds of laptop computers, breakfast cereals, or mobile phones, for example. As economies grow and develop, we typically see an increasing variety of goods available. But product variety is particularly noticeable with cars because automobile producers come out with new models each year.

New model introductions began early in the history of the automobile. In the 1920s, Ford faced stiff competition from other producers, particularly GM. In the mid-1920s, under the leadership of company president Alfred Sloan, GM had adopted a strategy of introducing new models.This discussion draws on the history of automobiles at David Gartman, “Tough Guys and Pretty Boys The Cultural Antagonisms of Engineering and Aesthetics in Automotive History,” Automobile in American Life and Society, accessed March 14, 2011, http://www.autolife.umd.umich.edu/Design/Gartman/D_Casestudy/D_Casestudy3.htm. In part, the strategy came from recognizing that automobiles were durable goods that households kept for many years. The introduction of new models was a strategy to motivate the exchange of old for new cars. This strategy worked. Ford’s sales of the Model T fell off and, at the end of the decade, Ford also adopted the strategy of model turnover.

The tactic remains in place today. Each year, car companies introduce new models. In some years, they make radical changes, while in other years new cars do not deviate much from previous models. The design and production of new models is one element of the competition among automobile producers. Although we often emphasize price competition, producers also compete in terms of the attributes of their models. Thus competition is very complex.

Plant Location

You have probably given little thought to why firms build factories in one location rather than another. But imagine for a moment that you must decide where to construct a new automobile plant. What kinds of factors might influence your decision?

You would certainly think about the cost of your inputs—that is, the items you need to manufacture new vehicles. Cars require substantial amounts of raw materials, such as steel, that have to be brought to your factory. If those inputs have to be brought in from a long way away, then your inputs will be more expensive. These costs depend also on the local infrastructure: are there good road and rail links to your prospective site? Another input, of course, is labor. Ideally, you want to locate your factory where labor is cheap but also sufficiently skilled for the positions you need to fill.

Once you have manufactured the cars, you have to get them to their final destinations: dealers throughout the country or even throughout the world. Because cars are large and heavy, they are expensive to ship to other locations. Thus, other things being equal, you would also like to locate your manufacturing site near your final demand. Of course, producers must usually serve many markets from a single plant.

Where you ultimately choose to locate the plant will depend on the costs of transporting both inputs and output. If your inputs are very costly to transport, then you will produce near the source of inputs and ship your finished goods to your markets. Alternatively, if your inputs are easy to transport but your output is costly to ship, then you might locate your production near some of your markets. You might even consider multiple production plants to lower the costs of transporting the final good.

You also care about local policies, such as the level of taxes. Countries, states, regions, and cities often compete to attract factories. They do so because a factory brings with it jobs and greater prosperity for a region. In some places in the world, you also have to worry about whether your property rights are well protected. If you set up a factory in the United States, you can be reasonably confident that the government will not try to confiscate either your capital or profits. In some other countries, however, you may justifiably be concerned for the safety of your assets.

The automobile industry in the United States was initially located in and around Detroit. This was partly due to the fact that access to the Great Lakes provided low-cost transportation of the necessary inputs into the production process. As time passed, plants began to appear outside the Detroit area, particularly in the southern part of the United States.

One of the factors motivating these location decisions was labor costs. The automobile plants in and around Detroit were dominated by a union, the United Auto Workers (www.uaw.org/node/39), which was formed near the end of the Great Depression. In the short run, firms must negotiate with the unions that represent its workers. In the longer run, though, firms have other options. One of them is to locate plants in areas with cheaper labor costs. Over time, firms have indeed shifted some of their production facilities to other parts of the United States and other countries around the world where labor is cheaper.

Key Takeaways

- The price of a car is a markup over the marginal cost of production. The markup depends on the elasticity of demand.
- Car producers compete by introducing new models.
- Plant location choices are made in an effort to reduce the costs of production as well as the costs of transporting intermediate goods to a plant and finished goods to the market.

Exercises

1. How might you explain the differences in labor input per car across automobile assembly plants?
2. Under what conditions would a car producer locate a production plant in Alaska?
3. What other goods have new models introduced into the market? Does this happen every year? Why do the producers of these goods change models?

This page titled 17.3: Supply of Cars is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

17.4: Market Outcomes in the Automobile Industry 17.4: Market Outcomes in the Automobile IndustryEdit Update titlecancel

Learning Objectives

1. What kind of competition is there in the automobile industry?
2. How do market outcomes differ in the short run compared to the long run?

The interactions among buyers and sellers in the car market ultimately lead to prices and quantities of all the different cars that are produced. But what is the right way to think about that interaction? Automobile markets are not examples of **competitive markets**—many firms each producing an identical product. Nor is there a single car producer acting as a monopolist. To study markets such as the car market, we have an intermediate situation where firms

- sell goods that are imperfect substitutes for other goods in the markets in the short run, and
- enter and exit in response to profit opportunities in the long run.

Competition among Producers in the Short Run

When we think about market outcomes for automobiles, there are two different markets to consider. There are **business-to-business** markets in which manufacturing firms sell cars to dealerships, and there is the **business-to-consumer** market in which dealerships sell cars to the final consumer. This pattern of trade is quite normal: most firms do not sell directly to the final consumer but instead sell their goods through retailers.

So far we have said that automobile producers determine prices for their cars. But the companies do not actually set the price you will ultimately pay for a new vehicle. That price is determined through a bargain between you and a dealer. The price that the company sets is the price at which it sells to the dealer. Given the numerous dealers, you would not expect them to be able to make much profit. Competition will force the price to be close to the cost of the car to the dealer. But the producer retains market power and can dictate a price for selling the car to the dealer.

This might make it tempting to think about the final market for cars as being roughly competitive. After all, one of the conditions for a competitive market is that there should be a large number of buyers and sellers. Another condition, though, is that sellers should be selling identical goods. In the case of cars, this is evidently not the case. We have already pointed out that firms produce many different models of vehicles with various options available. On top of that, dealerships may differ in terms of the quality of service they offer both before and after the sale of a vehicle. Consumers, when choosing which car to buy and where to buy it, are choosing from a large set of different, imperfectly substitutable products. We call these **differentiated products**.

Each dealer therefore possesses a degree of market power. Some of this market power comes from the fact that there will be only a small number of sellers of a particular model in a given region. Some of the dealer's market power stems from specific features of the dealership, such as location and after-sales service. The key point is that each dealership faces a downward-sloping demand curve for the cars that it sells. The seller chooses a point on the demand curve. Because there are competing cars available from other dealerships in the market, the position of the demand curve depends on the prices set by other firms for other models.

Although dealerships possess some market power, the retail market for automobiles is still quite competitive. Demand is relatively elastic because consumers have different dealerships and cars to choose from. In addition, information about the price at which dealers obtain vehicles from manufacturers is readily available. Under most circumstances, therefore, dealers are able to enjoy only a small markup over this price. (The exception is when a particular model of vehicle is in particularly high demand for some reason.)

From a dealer's perspective, marginal cost is determined largely by the price at which it obtains the car from the manufacturer. The producer sets the price to the dealer to maximize its own profit. Producers understand that the demand for their products is affected by the prices of competing vehicles. This strategic interaction means that the elasticity of demand (and hence the markup) for a particular car depends on the prices set by other manufacturers. We explained this in detail in Chapter 15. Likewise, dealers set their prices based in part on the prices at other dealerships. On the demand side, households take the set of products offered in the market and their prices (subject to a little bargaining with dealers) as given as well. Their decisions about which cars to purchase and when to purchase them generate the market demand curves faced by dealerships.

Market Dynamics in the Long Run

So far we have taken as given the types of cars produced, the location of plants, and the identity of the automobile producers. Over a short period of time, such as a year, this is a good way to think about the market for cars. But over longer periods of time, the market is much more dynamic. There are changes in the models of vehicles; there are changes in the location of manufacturing plants; and there is entry and exit of manufacturers. One way to see this is to look at the evolution of the automobile market since the early part of the 20th century.

The beautiful car shown in this picture is called a Marmon.This discussion draws a history of the Marmon that can be found at Bill Vance, "Motoring Memories: Marmon," accessed March 14, 2011, http://www.canadiandriver.com/2000/03/16/motoring-memories-marmon.htm. The photo is of a 1932 model. A Marmon won the first Indianapolis 500, and nearly 22,000 models were sold in 1929. But by 1934, the company was gone, a casualty of the Great Depression. Small fringe producers like Marmon disappeared from the automobile industry. Left behind were the large producers who were to dominate the US automobile industry from that time onward. By the mid-1930s, the US market was largely ruled by three manufacturers.



Figure 17.4.1. A 1932 Marmon.

Economists Tim Bresnahan and Daniel Raff looked at data on automobile plants during this time period. See Timothy Bresnahan and Daniel Raff, "Intra-industry Heterogeneity and the Great Depression: The American Automobile Industry, 1929-1935," *The Journal of Economic History*, June 1991, 317–31. They found that the number of plants (remember that one firm may have multiple plants) that were producing cars fell from 211 in 1929 to 121 in 1935. There is no single explanation of exactly why these producers failed and had to close their plants. The Great Depression evidently led to a large decrease in the demand for automobiles. But on top of that, surviving firms were marked by advances in product and process development. In the early stages of the automobile industry, small producers operated at a small scale. Such producers simply could not compete with Ford's lower-cost production process. This competition from Ford led to the exit of producers of cars like the Marmon. In the end, the industry was left with a small number of powerful firms.

In this market, firms were selling differentiated products, so they had market power. Over the long run, there was entry and exit of competing products (that is, firms introduced new products and retired old ones). There was also entry and exit of entire firms. The conditions governing entry and exit are the same as those that we explained in Chapter 9. A firm will introduce a new product if it expects to make sufficient profits (in terms of discounted present value) to justify the fixed entry costs. A firm will discontinue a product if the discounted present value of profits that it expects from that product is less than the value of the firm's recoverable assets. Similar conditions apply to entire firms in the market.

Over the past 70 or so years, after the shakeout in the 1920s and 1930s, the big three producers have remained the dominant sellers. From that perspective, you might think that there was little entry and exit. However, the story is more complicated. First, the market share of the three main producers declined due to foreign competition. American consumers started buying cars made in Europe, Japan, and elsewhere. Second, the products produced by the firms have evolved considerably over time. This is a very dynamic market in terms of product innovation. Although there may not have been very much entry and exit of firms, there was considerable entry and exit of products. Sometimes, manufacturers retire entire brands, such as the Hummers that General Motors (GM) stopped producing in 2010.

The Used Car Market

When households choose a car, one option is not to purchase a new car but instead to buy a (as the dealers like to put it) "preowned" vehicle. From the perspective of the buyer, there is one critical difference between a new car and a used car. With a

new car, it is relatively easy to make a reasonably good judgment about the attributes of a product, partly from reviews in magazines and on the Internet. With a used car, it is much harder to judge the quality of the product and thus place an accurate valuation on it. We explained a similar problem in terms of health care in Chapter 16.

With new cars, you bear only a small risk that the car will not perform properly when you buy them. This is not the case in the market for used cars. Imagine (or perhaps you have actually experienced this) going to a used car lot to look for a car. Here is what you might hear from a member of the sales force: “This is the best used car I have ever seen. No lie—it was purchased new by an elderly woman a few years back, and she treated it like one of her kids. It is only here on our lot because she has decided to stop driving. At this price, it is a steal.” You are much less likely to hear this: “Yeah, that car is a lemon. Some guy bought it from the dealer a few months back, and it never was right. One problem after another; it was back in the shop every week. Sure there is low mileage, but my guess is that there are no more miles from that car anyway. Go ahead, buy it if you like. But don’t say I didn’t warn you.”

When you see a used car for sale, ask yourself: why is that car here? The true answer could be one of these two stories. If it is the first situation, then the car is probably a good buy. But if it is the second, then you could be getting ripped off. And the problem is that the seller may give you the first story even when the second is the truth.

The fundamental difficulty here is that you and the seller have very different information. The seller of the product knows its quality (is the car good or bad?) while you, as a buyer, do not know its quality. This does not mean you should never buy a used car. But it does mean that your willingness to pay for a used car should reflect the uncertainty you face with regard to the quality of that car. Because all buyers face the same problem, the end result is that the market valuation of used cars will be low. Accordingly, the price of a used car is lower than it might be if the quality of cars was known. And this can also mean that there are fewer good used cars on the market. This is the problem that economists call **adverse selection**.

You can also perhaps spare some sympathy for the used car dealer as well. We have described this problem from the perspective of a buyer. Even if a dealer really does have a car that is of high quality, it is hard for him to convince prospective buyers of that fact. If you want to sell a car you own, you will probably encounter this problem: you may know that your car is high quality, but you cannot convince buyers.

Key Takeaways

- Car companies compete in markets where they sell differentiated products.
- In the long run, the entry of competitors (in the form of either new firms or new products) continues until profits are equal to zero.

Exercises

1. Use the condition that marginal revenue = marginal cost (consult the toolkit if needed) to explain the difference in the price of two cars of your choice.
2. Some used car sellers include a warranty with your purchase. Would that help overcome the lemons problem?
3. If two cars are close substitutes, what do you predict about their prices?

This page titled 17.4: Market Outcomes in the Automobile Industry is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

17.5: Policy Issues 17.5: Policy IssuesEdit Update titlecancel

Learning Objectives

1. What are some policies to control the emissions from cars?
2. How do trade policy and international investment affect the car market?
3. What is the problem of congestion, and what can governments do about it?

There are many ways in which government policies impinge on automobiles. Here we highlight a few such issues.

Environmental and Resource Concerns

In Chapter 2, we showed a photograph of smog in Mexico City. At the same time that cars have transformed the economic world, they have also transformed our natural environment. The exhaust from cars contributes to air pollution, which is hazardous to health. Car exhausts are a source of greenhouse gas emissions and thus contribute to climate change.

Pollution from cars is a classic example of an **externality**. (We discussed externalities in detail in Chapter 14.) An individual's decision to purchase and drive a car does not take into account the effects on third parties. In this case, some of the affected third parties are those in the immediate vicinity who suffer from a reduction in air quality. To the extent that emissions contribute to climate change, however, the third parties potentially include everyone in the world.

Toolkit: Section 31.19

You can review the definition and use of externalities in the toolkit.

Governments in the United States and elsewhere have enacted various policies that are motivated, at least in part, by the desire to take into account such environmental externalities and resource use. First, there are taxes on gasoline. These are relatively low in the United States but are much higher in Europe. Second, there are technological restrictions, such as the requirement that automobiles be fitted with catalytic converters and designed to run on unleaded fuel. In the United States, the government has taken action to improve the fuel consumption of cars produced within US borders. These are called Corporate Average Fuel Economy (CAFE) standards.CAFE standards are described in detail at the National Highway Traffic Safety Administration site, "CAFE Overview—Frequently Asked Questions," accessed March 14, 2011, www.nhtsa.gov/CARS/rules/CAFE/overview.htm. You can notice two things from this term: (1) the restrictions are in terms of fuel economy (miles per gallon), and (2) the restriction does not apply to individual automobiles but rather to the set of cars sold by a corporation. For example, the standard is applied to the entire set of models produced by General Motors (GM), not model by model, so GM makes some cars that are below the standard and others that are above. Corporations that do not meet the standard are fined.

The CAFE standard comes from legislation passed in 1975 in response to the embargo by oil-producing countries in 1973. The initial motivation was to reduce energy consumption and, in part, make the United States less dependent on imported oil. The arguments today for these standards also include the effect of car emissions on global warming.

Trade and Investment Policies

A second government policy that has had a huge impact on the automobile industry is the opening of the world economy to trade and international investment. The current automobile market is no longer just a US market. The United States is part of the world market. US producers interact with the rest of the world by

- selling cars in many countries,
- buying parts from suppliers throughout the world,
- producing in many countries,
- being financed by debt and equity held in foreign countries.

Meanwhile, US citizens

- own cars produced in other countries and imported into the United States,
- consume imported oil,
- work for foreign companies that produce cars in the United States,

- work at car production facilities in other countries.

For example, let us look at Ford Motor Company. In 2007, Ford had 95 plants worldwide and employed about 246,000 people. The Ford operations in North America (United States, Canada, and Mexico) had 94,000 employees. In other words, 62 percent of the workforce was employed outside North America. There are Ford plants all over the world. Ford's 2009 annual report tells us that Ford sold 4.82 million cars in 2009. Of these, 2.0 million were sold in North America, 1.6 million in Europe, and the remainder in South America and Asia.Ford Motor Company, "Annual Reports," accessed March 14, 2011, corporate.ford.com/microsites/annual-reports.

This international structure permits diversification. Ford produces and sells cars in China, South America, and elsewhere around the world. If you browse Ford's global activities,Ford Motor Company, "About Ford: Global Vehicles Sites," accessed March 14, 2011, corporate.ford.com/about-ford/global-vehicles-sites. you will get a sense of its worldwide sales and production operations.Ford Motor Company, "About Ford: Global Operations," accessed March 14, 2011, corporate.ford.com/about-ford/global-operations.

Each producer of cars has its own story of expansion across international borders, both through trade and through production. Honda began operations in the United States by creating a motorcycle sales division in the late 1950s. This eventually led to the production of motorcycles in the United States in 1978 and ultimately the production and sales of Honda cars in the United States.Honda Worldwide, "Establishing American Honda Motor Co. (1959)," accessed March 14, 2011, http://world.honda.com/history/challenge/1959establishingamericanhonda/index.html.

None of this would be possible without governments permitting the movement of goods and capital.

The first trade policy action directly impacting car production was the Canadian-US Automotive Products Trade Agreement of 1965. The goal of this agreement was to create an integrated market for cars between the United States and Canada by eliminating tariffs. Concerns that US companies would sell but not produce cars in Canada were met by some restrictions on production, including requirements that cars built in Canada had to have a certain domestic content.

The second trade policy action was called the North American Free Trade Agreement (NAFTA).US Department of Agriculture, Foreign Agricultural Service, "North American Free Trade Agreement (NAFTA)," accessed March 14, 2011, www.fas.usda.gov/itp/Policy/NAFTA/nafta.asp. NAFTA was a controversial trade agreement. One of the big issues was whether the reduction in trade barriers would lead to job destruction in the United States. (We discussed this in Chapter 9.) A 2001 study looking back at the effects of NAFTA directly on the production of cars did not find large effects at all.

Most fears about the ill effects of NAFTA on the U.S. auto industry, whether in term of employment, wages, or investment, have been proven wrong. The U.S. auto industry did experience rationalization of production and hence job displacements. But overall, NAFTA appears to have helped the U.S. auto sector (U.S. Trade Representative, 1997). Employment in the American automotive industry grew by 14.1 percent overall, with an increase of 16.1 percent in the auto parts sector and 10.1 percent in the motor vehicle assembly sector from 1994 to 1996. Hourly earnings for production workers in the U.S. automotive sector grew by 5.6 percent between 1993 and 1996. The Big Three U.S. automobile manufacturers invested $39.1 billion from 1993 to 1996 in new manufacturing plants and equipment in the United States, while investing only $3 billion in Mexico over the same period.Mary E. Burfisher, Sherman Robinson, and Karen Thierfelder, "The Impact of NAFTA on the United States," *Journal of Economic Perspectives* 15 (Winter 2001): 125–144.

These statistics, of course, refer to what actually happened in the auto sector over this time period. What *would* have happened had NAFTA not been implemented requires a more sophisticated analysis.

Congestion

If you travel to Mexico City or Manchester, Beijing or Buenos Aires, Jakarta or Johannesburg, Los Angeles or Lagos, you will see that these cities all have something in common: traffic jams. Such road congestion is another example of an externality. The decision of one person to drive has an effect on other drivers.

One way of solving externality problems is to create new markets. In most cases, there is no market for the use of roads. However, if we charge people to use roads, then market incentives come into play. Toll roads are an example of the introduction of a market mechanism to combat congestion problems.

Congestion fees and tolls are in use in some cities around the world, such as London and Singapore.This June 21, 2006, press release from the UK Commission for Integrated Transport provides some discussion of the London system and others around the

world: “New study shows road pricing progress,” accessed March 14, 2011, http://cfit.independent.gov.uk/pn/060621/index.htm. The system in London, started in February 2003, charges drivers for entering the central city area between certain hours. Details of the system are available from Transport for London.Transport for London, “Congestion Charging,” accessed March 14, 2011, http://www.tfl.gov.uk/roadusers/congestioncharging/default.aspx. The cost in May 2008 was £8 (about $15.60) for access to the charging zone in Central London between 7 a.m. and 6 p.m., Monday through Friday. The system is enforced by a series of cameras that record license numbers and then check them against a record of who has paid for access to the zone. According to the Transport for London, the traffic flow into the zone has been reduced by 21 percent, and there is now less pollution and more cycling in the area.

The Electronic Road Pricing (ERP) system in Singapore, although older, is much more sophisticated. It was introduced in April 1998 along Singapore’s expressways and in the city’s central business district. All vehicles contain a transponder, mounted on the windscreen, into which the driver inserts a prepaid cash card. There are gantries located at various points around the city, and whenever a car passes under a gantry, a toll is automatically deducted. The rates differ for different categories of vehicle: motorcycles and light goods vehicles pay less than cars; heavy goods vehicles pay more than cars.

The most striking feature of the Singapore system is that the charges vary by time of day. Charges are imposed only at the peak hours, and the charges vary within those hours. Thus, for example, a driver passing a typical gantry might pay SGD 0.80 (about $0.58) from 08:00 to 08:05, SGD 1.50 from 8:05 to 8:30, SGD 2.00 from 08:30 to 09:00, SGD 1.50 from 09:00 to 09:25, SGD 1.00 from 09:25 to 09:30, and so on. You can see that these rates are quite finely tuned, with some rates being in effect for only a five-minute period.

The rates just quoted were in effect in mid-2008. By now, they may be quite different because a second feature of the system is that these rates are revised frequently. The Singapore Land Transport Authority has targets for the desired average speed of traffic on Singaporean roads: the target speed for expressways is 45–65 kilometers per hour (28–41 miles per hour), and the target speed for arterial roads is 20–30 kilometers per hour (13–19 miles per hour). Thus if they observe that traffic is flowing below these speeds, they consider raising the rates; if traffic is flowing smoothly, they consider reducing rates. They also adjust rates on a seasonal basis —for example, ERP charges are lower during school vacations.

Key Takeaways

- Gas taxes and actions to improve fuel efficiency of cars are policies that reduce pollution from cars.
- In the United States, households benefit from the importation of foreign-produced cars and also from the ability to work at automobile factories owned by foreign companies.
- The opening of the car market to imports creates some job displacements.
- In some countries, governments tax the use of roads when they are congested.

Exercises

1. Give two reasons why the government taxes gasoline.
2. Why might a government choose to limit car access to a city center? What policies are available to a government for doing that?

This page titled 17.5: Policy Issues is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

17.6: End-of-Chapter Material 17.6: End-of-Chapter MaterialEdit Update titlecancel

In Conclusion

When you study economics from a textbook such as this, you learn different economic theories. In this book, we have looked at the theory of the consumer, which helps explain how individuals make their choices about what goods to buy. We have looked at the theory of the firm, which explains how firms make decisions about which goods to produce and what price to sell them at.

Our goal in this book is to help you see that economics is not only a matter of graphs and definitions but also the study of the world around you. Economists see economic decisions and economic forces everywhere they look. This chapter gave you many examples linked to one particular and very familiar product. Yet we have only scratched the surface in terms of the ways in which we could apply economic analysis to cars and the car industry.

Perhaps you can think of other ways in which you could apply the things you learned from your study of economics to the market for automobiles. In any case, we hope that, now, every time you see a car, you will remember that you are also seeing economics in action.

Key Links

- GM: http://www.gm.com/corporate/about/history
- Rouge River Plant: http://www.hfmgv.org/rouge/historyofrouge.aspx
- History of automobiles from the Smithsonian Institution: http://www.si.edu/Encyclopedia_SI/nmah/autohist.htm
- Bureau of Labor Statistics, motor vehicle and parts manufacturing: www.bls.gov/oco/cg/cgs012.htm
- BMW plant in South Carolina: http://www.bmwusfactory.com
- Electronic Road Pricing in Singapore: www.lta.gov.sg/motoring_matters/index_motoring_erp.htm

exercises

1. (Advanced) A car is an asset, like a house. How would you use the principles of asset valuation to ascertain the value of a car? (Hint: see Chapter 10.)
2. List the products and/or services that are substitutes for cars.
3. Because a household can delay the purchase of a new car when the price increases, what does this do to the price elasticity of demand for cars?
4. List the car characteristics and household characteristics that would increase a household's valuation of a car.
5. Would an increase in the price of gasoline decrease the demand for all cars or only some? What are the effects of an increase in gas prices on alternative forms of transportation?
6. In the coordination game for automobile insurance, is it desirable for the government to require that everyone buy automobile insurance?
7. There are three people in a community. The government is proposing to build a bridge in that community. The bridge costs $120 to construct. Everyone values the bridge at $60. Is it efficient to build the bridge? Suppose the government asks people what the bridge is worth to them and plans to tax them the amount that they say. If two people truthfully reveal to the government that the bridge is worth $60 to them, will the third person give his true valuation as well?
8. Using the equation linking the price of a car to markup and marginal cost, if the markup is 60 percent and marginal cost is $20,000, what is the price? Show with this example that a firm facing a more elastic demand curve will set a lower price.
9. Is a car dealership more valuable in a small city with little competition or in a big city with more competition but also more potential buyers?
10. Car dealers make profits from goods and services that are complementary to the cars they sell. List and discuss some of these complementary products.
11. Why do automobile companies offer warranties with new cars? Why do these warranties expire after a few years?
12. What factors determine the choice of an automobile producer to create a new model?
13. If you were a seller of a used car and you knew it was high quality, would you have an incentive to offer the buyer a warranty on the car? Would you have that same incentive if you knew the car was low quality? What then should you as a buyer infer if you see a seller willing to offer a used car with a warranty?
14. One concern with opening markets in the United States is that domestic car producers are forced to compete with foreign producers. In some cases, this competition leads to job losses in the United States. Who gains from this competition?

15. (Advanced) If you were designing a policy of charging for road use, would you include or exempt taxis?
16. In what sense is buying a car like buying health care? What do car companies do to provide buyers with assurance over the functioning of the car? Why doesn't this happen in the health-care market?
17. If you look at the market for cars, high-income households are more likely to buy new cars, and lower income households are more likely to buy used cars. Why is that the case?
18. Following a recession, it is likely that the average age of cars will be higher than prior to the recession. Why?
19. Would you say that the market for higher education is a market with monopolistic competition or a monopoly? How does competition occur in that market?

Economics Detective

1. In addition to car production being shifted to the southern part of the United States to reduce labor costs, car production has also moved overseas. What information can you find on the wages of workers at automobile plants outside the United States?
2. What is the current penalty on companies that do not meet the CAFE standard?
3. In June 2009, the US government became a majority owner of GM. What prompted that action? Why did the US government not buy Ford Motor Company? What has happened to US government ownership of GM? What has happened to the sales and profitability of Ford and GM in recent years?
4. What products other than cars have a "model year"? Are new models introduced simultaneously or at different times of the year? What factors might determine the timing of new model introductions?

Spreadsheet Exercise

1. Based on the equations for markup pricing, create a spreadsheet to calculate the price you would set as a car producer. To do this, input the elasticity of demand and marginal cost. Use your program to calculate the markup and product price. Produce graphs to show how the product price will change as the elasticity of demand changes.

This page titled 17.6: End-of-Chapter Material is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

18: The State of the Economy

18.1: The IMF Comes to Town
18.2: Measuring Economic Activity
18.3: Measuring Prices and Inflation
18.4: The Circular Flow of Income
18.5: The Meaning of Real GDP
18.6: End-of-Chapter Material

This page titled 18: The State of the Economy is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

18.1: The IMF Comes to Town 18.1: The IMF Comes to TownEdit Update titlecancel

In early 2002, a team from the International Monetary Fund (IMF) flew to Buenos Aires, Argentina. Argentina had been prospering during most of the 1990s, but more recently it had begun to run into economic problems. The IMF is an organization that attempts to help countries having financial difficulties.

An IMF team consists of professionally trained economists. These teams visit many countries, such as Argentina, on a regular basis. In this chapter, we imagine that the IMF added you to this mission and asked *you* to report back on the state of the Argentine economy. As we proceed, we think about how you might have approached this task.

You arrive at Aeropuerto Internacional Ministro Pistarini de Ezeiza Airport, which is a clean and modern airport on the outskirts of Buenos Aires. You ride into the city in a new car along modern highways lined with fancy billboards. When you get to the city center, you notice that there are luxurious shopping malls. You see high-end stores selling luxury brands, such as Louis Vuitton, Versace, Hermes, and Christian Dior. The city seems prosperous, reminiscent of Paris or New York. Just looking around, you see immediately that you are not in one of the really poor countries of the world.



Figure 18.1.1. Source: Image taken by authors.

As you explore the city, though, you begin to look more closely and notice that things are not quite what they seemed at first glance. The luxury stores do not have many customers in them. Some buildings show signs of a lack of maintenance; it has been a while since they were repainted. Some stores are boarded up or bear signs saying that they are going out of business. There seem to be a lot of people who are not working or who are making a living selling goods on the street.

Reflecting on these conflicting clues to Argentina's prosperity, you quickly realize that it is difficult to assess the health of an economy by casual observation. In addition, you have seen almost nothing of the country. Argentina covers over one million square miles; it is almost one-third of the size of the United States and has a population of nearly 40 million. The more you think about this, the harder the problem seems. Forty million people are buying things, selling things, making things, and consuming things every day. It seems an impossible task to make sense of all this activity and say anything useful about the economy as a whole. That challenge is the subject of this chapter.

How can we evaluate the overall performance of something as complicated as an economy?

Road Map

If you think about this question for a bit, you will realize that it has more than one dimension.

- First, we need *measurement*. We must summarize the economy in a manageable way, which is impossible unless we find some way of measuring what is going on in the economy. One of the primary tasks of economics is accounting. That leads to other questions: what should we count, and how should we count it?
- Data are not enough. Measurement will not take us very far unless we can combine it with some understanding of how the economy works. We need to know how to interpret the things we count. We need to know what our numbers mean. For this, we need *frameworks* that help us make sense of the economy.

These two ideas guide our discussion in this chapter.

Think for a moment in very general terms about what happens in an economy. An economy possesses some resources. These include the time and abilities of the people who live in the economy, as well as natural resources, such as land or mineral deposits. An economy also possesses various means of changing, or transforming, one set of things into other things (see the following figure). For example, we have a process for making tea. We produce tea by taking cold water, energy, and dried leaves and transforming those inputs into a hot beverage that people like to drink. The simple act of making a cup of tea is an example of production.

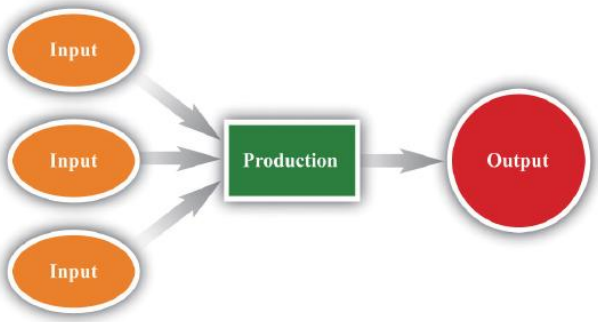


Figure 18.1.2: From Inputs to Output. One of the main economic activities is production: the transformation of inputs (raw materials, labor time, etc.) into output (goods and services that people value).

We are interested in measuring how much production occurs in an economy. Obviously, however, we cannot hope to count all the times that people drop a teabag into a cup, and it would not make much sense to do so. *Economic* activity typically involves more than production; it also includes the notion of exchange—buying and selling. If you make a cup of tea for yourself at home, we do not think of this as economic activity. If you buy a cup of tea at your local coffee shop, we *do* think of this as economic activity. A very rough definition of economic activity is as follows.

Economic activity is the production of goods and services for sale.

Any definition this straightforward is bound to be too simple, and we will see that there are several subtleties in the actual measurement of economic activity, particularly since some goods and services are not actually bought and sold. Still, if you keep this idea in mind, it will help you as we progress through the basics of economic measurement in this chapter.

This page titled 18.1: The IMF Comes to Town is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

18.2: Measuring Economic Activity 18.2: Measuring Economic ActivityEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What is the measure of total output of an economy?
2. What is the difference between real and nominal gross domestic product (GDP)?

Macroeconomics is data driven. Government statisticians and other organizations gather vast amounts of data on the performance of various aspects of the macroeconomy, and macroeconomists try to make sense of all this information.

If we want to explain economic data, then we first have to get the measurement right, and a big part of this is ensuring that we get the accounting right. To make sure that we do, we begin by constructing simple examples. This is not because a simple example is enough to describe an economy; but because cannot hope to understand the complicated accounting unless we do the simple accounting correctly.

The Pizza Economy

To understand the economic health of Argentina—or any other country—we begin by looking at production in the economy. Let us imagine that Argentina produces a single good—pizza. Each pizza is sold for 10 pesos (which is about US$3.33). To be concrete, suppose that every worker in the economy works in a pizza factory in which (1) each hour worked produces 1 pizza, (2) each worker works 40 hours per week, and (3) each worker works 50 weeks per year. Suppose there are about 15 million workers in the economy.

We measure total economic activity by determining the *total value of the pizzas produced* in this economy. We obtain this by multiplying the previous numbers together. There are

$$40 \ pizzas \ per \ worker \ per \ week,$$

so there are

$$2,000 \ pizzas \ per \ worker \ per \ year \ (= 40 \times 50),$$

which means that there are

$$30,000,000,000 \ pizzas \ per \ year \ (= 40 \times 50 \times 15,000,000).$$

The value of those pizzas is

$$300,000,000,000 \ pesos \ per \ year \ (= 40 \times 50 \times 15,000,000 \times 10).$$

The total value of all the production in the economy is called **nominal gross domestic product (nominal GDP)**. The word *nominal* indicates that something is being measured in terms of money—in this case, Argentine pesos. For this economy, nominal GDP is 300 billion pesos per year.

The economy we have just described is extremely stylized and somewhat dull from a culinary perspective. We begin with such a simple economy because it allows us to understand the basic workings of the economy without getting bogged down in a lot of details. We did, however, choose numbers that are the right order of magnitude for the Argentine economy in 2002: the total number of workers in Argentina in 2002 was about 15 million, and nominal GDP was about 300 billion pesos. In 2010, estimated GDP for Argentina was 1.4 trillion pesos, and the workforce was over 16 million.

Measuring Nominal GDP

We now consider a more formal definition of nominal GDP and go through it term by term.

Nominal GDP is the market value of the final goods and services produced by an economy in a given period of time.

Market Value

Our example pretended that there was only a single good produced in the economy—pizza. In real economies, millions of different goods and services are produced, ranging from cars at an assembly plant to haircuts sold by a local barber. If our goal is to measure the overall output of an economy, we are faced with the problem of how to add together these goods and services. How do you add 60,000 cubic meters of natural gas, 1,000 trucks, and 2,000 head of cattle (to pick just a few examples of goods produced in Argentina)?

We need a common denominator. Economists use the *market value* of the goods and services. This means that the common denominator is dollars in the United States, pesos in Argentina, kroner in Sweden, euros in Portugal, and so on. Nominal GDP equals total output produced in a year, valued at the actual market prices prevailing in that year. We choose market value for two reasons. One is simplicity: data on the market prices of goods and services are relatively easy to come by. The second reason is much more important. Market value tells us how much people are willing to pay for different goods and services, which gives us a measure of the relative value of different commodities. For example, if a new laptop computer costs $2,000 and a new hardcover novel costs $20, then the market is telling us that people are willing to trade off these goods at the rate of 100 novels to 1 laptop. In effect, the market is telling us that the laptop is 100 times more valuable than the novel.We take as given here that the market price —which tells us how much people are willing to spend—is a reasonable measure of the value of a good or a service. More precisely, it measures the value of the good or service "at the margin," meaning it measures the value of having one more unit of the good or the service. Explaining why this is usually a sensible interpretation of the market price (and when it is not) is a topic covered in microeconomics courses.

Let's look at an example of the calculation. Table 18.2.1 "Calculating Nominal GDP" considers a very small economy that produces three goods and services: T-shirts, music downloads, and meals. We show data for two years. To calculate GDP in 2012, we take the market value of the T-shirts ($\$20 \times 10 = \$200$), the market value of the music downloads ($1 × 50 = $50), and the market value of the meals ($\$25 \times 6 = \$150$). Adding these, we discover that nominal GDP is $400:

$$\$20 \times 10) + (\$1 \times 50) + (\$25 \times 6) = \$200 + \$50 + \$150 = \$400$$

Doing the same operations for 2013, we find that nominal GDP is $442:

$$(\$22 \times 12) + (\$0.80 \times 60) + (\$26 \times 5) = \$264 + \$48 + \$130 = \$442$$

We can see that lots of things changed between the two years. The price of T-shirts and meals slightly increased, while music downloads became cheaper. Firms produced more T-shirts and music downloads but fewer meals.

Table 18.2.1: Calculating Nominal GDP

Year	T-shirts		Music Downloads		Meals		Nominal GDP ($)
	Price ($)	Quantity	Price ($)	Quantity	Price ($)	Quantity	
2012	20.00	10	1.00	50	25.00	6	400.00
2013	22.00	12	0.80	60	26.00	5	442.00

On the surface, 2013 appears to have been a good year in this economy. Nominal GDP increased substantially relative to 2012. Dig a little deeper, however, and it is harder to interpret this change. Production increased for some products and decreased for others. Some prices increased, and others decreased. Was 2013 really better than 2012? We come back to this question shortly.

Final Goods and Services

In Table 18.2.1 "Calculating Nominal GDP", we assumed that all of the goods and services purchased were purchased by their *final* users. That is, the T-shirts, music downloads, and meals were all purchased by households for consumption purposes. (Households are not the only group that consumes final goods and services in an economy. Firms, the government, and households in other countries can also be final consumers.) We term these final goods (T-shirts) and final services (music downloads and restaurant meals).

In contrast, **intermediate goods and services** are products such as raw materials and energy that are used—and completely used up —in the production of other goods and services.There are two kinds of goods used in the production of other goods. Intermediate

goods are completely used up as part of the production process. Capital goods—such as factories and machines—are not completely used up but live to produce another day. We discuss capital goods in more detail in Chapter 20. We do not include intermediate goods in GDP. Think about a bottle of wine, for example. It might be bought by a consumer at a wine store, in which case it is counted in GDP. Alternatively, it might be bought by a restaurant to sell with its meals. In this case, the cost of the meal is included in GDP, and *the cost of the wine is already included in the cost of the meal*. The restaurant may have purchased the wine from a supplier, but that purchase is *not* included as part of GDP. If both the sale of wine to the restaurant and the sale of that wine to a customer of the restaurant were counted in GDP, the same bottle of wine would be counted twice. By excluding the sale of intermediate goods in calculating GDP, we avoid such double counting.

Being intermediate is therefore not a feature of the good itself. It depends on how the good is used. Wine sold to a consumer directly is a final good; wine sold to a restaurant is an intermediate good. This fits with the idea that we want GDP to measure goods as they are valued by consumers.

Produced by an Economy

Most of the time when we talk about an economy, we are speaking of a particular country. Thus we talk about US GDP, Argentine GDP, Indian GDP, or Uruguayan GDP. Similarly, most of the statistics that are collected refer to economic activity within a country. The term *economy* can be much more general, though, for it simply means a particular set of households and firms. We can speak of the world economy, the North Dakota economy, the Buenos Aires economy, or even the economy of a street of your hometown. The basic concepts are the same no matter what region we choose to discuss.

Over a Given Period

GDP is measured over a specified period of time. In principle, that time period could be anything—a week, a month, a quarter (three months), or a year. In the United States and many other countries, GDP is *measured* on a quarterly basis. However, it is typically *reported* on an annual basis. In other words, government statisticians might measure GDP for the first three months of 2012 and find that it was $4 trillion. That is, over that three-month period, $4 trillion worth of goods and services was produced. The number would typically be reported as “$16 trillion on an annual basis.”

It does not make any sense to talk about US GDP at the instant the clock strikes noon on February 29, 2012. The amount of GDP produced at any instant of time is, for all intents and purposes, zero. Instead, we think of GDP as a **flow**. We can count the number of pizzas produced only if we specify some interval of time. Other variables can be sensibly measured even at a given instant. For example, we could—in principle at least—count the number of pizza ovens in existence at any given time. The number of pizza ovens at a point in time is an example of a **stock**.

The requirement that we count goods and services produced in a certain period means that we should also ignore the *resale* of goods produced in earlier periods of time. If a construction company builds a new house and sells it to you, the production of that home is counted as part of GDP. By contrast, if you buy a house that is 10 years old, the sale of that house is not counted in GDP. (However, if you employed a real estate company to find the old house for you, payment to that company would be included as part of GDP.) In the same way, if you purchase a used textbook that was produced 3 years ago, that purchase is not counted in GDP.

Nominal GDP in the United States and Argentina

In macroeconomics, our data come to us in the form of *time series*. Time series are a sequence of dated variables: GDP in 2000, GDP in 2001, GDP in 2002, and so on. Usually these data are annual, but they could also be quarterly or monthly (or even daily or hourly). If we go to the *Economic Report of the President* (www.gpoaccess.gov/eop), we can find data for nominal GDP. In the United States, the Bureau of Economic Analysis (BEA; http://www.bea.gov/national/index.htm) in the Department of Commerce is responsible for calculating nominal GDP. Table 18.2.2 "Nominal GDP in the United States, 2000–2010" gives an example of a time series.

Table 18.2.2: Nominal GDP in the United States, 2000–2010

Year	Nominal GDP (Billions of Dollars)
2000	9,951.5
2001	10,286.2
2002	10,642.3

Year	Nominal GDP (Billions of Dollars)
2003	11,142.1
2004	11,867.8
2005	12,638.4
2006	13,398.9
2007	14,061.8
2008	14,369.1
2009	14,119.0
2010	14,660.2

It is often more revealing to show a time series as a picture rather than a list of numbers. Figure 18.2.1 shows the data from Table 18.2.2 "Nominal GDP in the United States, 2000–2010" in a graph. Looking at this figure, we see immediately that the US economy grew over these years. The level of nominal GDP (in billions) was $9.8 trillion in 2000 and $13.2 trillion in 2006.

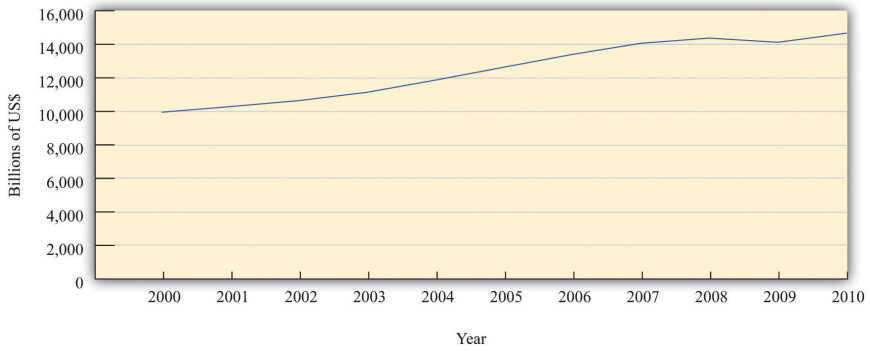


Figure 18.2.1: Nominal GDP in the United States, 2000–2010. Nominal GDP in the United States grew for most of the last decade but declined in 2009.

Source: 2011 Economic Report of the President, accessed July 29, 2011, www.gpoaccess.gov/eop/tables11.html, Table B-1.

Let us return to your International Monetary Fund (IMF) mission in Argentina. From talking to other members of the team, you learn that the Argentine government has statistics on nominal GDP. This is good news, for it means you do have information on the total value of production in the economy. Figure 18.2.2 shows nominal GDP for Argentina over the decade prior to your arrival (1993–2002). In 1993, it was 237 billion pesos. In 2002, it was 313 billion pesos. Thus nominal GDP grew by about one-third over the course of the decade.

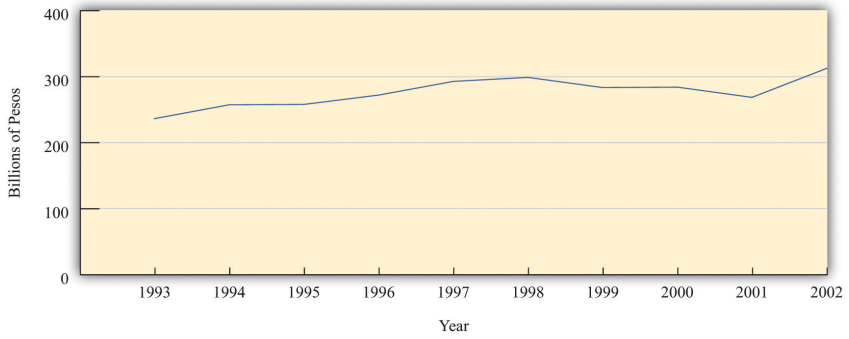


Figure 18.2.2: Nominal GDP in Argentina, 1993–2002. The graph shows nominal GDP in Argentina between 1993 and 2002. Nominal GDP grew overall during this period, although it decreased for several years in the second half of the decade.

Source: International Monetary Fund World Economic Outlook database (
http://www.imf.org/external/pubs/ft/weo/2010/01/index.htm).

Now suppose that in your hotel room one morning you hear on the radio that government statisticians in Argentina forecast that nominal GDP next year will be 300 million pesos greater than this year. How should you interpret this news? Without some context, it is difficult to make any judgment at all.

The first thing to do is to work out if 300 million pesos is a big number or a small number. It certainly *sounds* like a big number or looks like a big number if we write it out in full (300,000,000). If we stacked 300 million peso bills on top of each other, the pile would be over 100 miles high. But the real question is whether this is a big number relative to existing nominal GDP. We have been told that the change in nominal GDP is 300 million, but we would like to know what this is as a **growth rate**, which is a *percentage change*.

Toolkit: Section 31.21

A growth rate is a percentage change in a variable from one year to the next. That is, a growth rate is the change in a variable over time divided by its value in the beginning period.

For example, the growth rate of GDP is calculated as follows:

In our example for Argentina, the percentage change is equal to the change in nominal GDP divided by its initial value. Remember than nominal GDP in 2002 was about 300 billion pesos, so

When we express this change in nominal GDP as a percentage, therefore, we see that it is in fact very small—one-tenth of 1 percent. If you heard on the radio that nominal GDP was expected to grow by 300 million pesos in a 300-billion peso economy, the correct conclusion would be that nominal GDP would hardly change at all. By contrast, if the news announced a projected increase in nominal GDP of 30 billion pesos, the percentage change is 30 billion/300 billion = 0.1 = 10 percent. This is a substantial change in nominal GDP.

Measuring Real GDP

In your bid to understand the economy of Argentina, you have seen that nominal GDP increased by one-third between 1993 and 2002. One possibility is that Argentina is producing one-third more pizzas than it was a decade ago—30 billion pizzas instead of 22.5 billion pizzas. This would be good news. Producing more pizzas is something we would normally think of as a good thing because it means that we are experiencing economic growth: there are more goods and services for people to consume.

In talking to people about the Argentine economy, however, you learn something disconcerting. They tell you that the prices of goods and services are greater this year than they were last year and much greater than they were a decade ago. You begin to wonder: perhaps Argentina is producing no more pizzas than before but instead pizzas have become one-third more expensive than they formerly were. We would typically feel very differently about this outcome. Yet another possibility is that there has been an increase in both the number of pizzas produced *and* the price of pizza, and the combined effect doubled nominal GDP. We need a way of distinguishing among these different possibilities.

Separating Nominal GDP into Price and Output

In our pizza economy, it is easy to tell the difference between an increase in production and an increase in prices. We can measure increased production by counting the number of pizzas, and we can measure increased prices by looking at the price of a pizza. We call the number of pizzas **real gross domestic product (GDP)** (the word *real* here indicates that we are effectively measuring in terms of goods and services rather than dollars), and we call the price of a pizza the **price level** in the economy.

Then it follows that

$$nominal\ GDP\ =\ price\ level \times\ real\ GDP.$$

In our example, the price level is 10 pesos, and real GDP is 30 billion pizzas. Multiplying these numbers together, we find that nominal GDP is indeed 300 billion pesos. Sometimes, for shorthand, we use the term *price* to mean the price level in a given year and the term *output* to mean real GDP in a given year.

Real GDP is the variable that most interests us because it measures the quantity of goods and services produced in an economy. We would therefore like to find a way to decompose nominal GDP into the price level and the level of real GDP in actual economies. But real economies produce lots of different goods and services, the prices of which are continually changing. In addition—unlike

our fictional economy, where it makes sense to measure real GDP as the number of pizzas—there is no “natural unit” for real GDP in an actual economy.

In fact, even in our pizza economy, there is still an arbitrariness about the units. Imagine that we cut each pizza into 10 slices. Then we could just as easily say that real GDP is 300 billion pizza slices instead of 30 billion pizzas, but that the price level—the price per slice—is 1 peso. We would still conclude that nominal GDP—the number of slices multiplied by the price per slice—was 300 billion pesos.

So is it possible to say, in a real economy producing multiple goods and services, that nominal GDP is equal to the product of the price level and the level of real GDP? Does it still make sense to write

$$nominal\ GDP\ =\ price\ level \times\ real\ GDP$$

as we did for the pizza economy? The answer, as it turns out, is yes.

To see how this works, we begin by looking at how prices and output change from one year to another. Specifically, we divide 2013 nominal GDP by 2012 nominal GDP. This is one measure of the growth in nominal GDP from 2012 to 2013.Specifically, this measures the *gross growth rate* of nominal GDP. It is equal to 1 + the percentage change in nominal GDP. See the toolkit for details of the mathematics of growth rates. Remember that nominal GDP equals total output produced in a year, valued at the prices prevailing in that year. Comparing nominal GDP in 2012 and 2013 therefore gives us

Now we use a trick. Multiply above and below the line by “output in 2013 valued at 2012 prices” and then rearrange:

Look carefully at this calculation to make sure you understand what we did here.

Now examine the two ratios on the right-hand side of the second line. The first compares the cost of the same bundle of goods (output in 2013) at two different sets of prices—those prevailing in 2013 and those prevailing in 2012. Think of the bundle as being a grocery cart full of goods. If you compare how much it costs to buy *exactly the same collection of goods at two different times*, you have a measure of what has happened to prices.

The second ratio on the right-hand side is a measure of the increase in real GDP. It uses the same prices to compare the value of output in 2012 and 2013. In other words, it tells you how much it costs to buy *two different collections of goods at exactly the same prices*.

To reiterate, the first ratio compares the same bundle of goods at two different sets of prices. The second ratio compares two different bundles of goods at the same prices. We have succeeded in separating the change in nominal GDP into two components: a price change and a change in real GDP.

Measuring Real GDP and the Price Level

We can illustrate this technique using the data in Table 18.2.1 "Calculating Nominal GDP". In that example, the growth in nominal GDP equals 10.5 percent because

Now we choose an arbitrary year that we call the *base year*. For the base year, we set the price level equal to 1. In our calculations, we choose 2012 as our base year. Because nominal GDP equals the price level times real GDP, this means that real GDP in 2012 is $400.

When we choose 2012 as our base year, we use the prices of T-shirts, music downloads, and meals in 2012 for our calculations of real GDP for 2012 and 2013. Table 18.2.3 "Real GDP Using 2012 as the Base Year" shows what we find. The first row is exactly the same as in Table 18.2.1 "Calculating Nominal GDP". Nominal GDP in 2012 is—by definition—the same as real GDP in 2012 because we are using 2012 as the base year. The second row of the table calculates real GDP for 2013; it uses 2013 quantities but 2012 prices. Notice also the heading in the final column of the table: “Real GDP (Year 2012 dollars).” The term in parentheses tells us that everything is being measured according to the prices that prevailed in our base year of 2012.

Table 18.2.3: Real GDP Using 2012 as the Base Year

Year	T-shirts		Music Downloads		Meals		Real GDP (Year 2012 Dollars)
	2012 Price ($)	Quantity	2012 Price ($)	Quantity	2012 Price ($)	Quantity	
2012	20	10	1	50	25	6	400

Year	T-shirts		Music Downloads		Meals		Real GDP (Year 2012 Dollars)
2013	20	12	1	60	25	5	425

We previously calculated that 2013 nominal GDP—output in 2013 valued at 2013 prices—was $442. By contrast, Table 18.2.3 "Real GDP Using 2012 as the Base Year" shows that, when valued in year 2012 dollars, the total output of this economy in 2013 is $425. In other words,

Nominal GDP increased by 10.5 percent between the two years, but real GDP is increased by only 6.25 percent. From this we see that not all of the increase in nominal GDP is due to increased output. Some of the increase is because prices increased between 2012 and 2013.

In our pizza economy, we said that nominal GDP was equal to the price per pizza multiplied by the quantity of pizza. In our example here, we have calculated something very similar. Nominal GDP equals the price level multiplied by real GDP. In the base year, the price level equals 1 (that is what it means to choose the base year), and so real GDP equals nominal GDP in that year. Because we can calculate the increase in the price level and the increase in real GDP from one year to the next, we can obtain a time series for the price level and a time series for real GDP. In each year, nominal GDP equals the price level in that year times real GDP in that year.

There is, however, one difference between the calculation for our pizza economy and measurement in real economies. In the pizza economy, because there was a single good, we were able to measure real GDP in physical units—the number of pizzas. In real economies, there is no single good, and so we measure real GDP in base year dollars rather than as a physical quantity. The price level in, say, 2013 is not, strictly speaking, the price of real GDP in terms of 2013 dollars but rather is the price of a base year dollar in terms of 2013 dollars.

But this is a technical difference. From an intuitive point of view, it is simplest to think about real GDP as being a physical quantity—a number of pizzas. In this book we therefore imagine that real GDP is actually a bundle of goods and services all melded together to create a composite good. We call that good “units of real GDP,” and we call the price level the price of a unit of GDP. In fact, we could think about the pizza economy in that same way. Even a basic pizza is itself composed of dough, sauce, and cheese: it is a bundle of items melded into one. So when we talk about the physical quantity of pizza, we are really talking about the number of bundles of these ingredients. Likewise, when we talk of real GDP, we are talking about a bundle of goods that we measure in base year dollars.

Real GDP is our most basic measure of economic performance. It is a very broad measure because it tells us how much economic activity of any kind (at least, any kind that we can measure) is going on in our economy. Real GDP tells us how much we have produced of all the different goods and services that people enjoy and want to consume. For this reason, real GDP statistics are among the most closely watched of all the figures released by a government.

Real GDP in the United States and Argentina

Figure 18.2.3 shows real GDP for the US economy from 1929 to 2008 in year 2000 dollars. The figure shows that the US economy grew substantially over those years. The level of real GDP was $865.2 in 1929 and $10,842 in 2008 (in billions of $2000).If you look at this figure, you will see that real GDP is listed as “chain weighted.” This method of calculating real GDP averages growth rates by using different base years. By averaging, this measure has the virtue that calculations of real GDP are less sensitive to the selection of an arbitrary base year. For more information on chain-weighted measures, see Charles Steindel, “Chain-Weighting: The New Approach to Measuring GDP,” *Current Issues in Economics and Finance* 1, no. 9 (1995): 1–6, accessed June 28, 2011, http://www.newyorkfed.org/research/current_issues/ci1-9.pdf.

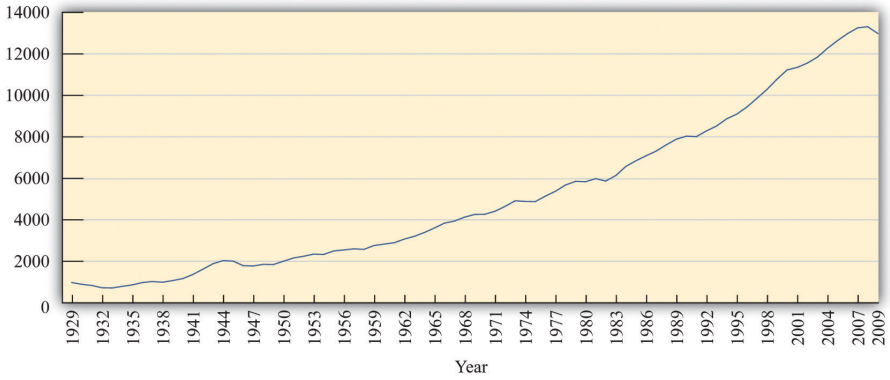


Figure 18.2.3: Real GDP in the United States, 1929–2009. Source: *http://www.bea.gov/national/Index.htm*

Figure 18.2.4 shows real GDP in Argentina and thus reveals that our earlier data for nominal GDP were indeed misleading. Nominal GDP may have increased between 1993 and 2002, but real GDP in 2002 was at the same level as in the previous decade. Moreover, real GDP had been decreasing for the prior four years before the IMF visit.

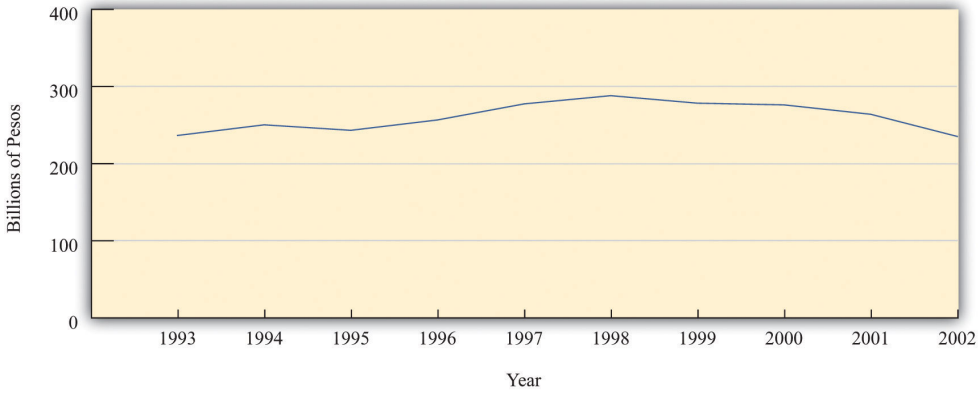


Figure 18.2.4: Real GDP (in 1993 Pesos) in Argentina in the 10 Years Prior to 2002. Real GDP in Argentina was essentially flat between 1993 and 2002.

Source: International Monetary Fund World Economic Outlook database (
http://www.imf.org/external/pubs/ft/weo/2010/01/index.htm).

This helps you to make sense of your contradictory impressions of Buenos Aires. Argentina became poorer, not richer, in the late 1990s and early 2000s. The presence of luxury goods stores, for example, is a reminder that Argentina was a relatively rich country, but the absence of shoppers in those stores tells you that people are not feeling very rich at this time.

Key Takeaways

- Economists and policymakers measure output as GDP. This is a measure of the total value of all production in an economy.
- Nominal GDP measures the total value of all production using current prices, while real GDP measures total output and corrects for changes in prices relative to a base year.

Exercises

1. Why is there no “natural unit” for calculating real GDP in an actual economy compared to the pizza economy?
2. If your income is currently $150 each week and you received a raise of $50, what is the percentage change in your weekly income?

This page titled 18.2: Measuring Economic Activity is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

18.3: Measuring Prices and Inflation 18.3: Measuring Prices and InflationEdit Update titlecancel

Learning Objective

After you have read this section, you should be able to answer the following questions:

1. How are price indices such as the Consumer Price Index (CPI) calculated?
2. What is the difference between the CPI and gross domestic product (GDP) deflator?
3. What are some of the difficulties of measuring changes in prices?

If nominal GDP increased in Argentina but real GDP did not, then prices must have increased. So now we look in more detail at the measurement of prices.

The Price Index

Remember that we defined the change in prices as follows:

We can use the data in Table 18.2.1 "Calculating Nominal GDP" to calculate this ratio as well. This time, however, we compare the cost of the *same* basket of goods (in this case, output in 2013) according to the prices prevailing at two different times. The basket of goods in 2013 is shown in Table 18.3.1 "Calculating the Price Index" as the quantities of the three goods and services produced that year: 12 T-shirts, 60 music downloads, and 5 meals. As we saw earlier, the cost in dollars of this basket of goods and services is $442.

Table 18.3.1: Calculating the Price Index

Year	T-shirts		Music Downloads		Meals		Cost of 2013 Basket ($)	Price Index
	Price ($)	Quantity	Price ($)	Quantity	Price ($)	Quantity		
2012	20	12	1	60	25	5	425	1.00
2013	22	12	0.80	60	26	5	442	1.04

Table 18.3.1 "Calculating the Price Index" also shows the total cost of consuming the 2013 basket in 2012, which we already know is $425. Thus the price index for 2012 is $425/$425 = 1, and the price index for 2013 is $442/$425 = 1.04.Frequently, the value for the price index is multiplied by 100, so the price index for 2013 would be given as 104. For the simple three-good economy described in Table 18.2.1 "Calculating Nominal GDP", we therefore have the following:

Prices increased by 4 percent, real GDP increased by 6.25 percent, and nominal GDP increased by 10.5 percent.

To summarize, the basic principle for calculating inflation is as follows: (1) We decide on a bundle of goods and look at how much it costs in a given year. (2) Then we look at the *same bundle of goods in the following year* and see how much it costs. (3) The ratio of the two is called a **price index** and provides a measure of one plus the **inflation rate**.

Toolkit: Section 31.8

A price index for a given year is calculated as the cost of a bundle of goods in that year divided by the cost of the same bundle in the base year. The growth rate of the price index from one year to the next is a measure of the inflation rate.

Different Price Indices

There are many different price indices that are constructed and used for different purposes. They can be constructed for particular categories of goods or regions, for example. If you listen to the news, you may hear references to the Producer Price Index or the Wholesale Price Index. Ultimately, the differences among different price indices simply come down to the bundle of goods that is chosen.

Figure 18.3.1 shows an example of a very particular price index that was used by a supermarket in Thailand to advertise its prices. The store placed two supermarket carts at the entrance with the same bundle of goods in each. The one on the left, with the black label, showed the cost of this cartload of goods at the old prices. It used to cost 1,059.50 Thai baht (approximately US$28). The one on the right, with the red label, showed that the cost of this same bundle of goods was now 916.00 Thai baht. The reduction in price for the basket of goods was 143.50 Thai baht, or about 13.5 percent.



Figure 18.3.1: An Example of a Price Index. A supermarket in Phuket, Thailand, used an actual basket of groceries to show that its prices had been reduced. This is an example of a price index.

Source: Image taken by the authors.

The Consumer Price Index

In this book, we use price indices that measure the general level of inflation. There are several such measures, but we do not need to worry about this. The differences among these different measures are usually small and typically unimportant for our basic understanding of the economy. The measure of inflation that we have used so far is called the **GDP deflator**, a price index that uses as the bundle of goods everything that goes into GDP. A more common measure of inflation is the **Consumer Price Index (CPI)**, which uses as the bundle of goods the typical purchases of households.

The CPI is the most familiar measure of prices. When economic commentators speak of inflation, they usually mean the percentage change in the CPI. As the name suggests, the CPI is intended to measure inflation as consumers experience it. The bundle of goods included in the CPI is supposed to correspond to the bundle of goods purchased by a typical household. This means that certain goods that are included in GDP do not show up in the CPI. For example, an increase in the price of stealth bombers does not show up in the CPI because (we hope!) households do not buy stealth bombers. However, stealth bombers do show up in the GDP deflator. At the same time, certain goods that are not part of GDP are included in the CPI—most importantly, consumer goods that are imported from other countries. Because imported goods are not produced in the domestic economy, they do not show up in the GDP deflator; however, because domestic consumers purchase imported goods, they do show up in the CPI.

Households differ dramatically in their consumption patterns, so different households have very different experiences of inflation. An individual who drives 100 miles daily to get to work views variations in the price of gasoline very differently from someone who rides a bicycle to work. The CPI captures the average experience of all households, which can be quite different from the actual experience of an individual household.

Figure 18.3.2 shows the CPI inflation rate (that is, the percentage change of the CPI) from 1914 to 2008 in the United States.Inflation data and more details about the construction of price indices can be found at the website of the Bureau of Labor Statistics (BLS; http://www.bls.gov). In some early years, prices actually decreased from one year to the next, meaning that the inflation rate was negative. Since 1960, however, the United States has experienced a positive inflation rate.

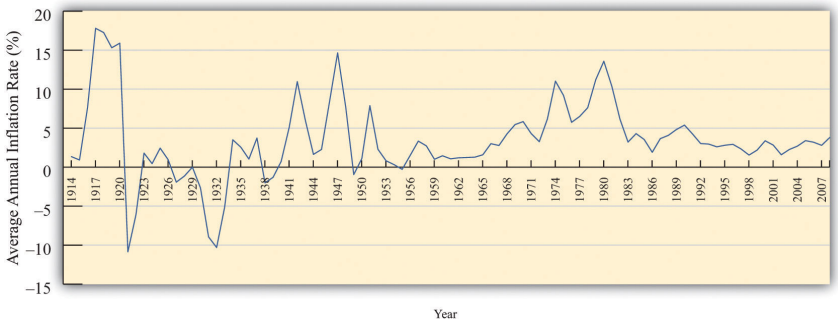


Figure 18.3.2: The Inflation Rate in the United States, 1914–2008. Source: *ftp://ftp.bls.gov/pub/special.requests/cpi/cpiai.txt*

Figure 18.3.3 shows the price level in Argentina between 1993 and 2002. The most striking thing about this picture is that there was very little inflation for most of this period. In the final year, however, prices increased substantially. Notice that our picture for the United States shows the inflation rate, whereas for Argentina we are looking at the level of prices. Either way of presenting the data is valid, but it is critical to understand the difference between them. Make sure you understand the difference between the *level* of prices and the *percentage change* in prices.

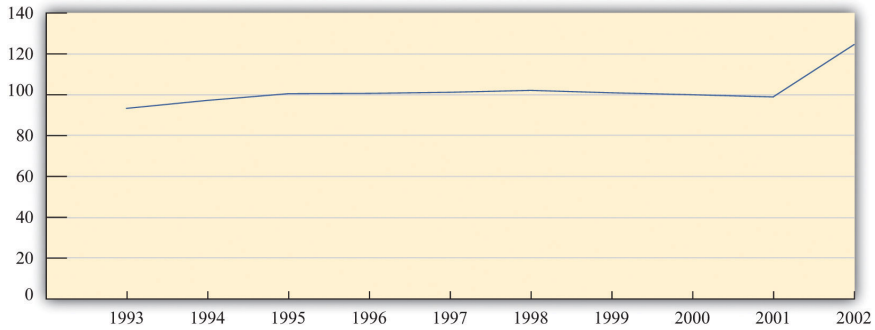


Figure 18.3.3: The Price Level in Argentina. The price level in Argentina was roughly constant between 1993 and 2001. However, there was a big jump in the price level in 2002.

Source: International Monetary Fund World Economic Outlook database (
http://www.imf.org/external/pubs/ft/weo/2010/01/index.htm).

Calculation of the CPI in Practice

The actual calculation of the CPI is more complicated than our example suggests. The Bureau of Labor Statistics (BLS; http://www.bls.gov/cpi) is the US government agency that is responsible for this calculation, while other countries have similar agencies. The BLS procedure is, in essence, the one we have described: it compares the cost of the same bundle of goods in different years. However, the BLS confronts several difficulties that we have ignored so far.

1. **Quality changes.** Imagine that you now work for the BLS (you took this job after you left the International Monetary Fund [IMF]) and are asked to look at changes in the price of laptop computers. You decide to use the IBM ThinkPad computer.For the history of the ThinkPad, see “ThinkPad: A Brand That Made History,” Lenovo, accessed June 28, 2011, http://www.pc.ibm.com/ca/thinkpad/anniversary/history.html; for the 2011 specifications and prices, see “Lenovo Announces Premium ThinkPad Edge E220s, E420s SMB Notebooks,” ThinkPads.com, January 3, 2011, accessed July 20, 2011, www.thinkpads.com/2011/01/03/lenovo-announces-premium-thinkpad-edge-e220s-e420s-smb-notebooks. You discover that in 1992 a ThinkPad cost $4,300 on average. Then you find that it is possible to purchase a ThinkPad in 2011 for $899. You calculate the percentage change in the price as ($899 − $4,300)/$4,300 = −$3,401/$4,300 = −0.79 and conclude that the ThinkPad is 79 percent cheaper than two decades previously. You report this to your boss and then go home.

   But then you start to worry. The 2011 ThinkPad is nothing like the 1992 version. The 1992 computer had 120 MB of memory and weighed over 5.5 pounds. The 2011 ThinkPad has 4 GB of memory and weighs 2 pounds less. It has a vastly bigger hard drive, wireless Internet connection, and a superior display. In short, there were huge quality improvements over this period. A computer with the specifications of the 1992 ThinkPad would be worth much less than $899. By ignoring the improvements in quality, you have understated how much the price of computers has fallen.

This problem is particularly acute for computers, but it applies to all sorts of different goods. The new car that you purchase today is very different from a car that your mother or your grandfather might have bought. Cars today come equipped with computerized braking systems, global positioning system (GPS) navigational tools, and numerous other sophisticated engineering features. They are also much more reliable; your grandparents will tell you that cars used to break down all the time, whereas now that is a relatively rare event. It would be a big mistake to say that a 2012 automobile is the same as a 1961 automobile.

2. **New goods and old goods.** The typical basket of goods bought by consumers is changing. In 1970, no one had a mobile phone, an MP3 player, or a plasma television. Similarly, people today are not buying vinyl records, videocassette recorders, or Polaroid cameras. The BLS needs to keep up with every change. As the economy evolves and new goods replace old ones, they must change the basket of goods.
3. **Changes in purchasing patterns.** The bundle purchased by the typical household also changes over time because of changes in the prices of goods and services. The typical household will substitute away from expensive goods to relatively cheaper ones. If the basket of goods is held fixed, the calculation of the CPI will overstate the increase in the cost of living. This effect is most severe if there are two goods that are very close substitutes and the price of one increases significantly relative to another.

Perhaps these seem like minor details in the calculation of the CPI. They are not. A government commission chaired by the economist Michael Boskin provided an extensive report on biases in computing the CPI in 1996. The Boskin Commission concluded the following: “The Commission’s best estimate of the size of the upward bias looking forward is 1.1 percentage points per year. The range of plausible values is 0.8 to 1.6 percentage points per year.” That is, the Boskin Commission concluded that if inflation as measured by the CPI was, say, 3.1 percent, the true inflation rate was only 2 percent. In response to these concerns with measurement, the BLS responded by taking actions to reduce the biases in the measurement of the CPI and deal more effectively with the introduction of new goods.For the complete Boskin Commission Report, see Advisory Commission to Study the Consumer Price Index, “Toward a More Accurate Measure of the Cost of Living,” Social Security Administration, December 4, 1996, accessed June 28, 2011, http://www.ssa.gov/history/reports/boskinrpt.html. For the BLS response to the report, see “Consumer Price Index: Executive Summary,” Bureau of Labor Statistics, October 16, 2011, accessed June 28, 2011, www.bls.gov/cpi/cpi0698b.htm.

Correcting for Inflation

The data on nominal and real GDP in Argentina illustrate the dangers of looking at nominal rather than real variables. Had you looked at only nominal GDP, you would have concluded that the Argentine economy had been growing between 1993 and 2002, when it was actually stagnating.

But many economic statistics—not only nominal GDP—are typically quoted in terms of dollars (pesos, euros, ringgit, or whatever the currency of the country is). To make sense of such statistics, we must understand whether changes in these statistics represent real changes in the economy or are simply a result of inflation.

Toolkit: Section 31.8

If you have some data expressed in nominal terms (for example, in dollars) and you want to covert them to real terms, use the following steps.

1. Select your deflator. In most cases, the CPI is the best deflator to use.
2. Select your base year. Find the value of the index in that base year.
3. For all years (including the base year), divide the value of the index in that year by the value in the base year. (This means that the value for the base year is 1.)
4. For each year, divide the value in the nominal data series by the number you calculated in Step 3. This gives you the value in base year dollars.

Here is an example of how to correct for inflation. Suppose that a sales manager wants to evaluate her company’s sales performance between 2000 and 2005. She gathers the sales data shown in Table 18.3.2 "Sales, 2000–2005".

Table 18.3.2: Sales, 2000–2005

Year	Sales (Millions of Dollars)
2000	21.0

Year	Sales (Millions of Dollars)
2001	22.3
2002	22.9
2003	23.7
2004	24.1
2005	24.7

At first glance, these numbers look reasonably encouraging. Sales have grown every year between 2000 and 2005. But then she remembers that these data are in nominal terms, and there was also some inflation over this time period. So she decides to correct for inflation. She first goes to the *Economic Report of the President* and downloads the data in Table 18.3.3 "Consumer Price Index, 2000–2005".See *Economic Report of the President*, 2011, Table B-60, accessed June 28, 2011, www.gpoaccess.gov/eop. She decides to use 2000 as the base year—she wants to measure sales in year 2000 dollars. So there are two steps to her calculations, as shown in Table 18.3.4 "Sales Data Corrected for Inflation, 2000–2005". First, she takes the CPI series and divides every term by the 2000 value (that is, 172.2). This gives the third column of Table 18.3.4 "Sales Data Corrected for Inflation, 2000–2005", labeled “Price Index.” Then she divides each of the sales figures by the corresponding price index to obtain the real (that is, corrected for inflation) value of sales. These are given in the final column of the table.

Table 18.3.3: Consumer Price Index, 2000–2005

Year	CPI
2000	172.2
2001	177.1
2002	179.9
2003	184.0
2004	188.9
2005	195.3

Source: ftp://ftp.bls.gov/pub/special.requests/cpi/cpiai.txt

Table 18.3.4: Sales Data Corrected for Inflation, 2000–2005

Year	CPI	Price Index (Base = 2000)	Sales (Millions of Dollars)	Real Sales (Millions of Year 2000 Dollars)
2000	172.2	1.00	21.0	21.0
2001	177.1	1.03	22.3	21.7
2002	179.9	1.04	22.9	21.9
2003	184.0	1.06	23.7	22.2
2004	188.9	1.10	24.1	22.0
2005	195.3	1.13	24.7	21.8

We can see that the sales data are much less rosy after we account for inflation. Sales were increasing between 2000 and 2003 in real terms, but real sales decreased in 2004 and 2005. Had she just looked at the dollar measure of sales, she would have completely missed the fact that the business had experienced a downturn in the last two years.

Economic statistics reported in the news or used by businesspeople are very often given in nominal rather than real terms. Perhaps the single most important piece of “economic literacy” that you can learn is that you should always correct for inflation. Likewise, you should be on your guard for misleading statistics that fail to make this correction. Here is an example from an article that

appeared in the *Washington Post*. “The Clinton recovery has been far less egalitarian than the much-criticized Reagan ‘era of greed.’ Between 1990 and 1995, the [real average] family income actually declined slightly while the number of people with a net worth over $1 million more than doubled.”See J. Kotkin and D. Friedman, “Keep the Champagne on Ice,” *The Washington Post*, reprinted in *The Guardian Weekly*, June 7, 1998. In fact, the quote in the newspaper was even more misleading because it did not even make it clear that the family income figure was adjusted for inflation.

Can you see why this sentence is so misleading? It mixes together a real measure and a nominal measure in the same sentence. Real family income—that is, family income corrected for inflation—declined in the first half of the 1990s. But the number of millionaires is a nominal measure. In a time of inflation, we would expect to have more millionaires, even if people are not really getting any richer.

Key Takeaways

- A price index is created by calculating the cost of purchasing a fixed basket of goods in different years.
- The CPI is a price index for goods and services, including imported goods, consumed by households, while the GDP deflator is based on all the goods and services that compose GDP.
- Calculating a price index is difficult due to the introduction of new products, quality changes, and changes in purchasing patterns.

Exercises

1. The BLS has an inflation calculator on its website (http://data.bls.gov/cgi-bin/cpicalc.pl), which is shown in Figure 18.3.4.

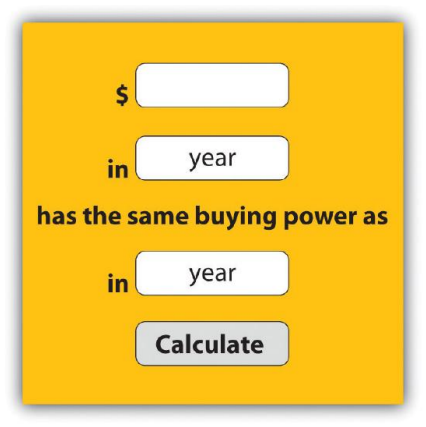


Figure 18.3.4: BLS Inflation Calculator

1. You enter an amount and two different years, and then it tells you the other amount. Explain the calculation that this program performs.
2. In Table 18.3.4 "Sales Data Corrected for Inflation, 2000–2005", calculate the inflation rate (that is, the percentage change in the price index) and the growth rate of sales in each year. What is the relationship between these two variables (a) when real sales are increasing and (b) when real sales are decreasing?

This page titled 18.3: Measuring Prices and Inflation is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

- **8.6: Beyond Perfect Competition** by Anonymous is licensed CC BY-NC-SA 3.0. Original source: https://2012books.lardbucket.org/books/theory-and-applications-of-economics.
- **Current page** by Anonymous is licensed CC BY-NC-SA 3.0. Original source: https://2012books.lardbucket.org/books/theory-and-applications-of-economics.

18.4: The Circular Flow of Income 18.4: The Circular Flow of IncomeEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What is the circular flow of income?
2. What is the national income identity?

Looking at some basic measurements of the economy has allowed you to be more concrete about the problems in Argentina. You report back to the International Monetary Fund (IMF) team that production has been declining in recent years. You also report that there was a recent increase in the price level. As yet, though, you do not know anything about either the causes or the consequences of these events. Measurement of the economy tells you *what* has happened, but it tells you neither why it happened nor what it means. Measurement is not enough. We need frameworks to help us make sense of the data that we gather.

Economists use many different kinds of frameworks to make sense of an economy. One of the most important is called the **circular flow of income**. To understand the circular flow, recall our working definition of economic activity: "goods and services produced for sale." So far, we have focused on production. Now we think about the "for sale" part.

Toolkit: Section 31.27

As individuals and firms buy and sell goods and services, money flows among the different sectors of the economy. The circular flow of income describes these flows of dollars. From a simple version of the circular flow, we learn that, as a matter of accounting,

$$gross\ domestic\ product\ (GDP)\ =\ income\ =\ production\ =\ spending.$$

This relationship lies at the heart of macroeconomic analysis.

There are two sides to every transaction. When you purchase a piece of computer software, you give money to the seller, and the seller gives the software to you. (You might literally hand over dollar bills and receive a CD, or you might enter a credit card number into a website entitling you to a download. The idea is the same either way.) There is a flow of money from you to the seller and a flow of goods or services from the seller to you. This is true for all transactions: as individuals and firms buy and sell goods and services, money flows among the different sectors of the economy. Macroeconomists follow the money. By tracking these flows, we can understand the links between different markets; by understanding these links, we gain insight into the functioning of an economy.

One linkage is between income and spending. The spending by households on goods and services is funded by the income that households earn. But this income comes from firms, and they get *their* income from the spending of households. Thus there is a circular flow of income in an economy as a whole.

Household income comes from two main sources: (1) Households contain workers who sell their time to firms and receive wages in return. (2) Households are the ultimate owners of the firms—shareholders live in houses too—and thus any profits that firms make are returned to households. All firms in an economy are owned by someone, and any profits they make do not vanish into thin air but must eventually show up as someone's income.

Households take this income and do one of two things: they either spend it or save it. To start, let us figure out what would happen if no household income is saved. Households spend all their income, and this money becomes the revenue of firms. Firms send these revenues back to households, either as labor income or profits, and so the circular flow continues.

The Simplest Version of the Circular Flow

We can make this idea more precise, using the pizza economy to illustrate. Imagine that our economy is composed of two *sectors*, which we call households and firms. Households supply labor to firms and are paid wages in return. Firms use that labor to produce pizzas and sell those pizzas to households. There is a flow of goods (pizzas) from firms to households and a flow of labor services

(worker hours) from households to firms. Because there are two sides to every transaction, there is also a flow of dollars from households to firms, as households purchase pizza, and a flow of dollars from firms to households, as firms pay workers.

For now, think of firms as very simple entities that pay out all the income they receive in the form of wages to workers. As a result, 300 billion pesos flow from the household sector to the firm sector (the purchase of pizzas) each year, while 300 billion pesos flow from the firm sector to the household sector (the payment of wages). These flows of pesos are illustrated in Figure 18.4.1. Think of this diagram as representing the interaction of many households with many firms. A particular household works for one (or perhaps a few firms) but purchases goods and services from many firms. (If you like, imagine that different firms specialize in different kinds of pizza.) A feature of modern economies is that individuals *specialize in production* of goods and services but *generalize in consumption* by consuming many varieties of goods and services.

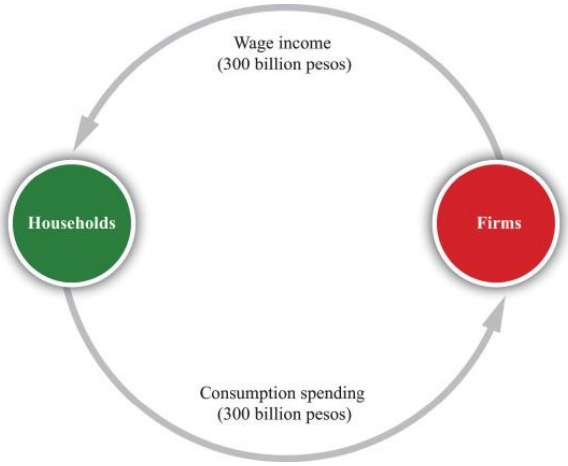


Figure 18.4.1: The Simplest Version of the Circular Flow. The circular flow of income follows the money in an economy. In the pizza economy, firms produce pizzas and sell them to households, while households sell labor to firms and purchase pizzas from them.

The circular flow reveals that there are several different ways to measure the level of economic activity. From the household perspective, we can look at either the amount of income earned by households or their level of spending. From the firm perspective, we can look at either the level of revenues earned from sales or the amount of their payments to workers and shareholders. In all cases, the level of nominal economic activity would be measured at 300 billion pesos.

Corresponding to the flows of pesos shown in Figure 18.4.1, there are flows of goods and services between these sectors, as shown in Figure 18.4.2. The wage income received by consumers is payment for labor services that flow from households to firms. The consumption spending of households is payment for the goods that flow from firms to households.

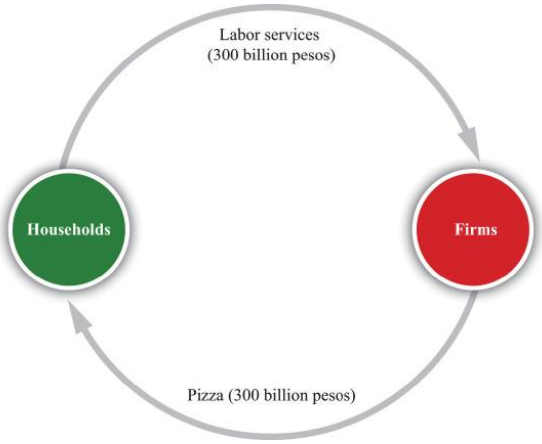


Figure 18.4.2: The Flows of Goods and Labor within the Circular Flow. There are flows of goods and labor services that correspond to the flows of pesos shown in *Figure* 18.4.1. *Three hundred billion pesos worth of pizza flows from firms to households, and 300 billion pesos worth of labor services flow from households to firms.*

Of course, there are also flows of dollars *within* the household and firm sectors as well as between them. Importantly, firms purchase lots of goods and services from other firms. One of the beauties of the circular flow construct is that it allows us to describe overall economic activity without having to go into the detail of all the flows among firms.

Figure 18.4.3 shows us that the flows in and out of each sector must balance. In the household sector, total spending by the household equals total income for the household. If spending equals income for each individual household, then spending also equals income for the household sector as a whole. Similarly, each firm has a balance sheet. Accounting rules ensure that all of a firm's revenues must ultimately show up on the other side of the balance sheet as payments for the inputs that the firm uses (in our simple example, the firm's only input is labor). As this is true for each individual firm, it is also true for the sector as a whole.

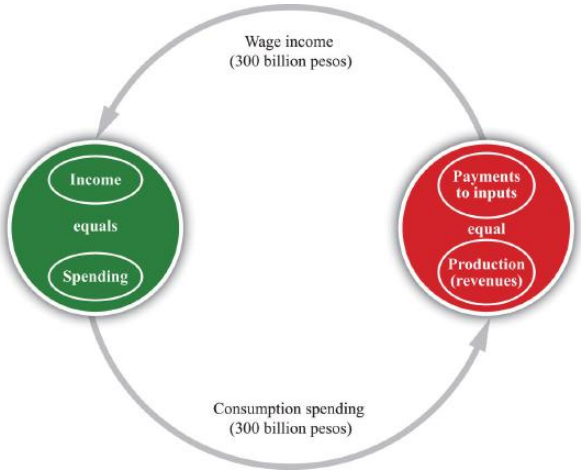


Figure 18.4.3: . Income, Spending, Payments to Inputs, and Revenues in the Simple Circular Flow

In each household, and thus in the household sector as a whole, income must equal spending. In each firm, and thus in the firm sector as a whole, revenues must equal payments to inputs. GDP measures the production of the economy and total income in the economy. We can use the terms production, income, spending, and GDP interchangeably.

Although this version of the circular flow is simple, it teaches us four key insights that remain true (albeit in slightly refined forms) in more sophisticated versions as well.

1. **Spending = production.** The total value of all spending by households becomes an inflow into the firm sector and thus ends up on the revenue side of a firm's balance sheet. The revenues received by firms provide us with a measure of the total value of production in an economy.
2. **Production = payments to inputs.** Flows in and out of the firm sector must balance. The revenues received by firms are ultimately paid out to households.
3. **Payments to inputs = income.** Firms are legal entities, not people. We may talk in common speech of a firm "making money," but any income generated by a firm must ultimately end up in the hands of real people—that is, in the household sector of an economy. The total value of the goods produced by firms becomes an outflow of dollars from the firm sector. These dollars end up in the hands of households in the form of income. (This ownership is achieved through many forms, ranging from firms that are owned and operated by individuals to giant corporations whose ownership is determined by stock holdings. Not all households own firms in this way, but in macroeconomics it is sufficient to think about the average household that does own stock in firms.)
4. **Income = spending.** We complete the circle by looking at the household sector. The dollars that flow into the household sector are the income of that sector. They must equal the dollars that flow out of the household sector—its spending.

The circular flow of income highlights a critical fact of national income accounting:

$$GDP = income = spending = production.$$

Earlier, we emphasized that GDP measures the production of an economy. Now we see that GDP is equally a measure of the income of an economy. Again, this reflects the fact that there are two sides to each transaction. We can use the terms *income*, *spending*, *production*, and *GDP* completely interchangeably.

What does this mean for your assessment of Argentina? For one thing, it tells you that the decline in real GDP implies a corresponding decline in income. Economists pay a great deal of attention to real GDP statistics for exactly this reason: such statistics provide information on the total amount of income earned in an economy.

The Complete Circular Flow

Figure 18.4.4 shows a more complete version of the circular flow. It includes five sectors: the household and firm sectors that we have seen already, a government sector, a financial sector, and a foreign sector. In every sector of the circular flow, accounting rules tell us that the flow of money in must equal the flow of money out. When we look at this sector by sector, we discover five accounting relationships, each playing an important role in macroeconomics. For now, we take a very quick look at each one in turn.When we revisit each sector in different chapters of this book, we include more precise definitions and more detailed discussion of the individual flows (such as consumption or government purchases).

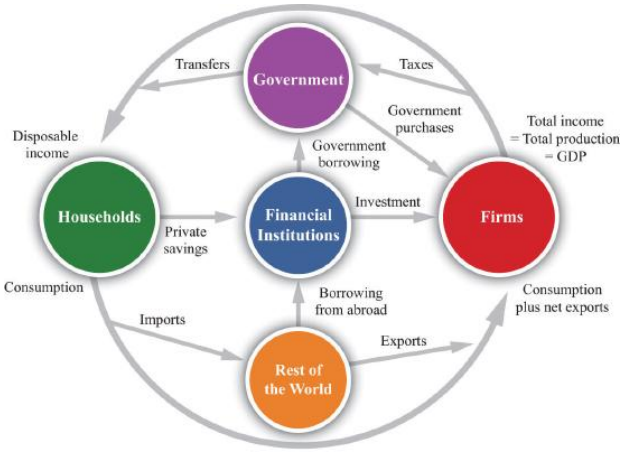


Figure 18.4.4: The Complete Circular Flow. The circular flow of income describes the flows of money among the different sectors of an economy. This representation includes the five main sectors: households, firms, government, the financial sector, and the rest of the world.

The Firm Sector

The flows in and out of the firm sector of an economy must balance. The total flow of dollars from the firm sector measures the total value of production in the economy. The total flow of dollars into the firm sector equals total expenditures on GDP, which we divide up into four categories.

Toolkit: Section 31.27

The **national income identity** is the condition that

$$production = consumption + investment + government\ purchases + net\ exports.$$

It is the most fundamental relationship in the national accounts.

Consumption refers to total expenditures by households on final goods and services. *Investment* refers to the purchase of goods and services that, in one way or another, help to produce more output in the future. *Government purchases* are all the purchases of goods and services by the government. *Net exports* are the difference between exports and imports: they measures the total expenditure flows associated with the rest of the world.These terms are explained in detail in Chapter 22.

The Household Sector

Households receive income from firms. They also receive money from the government (transfers) and must pay money to the government (taxes). Households spend some of their disposable income and save the rest. In other words,

$$income + transfers - taxes = consumption + private\ savings.$$

There are many different ways of saving, but we do not focus on these differences. We simply imagine that households take their savings to financial markets to purchase interest-bearing assets. Some individual households are net borrowers, but, overall, the household sector saves. There is, on net, a flow of dollars from the household sector to the financial sector of an economy. These dollars are then available for firms to borrow to build new factories, install up-to-date equipment, and so on. That is, they are available for investment.The flows in and out of the household sector are discussed in Chapter 27.

The Government Sector

From a macroeconomic perspective, the key functions of government are as follows:

- It purchases goods and services.
- It collects revenues through personal and corporate taxes and other fees.
- It gives transfers to households.

The amount that the government collects in taxes does not need to equal the amount that it pays out for government purchases and transfers. If the government spends more than it gathers in taxes, then it must borrow from the financial markets to make up the shortfall.

Figure 18.4.4 shows two flows into the government sector and one flow out. Since the flows in and out of the government sector must balance, we know that

$$government\ purchases\ =\ tax\ revenues\ -\ transfers\ +\ government\ borrowing.$$

Government borrowing is commonly referred to as the budget deficit. It is also possible that the government takes in more than it spends, in which case the government is saving rather than borrowing, so there is a budget surplus rather than a deficit.Government finances are discussed in Chapter 29.

The Financial Sector

The financial sector of an economy is at the heart of the circular flow. It summarizes the behavior of banks and other financial institutions. Most importantly, this sector of the circular flow shows us that the savings of households provide the source of investment funds for firms. On the left-hand side, the figure shows a flow of dollars from the household sector into financial markets, representing the saving of households. (Though we have not included it in Figure 18.14, firms also save, by means of profits that they retain to finance new investment rather than distribute to their shareholders. As far as the national accounts are concerned, it is as if firms sent these funds to the financial market and then borrowed them back again.) When we borrow from other countries, there is a second flow of dollars into the financial markets. On the right-hand side, there is a flow of money from the financial sector into the firm sector, representing the funds that are available to firms for investment purposes. The linkage between the saving of households and the investment of firms is one of the most important ideas in macroeconomics.

The financial sector is also linked to the government sector and the foreign sector. These flows can go in either direction. As we have already seen, if the government runs a deficit, it does so by borrowing from the financial markets. There is a flow from the financial sector to the government sector. This is the case we have drawn in Figure Runaway argument for \eatSpaces? If the government were to run a surplus, the flow would go in the other direction: government would provide an additional source of saving. The foreign sector can provide an additional source of funds for investment, if those in other countries decide they want to use some of their savings to purchase assets in our economy. In this case, there is a flow from the foreign sector into the financial sector. Again, this is the case we have drawn. If we lend to other countries, then the flow goes in the other direction.

The flows in and out of the financial sector must balance, so

$$investment\ +\ government\ borrowing\ =\ private\ savings\ +\ borrowing\ from\ other\ countries.$$

The Foreign Sector

The foreign sector is perhaps the hardest part of the circular flow to understand because we have to know how international transactions are carried out.

Some of the goods produced in an economy are not consumed by domestic households or firms in an economy but are instead exported to other countries. Whenever one country sells something to another country, it acquires an asset from that country in exchange. For example, suppose a US movie company sells DVDs to an Australian distributor. The simplest way to imagine this is to suppose that the distributor hands over Australian dollar bills to the movie company.. The movie company—and, more generally, the US economy—has now acquired a foreign asset—Australian dollars.

Because these Australian dollars can be used to purchase Australian goods and services at some time in the future, the US economy has acquired a claim on Australia. In effect, the United States has made a loan to Australia. It has sent goods to Australia in exchange for the promise that it can claim Australian products at some future date.

Similarly, some of the goods consumed in our economy are not produced locally. For example, suppose that a US restaurant chain purchases Argentine beef. These are imports. We could imagine that the restaurant chain hands over US dollars to the Argentine farmers. In this case, the United States has borrowed from Argentina. It has received goods from Argentina but has promised that it will give some goods or services to Argentina in the future.

Of course, international transactions in practice are more complicated than these simple examples. Yet the insight we have just uncovered remains true no matter how intricate the underlying financial transactions are. Exports are equivalent to a loan to the rest of the world. Imports are equivalent to borrowing from the rest of the world.

If we import more than we export, then we are borrowing from the rest of the world. We can see this by looking at the flows in and out of the foreign sector:

$$borrowing\ from\ abroad\ =\ imports\ -\ exports.$$

If we export more than we import, then—on net—we are lending to the rest of the world, and there is a flow of dollars from the financial markets to the rest of the world.

The Causes of a Decrease in Real GDP

We saw that, in Argentina, real GDP decreased between 1998 and 2002. The circular flow of income tells us that when real GDP decreases, it must also be the case that real production decreases and real spending decreases. The IMF team in 2002 wanted to understand *why* real GDP decreased. We are not going to answer that question in this chapter—after all, we are still at the very beginning of your study of macroeconomics. Still, the circular flow still teaches us something very important. If real GDP decreased, then there are really only two possibilities:

1. For some reason, firms decided to produce less output. As a consequence, households reduced their spending.
2. For some reason, households decided to spend less money. As a consequence, firms reduced their production.

Of course, it could be the case that both of these are true. This insight from the circular flow is a starting point for explaining what happened in Argentina and what happens in other countries when output decreases.

Key Takeaways

- The circular flow of income illustrates the links between income and spending in an economy. In its simplest form, revenue earned by firms by selling their output ultimately flows to households, which spend this income on the output produced by firms.
- The national income identity says that total spending must equal total output and also must equal total income.

Exercises

1. What changes in Figure 18.4.4 if the government takes in more revenue than it spends?
2. We said that borrowing from abroad equals imports minus exports. Is there an analogous relationship that holds for an individual?

This page titled 18.4: The Circular Flow of Income is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

18.5: The Meaning of Real GDP 18.5: The Meaning of Real GDPEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. When the focus is on real gross domestic product (real GDP), what aspects of economic welfare are then missed?
2. What are some other useful measures of economic welfare?

As you leave Argentina, you might well find yourself wondering about the implications of your work. You know that real GDP decreased, and from your study of the circular flow, you know that income decreased as well. Argentines have become poorer, as you might have guessed from the boarded-up stores you saw when you arrived in the country. You hope that, with the help of your observations, the International Monetary Fund (IMF) and the Argentine government will together find a way to enact good policies to increase the welfare—that is, happiness—of the individuals who live and work in the economy.

Our happiness is surely influenced by our material well-being—our ability to live in comfort; enjoy good food; have access to books, music, computers, and videogames; and so forth. In addition, it depends on our having the leisure time to enjoy these comforts; socialize with our friends; and go to movies, plays, and restaurants. However, our happiness depends on many other factors that are beyond the purview of economics and the influence of economic policymakers. Our happiness depends on our friends, families, health, and much more. Economics cannot help us very much with such matters. So, you wonder, is it enough to look at real GDP?

Real GDP and Economic Welfare

Real GDP is certainly a useful indicator of how well an economy is performing. This does not necessarily mean that it tells us about the welfare of those who live there. Some countries, such as China or India, have a large real GDP simply because they have large populations. Living standards in these countries are nonetheless relatively low because the large GDP must be shared by a very large number of people. To correct for this, we look at real GDP *per person*, which measures how much GDP would be available if we shared it equally across the entire population.

If two countries have substantially different levels of real GDP per person, we can fairly reliably infer that the richer country, by this measure, is also the country with higher living standards. Real GDP per person in Germany is about 25 times greater than real GDP per person in Kenya. Even a few minutes spent in the streets of Nairobi and Berlin would confirm that Germany enjoys much higher material living standards. However, when we compare countries with similar levels of real GDP per person, it is rash to assume that a richer country necessarily enjoys a higher standard of living. This is because there are several ways in which real GDP per person is flawed as an indicator of economic welfare.

Remember, first, that GDP measures market transactions only. National income accounts can measure activities that are traded only in markets. If people clean their own homes, tend their own gardens, repair their own cars, or cook their own meals, these activities are not included in our measurement of GDP. (There are a few exceptions. Most notably, GDP statistics impute a value to owner-occupied housing: GDP statistics effectively pretend that homeowners rent their houses from themselves.)

This leads to some unfortunate inconsistencies in GDP accounting. Suppose you and your neighbor both work as auto mechanics. If you each maintain and repair your own cars, these activities do not show up in GDP. But if you hire your neighbor to maintain your car, and she hires you to repair her car, then GDP *does* include this economic activity. Yet another possibility is that you barter with your neighbor, so she looks after your car and you look after hers but no money changes hands. Again, this work goes unrecorded in national accounts. Barter is more prevalent in developing countries than in developed countries and causes more of a problem for measurement of GDP in poorer countries. It is a particular source of difficulty when we want to compare economic activity in different countries.

People also value their leisure time. GDP measures the goods and services that people consume but does not tell us anything about how much time they must give up to produce those goods. For example, people in the United States are richer, on average, than people in Spain. But people in the United States work longer hours than people in Spain, and Spanish workers also enjoy much longer vacations. If we use measures of GDP to compare welfare in the United States and Spain, we will capture the fact that

Americans can afford more DVD players, but we will miss the fact that they have less time to watch DVDs. GDP measures *material* well-being rather than *overall* welfare.

The economic activity that goes into the production of GDP also often has negative consequences for economic welfare that go unmeasured. A leading example is pollution. Coal-generated power plants generate sulfur dioxide as a by-product of the production of electricity. When sulfur dioxide gets into the atmosphere, it leads to acid rain that damages forests and buildings. This damage is not accounted for in GDP. Emissions from automobiles contribute to the buildup of greenhouse gases in the atmosphere, contributing to global climate change. They also generate smog (technically, particulate matter) that is damaging to health. These adverse effects are not accounted for in GDP.There have been many attempts by economists to amend the GDP measure to take environmental issues into account. For an early discussion on this issue, see William D. Nordhaus and James Tobin, "Is Growth Obsolete?" Yale University, Cowles Foundation paper 398, accessed June 28, 2011, http://cowles.econ.yale.edu/P/cp/p03b/p0398a.pdf.

Critics sometimes argue that GDP not only fails to measure negative effects from production but also erroneously includes measures taken to offset those measures. This criticism is misplaced. Consider the 2010 oil spill in the Gulf of Mexico. The environmental damage from that spill is not included in GDP, which is indeed a problem with using GDP to measure welfare. The costs of cleaning up the gulf are included in GDP, and the inclusion of cleanup costs *does* make GDP a better measure of welfare. To clean up means to produce a cleaner environment from a dirty environment, which increases economic welfare. The problem is the failure to include the original environmental damage, not the inclusion of the cleanup costs.

Finally, real GDP is an aggregate measure. It does not reflect the ways in which goods and services are distributed across the many households of an economy. In comparing two economies, we may feel differently about an economy in which resources are distributed relatively equitably compared to one in which some people are very rich and others are very poor, even if overall real GDP per person is the same. Similarly, we may feel quite differently about changes in real GDP depending on who is reaping the benefit of those changes.

In summary, real GDP is far from a perfect measure of economic welfare, but then again it is not designed to be. It is designed to measure economic activity, and it is—at best—an imperfect measure of material well-being. Nevertheless, when we want to understand what is happening to overall economic well-being or get an idea of comparative welfare in various countries, we begin with real GDP per person. For all its flaws, it is the best single indicator that we have.

Other Indicators of Societal Welfare

Because real GDP is an imperfect a measure of well-being, we look at other statistics as well to gauge overall economic welfare. Here are some examples of economic statistics that we also use as indicators of economic welfare.

Unemployment

The **unemployment rate** is one of the most frequently cited statistics about the macroeconomy; it is the percentage of people who are not currently employed but are actively seeking a job. It signals the difficulty households face in finding employment. GDP data are reported on a quarterly basis only, but unemployment statistics are reported monthly and so contain more up-to-date information than GDP. Chapter 23 contains more discussion of the definition and measurement of the unemployment rate.

Figure 18.5.1 shows the unemployment rate in the United States from 1940 to 2010. On average, the rate of unemployment over this period was 6.0 percent. The unemployment rate was at its highest—14.6 percent—in 1940 and its lowest—1.2 percent—in 1944. The low unemployment in 1944 was largely due to World War II (and is an indication that low unemployment is not always a sign that all is well in an economy). From 1995 to 2008, the unemployment rate was never above 6 percent, but it jumped to 9.3 percent in the major recession of 2008 and by mid-2011 had still not fallen back below 9 percent.

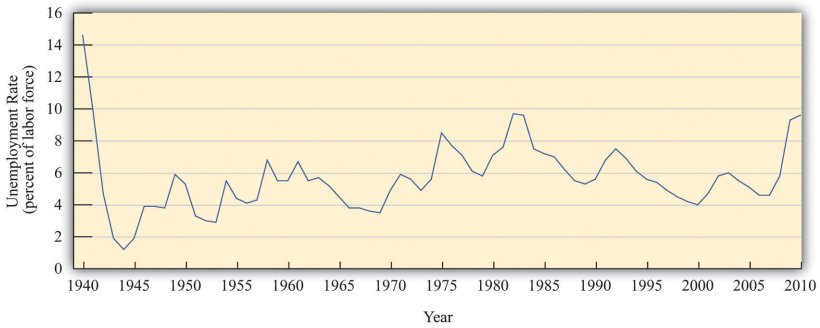


Figure 18.5.1: The Unemployment Rate in the United States. Source: *http://www.bls.gov/cps/home.htm*

In the United States, defining and measuring the unemployment rate and other labor market variables is the job of the Bureau of Labor Statistics (BLS; http://www.bls.gov/cps/home.htm). Each month, about 60,000 households are asked about their recent employment experience. The BLS takes care to be sure that the sample is representative of the entire population of the United States. Notice that it is *households* who are interviewed, not people. So when a household is interviewed, information is acquired about all household members age 16 and over.To be more precise, in addition to being age 16 or older, the survey excludes people in an institution (such as prison) or in the armed forces. As a consequence of the interview, individuals are placed in one of three categories: (1) out of the labor force, (2) in the labor force and working, and (3) in the labor force and looking for a job. Similar surveys are conducted to measure unemployment in other countries.

The (civilian) **labor force** is all individuals who are either working or actively looking for work. That is, it comprises all employed and unemployed workers. Individuals who are not in the labor force are neither employed nor looking for a job. These include those at school or choosing to stay at home. Individuals in the labor force are either employed or seeking work. Employment can be temporary or even part time; as long as someone has a job, he or she is counted as employed. Those who are not at work due to vacation, illness or family issues but who still have jobs are also counted as employed.

The other group in the labor force is a bit more problematic: what exactly does it mean to be looking for a job? The BLS considers you unemployed if you do not have a job and have been seeking one during the past four weeks. Here, “seeking” is intended to be active (going out for job interviews), not passive (reading want ads). Individuals on temporary layoff are considered to be unemployed even if they are not actively looking for a new job. The BLS does not directly ask individuals to classify themselves into one of these three categories. Instead, BLS interviewers ask a series of questions to facilitate the classification. The sum of the civilian labor force and those out of the labor force equals the civilian working age population. Figure 18.5.2 shows the unemployment rate in Argentina between 1993 and 2002. Unemployment was quite high throughout this period: it was in excess of 10 percent in every year from 1994 onward. In addition, the unemployment rate increased substantially in the period when real GDP was decreasing, from 12.8 percent in 1998 to almost 20 percent in 2002. The economic distress you witnessed on the streets of Buenos Aires is reflected in this statistic.

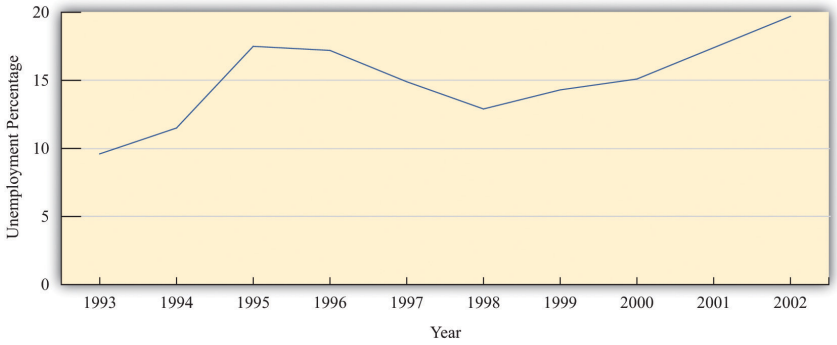


Figure 18.5.2: The Unemployment Rate in Argentina. The unemployment rate in Argentina was about 10 percent in 1993. It increased sharply over the next two years, decreased somewhat in the mid to late 1990s, and then increased again to almost 20 percent in 2002.

Source: Ministra de Economía y Producción de Argentina.

Real Wages

Average real GDP figures tell us nothing about how GDP is shared in an economy. They tell us how big the pie is but not who has the largest and smallest slices. Economists therefore also look at other measures that tell us about the economic environment as it is experienced by workers and households.

Wages in an economy provide a sense of how workers are doing. However, the wage in dollars—the nominal wage—is not the best indicator. While salaries and pay scales for jobs are quoted in dollar terms, decisions on whether or not to take a job and how many hours to work at that job depend on what those dollars can buy in terms of goods and services. If all prices in the economy were to double, then $10 would buy only half as much as it used to, so a job paying $10 an hour would seem much less attractive than it did before.

For this reason, we instead look at the **real wage** in the economy. As with real GDP, *real* here refers to the fact that we are correcting for inflation. It is real wages—not nominal wages—that tell us how an economy is doing. To convert nominal wages to real wages, we need a price index, and because we are looking at how much households can buy with their wages, we usually choose the Consumer Price Index (CPI) as the index.

Toolkit: Section 31.3

The real wage is the wage corrected for inflation. To obtain the real wage, simply divide the wage in dollars—the nominal wage—by the price level:

Figure 18.5.3 shows the nominal (hourly) wage paid to private sector industrial workers from 1964 to 2010. Over this period, the nominal wage rate increased almost eightfold from a low of $2.50 in January 1964 to nearly $19.00 by the end of the period.“Average Hourly Earnings: Total Private Industries” from the Bureau of Labor Statistics, *The Employment Situation*, which is seasonally adjusted. The CPI is the “Consumer Price Index for All Urban Consumers: All Items CPIAUCNS,” Bureau of Labor Statistics Consumer Price Index, 1982 − 84 = 100. The real wage series in Figure 18.5.3 shows the nominal wage divided by the CPI (times 100 so that the real and nominal wages are equal in the base year of the CPI). The nominal wage increased over this period by over five times, but the real wage actually decreased at times. It peaked at near $9.50 in 1973, decreased to $7.62 in 1995, and has risen only slowly since that time.

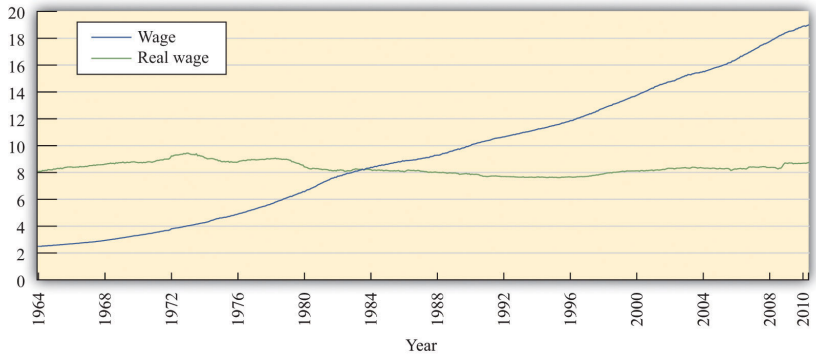


Figure 18.5.3: Real and Nominal Wages. Source: US Department of Labor, Bureau of Labor Statistics, *http://stats.bls.gov/lpc/data.htm*

It is a remarkable fact that, even though US real GDP is now more than 150 percent greater than it was in the early 1970s, real wages are still significantly lower than they were at that time. What is going on here? Part of the story is that other forms of nonwage compensation have become increasingly significant over the past few decades. The most important of these are health-care benefits. When these and other benefits are included, we find that overall compensation has increased reasonably steadily and is about 50 percent greater now than in the early 1970s.More precisely, real hourly compensation in the nonfarm business sector increased by 51.7 percent between 1970 and 2007. See “Labor Productivity and Costs,” Bureau of Labor Statistics, accessed June 28, 2011, http://stats.bls.gov/lpc/data.htm. Total compensation is, in fact, a better measure than real wages. Even so, total compensation has been increasing at a far slower rate than real GDP over the last few decades.

Noneconomic Indicators of Welfare

We turn finally to some noneconomic measures of societal welfare, such as statistics on health and education. Table 18.5.1 "Noneconomic Indicators of Welfare" shows some examples of indicators for four countries.“The World Factbook,” Central Intelligence Agency, accessed June 28, 2011, https://www.cia.gov/library/publications/the-world-factbook/index.html; “The Complete World Development Report Online,” World Bank, accessed June 28, 2011, wdr2011.worldbank.org/WDR2011_Data. Large differences in GDP per person, such as the difference between the United States and Argentina, are reflected in these other measures. GDP per person is about three times greater in the United States than in Argentina, and the United States also has higher adult literacy, higher secondary school enrolment, lower infant mortality, and higher life expectancy.

Table 18.5.1: Noneconomic Indicators of Welfare

Indicator	United States	United Kingdom	Greece	Argentina
GDP per person, 2005 ($US)	42,000	30,900	22,800	13,700
Infant mortality rate, 2006 (deaths per 1,000 live births)	6.43	5.08	5.43	14.73
Life expectancy at birth, 2006 (years)	77.85	78.54	79.24	76.12
Adult literacy rate, 2003 (%)	99.0	99.0	97.5	97.1
Secondary school enrollment ratio, 2002–3 (%)	88	95	86	81

Differences in GDP per person are a much less reliable guide when we compare relatively rich countries. For example, the United States has greater GDP per person than the United Kingdom or Greece. But both of those countries have lower infant mortality rates and higher life expectancy. They also have similar rates of literacy and school enrollment. In fact, based on these measures, the United Kingdom looks like a more attractive country to live in than the United States, even though its GDP per person is 25 percent lower.

Key Takeaways

- Real GDP measures total output and thus total income in an economy, but it does not measure economic activity at home, ignores income distribution, and excludes the effects of economic activity on the environment.
- Measures of unemployment, real wages, and indicators of health and education are also useful indicators of economic welfare.

Exercises

1. If there is an increase in investment and an associated increase in real GDP, why does this increase economic welfare?
2. If there is a decrease in real wages and an offsetting increase in a firm’s profits, does this affect overall household income? If not, what effect does it have on the household sector?

This page titled 18.5: The Meaning of Real GDP is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

18.6: End-of-Chapter Material 18.6: End-of-Chapter MaterialEdit Update titlecancel

In Conclusion

Understanding the meaning and measurement of macroeconomic variables is vital for your ability to evaluate the abundance of information you receive through various forms of the media about the state of the aggregate economy. The difficulties faced by the team of International Monetary Fund (IMF) economists with which we opened the chapter are not that different from the problems each of us faces in understanding what is happening in the economy.

The concepts and variables you have discovered in this chapter are used over and over again in the various applications discussed in this book. We use the concepts of real gross domestic product (real GDP), the inflation rate, the unemployment rate, and so forth almost everywhere in our study of macroeconomics.

Key Links

- *Economic Report of the President*: www.gpoaccess.gov/eop/tables06.html
- Bureau of Labor Statistics (BLS): http://www.bls.gov
  - Unemployment: http://www.bls.gov/cps/home.htm
  - CPI: http://www.bls.gov/cpi
- Bureau of Economic Analysis (BEA): http://www.bea.gov
- World Economic Outlook, International Monetary Fund: http://www.imf.org/external/pubs/ft/weo/2011/01/weodata/index.aspx
- World Bank: http://www.worldbank.org

exercises

1. Which of the following variables are stocks? Which are flows?
   1. The number of cars parked on the street where you live.
   2. The number of cars that drive past your house every day.
   3. The number of people losing their jobs and becoming unemployed.
   4. The blue jeans on the shelves of a GAP store.
   5. The amount of water in a reservoir.
   6. The amount of money you have on your person right now.
   7. The amount of money you spent this week.
2. Suppose an economy produces at least as much—and maybe more—of every good and service this year compared to last year. Also suppose that the price of every single good and service is at least as high this year as it was last year. What, if anything, can you conclude about nominal GDP, real GDP, and the price level between the two years?
3. Why do we exclude intermediate goods when calculating GDP?
4. Redo Table 18.2.3 "Real GDP Using 2012 as the Base Year" assuming that 2013 is the base year.
5. Suppose that Australia had nominal GDP last year equal to 1 trillion Australian dollars and that in the first quarter of this year, its nominal GDP is 252 billion Australian dollars. What is Australia's annualized growth rate of nominal GDP?
6. Suppose that, between 2012 and 2013, a country experiences 3 percent negative inflation (this is known as deflation). In other words, prices are on average 3 percent lower in 2013 compared to 2012. However, the economy also experiences real economic growth of 5 percent. Is nominal GDP in 2013 greater or less than in 2012?
7. If nominal GDP in country A grows faster than nominal GDP in country B, what, if anything, can you conclude about the inflation rates in the two countries?
8. Suppose that the price of Brazilian coffee decreases. What does that imply for the Consumer Price Index (CPI) in Germany? What does that imply for the GDP deflator in Japan?
9. Is it possible for prices to be increasing and the inflation rate to be decreasing at the same time? Explain why or why not.
10. Is it possible for an economy's production to increase at the same time that total income in the economy decreases? Explain why or why not.
11. Which of the following people are classified as unemployed?
    1. A full-time student who also works part time in a store selling CDs.

2. A worker who would like a job but has given up looking because she was unable to find one.
3. An autoworker who was recently laid off and is looking for a new job.
4. A member of the military who is not currently on active duty.
5. A woman on maternity leave from her job.
6. A 70-year-old man who is actively applying for jobs.

12. Give three reasons why real GDP is an imperfect measure of economic welfare.

Economics Detective

1. Update Table 18.2.2 "Nominal GDP in the United States, 2000–2010" and Figure 18.6.1 using data from the Bureau of Economic Analysis (BEA). Using the IMF World Economic Outlook Database (http://www.imf.org/external/pubs/ft/weo/2010/01/weodata/index.aspx), create tables to show nominal GDP, the GDP price deflator, and real GDP for Argentina.
2. A version of GDP that takes into account environmental effects is called “environmental accounting” or “green accounting.” Use the Internet to find a discussion of this alternative way of calculating GDP. List some of the differences between the usual way of calculating GDP and the environmental or green accounting method. Do other countries employ these alternative measures?
3. Try to find out whether people in richer countries are happier than those in poorer countries.

Spreadsheet Exercise

Year	T-shirts		Music Downloads		Meals	
	Price ($)	Quantity	Price ($)	Quantity	Price ($)	Quantity
2012	25	12	1	60	25	5
2013	30	10	1.80	50	23	2

TABLE 18.6.1: DATA

1. Using the data in the preceding table, reconstruct Table 18.2.1 "Calculating Nominal GDP" to calculate nominal GDP, Table 18.2.3 "Real GDP Using 2012 as the Base Year" to calculate real GDP, and Table 18.3.1 "Calculating the Price Index" to calculate a price index and the inflation rate.

This page titled 18.6: End-of-Chapter Material is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

19: The Interconnected Economy

19.1: A Financial Crisis in the News
19.2: Housing Supply and Demand
19.3: Comparative Statics- Changes in the Price of Housing
19.4: Three Important Markets
19.5: Linkages across Markets
19.6: End-of-Chapter Material

This page titled 19: The Interconnected Economy is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

19.1: A Financial Crisis in the News 19.1: A Financial Crisis in the NewsEdit Update titlecancel

Here are some headlines from the fall of 2008. If you were following the news during this time period, you probably saw stories like these. The first excerpt talks about houses in the United States.

Fallout from Financial Crisis Hammers Housing

The nation is on track to build fewer homes this year than at any time since the end of World War II, adding to the woes of an economy that analysts said Friday has almost certainly entered a recession.

[…]

David Seiders, chief economist for the group, said builders are being hit by a double whammy from the financial turmoil: It's harder for them to get loans to pursue new houses, and more difficult to sell those they do build.

[…]Martin Crutsinger, "Fallout from Financial Crisis Hammers Housing," *USA Today*, Money, October 17, 2008, accessed June 28, 2011, http://www.usatoday.com/money/topstories/2008-10-16-3784489146_x.htm.

The next excerpt also concerns housing but this time in the United Kingdom.

Financial Crisis: House-Price Slump to Cost Economy £50 Billion

House prices are set to fall 35 per cent from last year's peak, as the property slump costs the wider economy almost £50bn as people stop buying homes, economists warned.

With house prices predicted to make their biggest fall in British history by dropping 35 per cent by autumn next year, the associated consumer spending is expected to plunge, they said.

[…]

This is expected to have a huge impact on the wider economy as each house sale triggers around £4,000 in new spending on household goods, on items such as washing machines and other white goods.

[…]

The lack of spending in these areas will hit employment, with some analysts forecasting that the construction sector alone could see a loss of up to 350,000 jobs within the next five years.

[…]Myra Butterworth, "Financial Crisis: House-Price Slump to Cost Economy £50 Billion," *The Telegraph*, October 21, 2008, accessed June 28, 2011, www.telegraph.co.uk/finance/economics/houseprices/3235741/Financial-crisis-House-price-slump-to-cost-economy-50-billion.html.

Taking these excerpts together, we notice four things: (1) There was a housing slump—fewer houses being bought and sold, and house prices decreasing—in both the United Kingdom and the United States at around the same time. (2) Both are linked to a financial crisis. (3) These slumps affect other parts of the economy. (4) The housing problems lead to job losses.

The next excerpt tells us that the crisis also affected the value of the US dollar.

Financial Crisis Has One Beneficiary: The Dollar

The great market upheaval of 2008 has stripped 45 percent from the value of global equities, led bank lending to nearly dry up and caused commodity prices to crash from stratospheric heights. And now, paradoxically, it is helping to lift the long-suffering dollar.

[…]See David Jolly, "Global Financial Crisis Has One Beneficiary: The Dollar," *New York Times*, October 22, 2008, accessed June 28, 2011, http://www.nytimes.com/2008/10/22/business/worldbusiness/22iht-dollar.4.17174760.html.

This excerpt tells us that the financial crisis has also affected other prices in the economies of the world. The price of equities—shares in companies—decreased, as did the price of goods such as basic minerals (copper and tin, for example) and basic foods

(rice and coffee, for example). But even as these items became less valuable, the US dollar became more valuable. The price of the US dollar increased.

The Chinese economy was also affected by the crisis:

Agricultural Products Export Growth Slows Down in 2008

China's agricultural products exports rose 9.8 percent year-on-year in 2008 to $40.19 billion, the General Administration of Customs said on Wednesday.

According to the statistics, export growth declined 8.2 percentage points from a year earlier. Exports in the last two months of 2008 fell 6.9 percent and 7.2 percent to $3.47 billion and $3.76 billion respectively over the same period of 2007.

Although the country has increased export rebates for some agricultural products and lowered or even canceled the export tax, exports are unlikely to see a quick rebound in the near future. Poor overseas demand and falling prices in the international market amid the financial crisis, as well as the increasing distrust in China's food quality are likely to stifle export growth, the General Administration of Customs said.

[…]See Tong Hao, "Agricultural Products Export Growth Slows Down in 2008," *China Daily*, February 11, 2009, accessed June 28, 2011, http://www.chinadaily.com.cn/bizchina/2009-02/11/content_7467089.htm.

The excerpt tells us that China's exports of agricultural products have been growing rapidly, reaching a growth rate of nearly 10 percent in 2008. But they had been growing even faster in the previous year. The effects of the financial crisis and the economic downturn are clear: the amount of exports decreased at the end of 2008 (and in fact fell throughout 2009 as well).

We have shown a few headlines about the impact of the 2008 financial crisis. We could have picked thousands of others. For example, if you enter into a search engine the terms *financial crisis* and *XYZ*, where XYZ is just about any product or international currency, you will probably find dozens, perhaps hundreds, of articles. The financial crisis of 2008 affected just about every market—all around the world.

Our task in this chapter is certainly not to fully understand these events. Our goals here are much more modest. First, we want to develop the supply-and-demand framework—perhaps the most basic tool in economics—to understand how an event affecting some good or service leads to changes in the price of that good or service as well as changes in the quantity that is bought and sold. Second, we want to explore some of the ways in which different markets in the economy are linked, for linkages across markets are among the most important features of macroeconomic analysis. The financial crisis is a good illustration because this single event affected so many markets.

Understanding the sources and consequences of changing prices and quantities in the economy is one of the key tasks of an economist. There is an almost endless list of such analyses in economics. In fact, most of the applications in this textbook ultimately come down to understanding, explaining, and predicting changes in prices and quantities. The two questions that motivate this chapter are as follows:

What determines price and quantity in a market?

How are markets interconnected?

Road Map

The story of the crisis of 2008 is fascinating and worth understanding in some detail. We begin with the basics of supply and demand, looking at a single market—the market for houses. We explain how the equilibrium price and quantity in this market are determined, which allows us to understand why the price of housing changes. This is a first step to understanding the crisis of 2008 because the housing market was central to that story.

The story began in the housing market but did not end there. It spread across the economy and across the world. Hence we next look at three significant markets in the economy: the labor market, the credit market, and the foreign exchange market. Understanding how these three markets work is necessary for a good understanding of macroeconomics. We use these markets to provide more illustrations of supply and demand in action. Finally, we look at how markets are linked together to see how what might have seemed like a minor problem in one market turned into a cataclysmic event for the world's economies.

> Throughout this chapter, we use the term “the crisis of 2008” as shorthand, but the first signs of the crisis emerged well before that year, and the effects of the crisis are still being felt several years later. The crisis was a complex event, and right now, at the beginning of your studies of macroeconomics, we are not yet ready to delve deeply into a detailed analysis of those events. We return to the crisis in Chapter 30, which is a capstone chapter that brings together most of the tools of macroeconomics from this book.

This page titled 19.1: A Financial Crisis in the News is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

19.2: Housing Supply and Demand 19.2: Housing Supply and DemandEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What factors underlie the demand for housing?
2. What factors underlie the supply of housing?
3. What determines the amount of housing traded and the price of housing?

The first two articles we quoted from made it clear that the housing market was heavily affected by the financial crisis. More than that, it was where the crisis began—and so it is where we begin our story.

We start with the market for new homes, which are part of real gross domestic product (real GDP). (The buying and selling of existing homes is not counted in GDP.) New homes are supplied by construction firms and demanded by families wishing to live in a new home. New homes are also bought by speculators who purchase houses in the hope that they can resell them for a higher price in the future.

Toolkit: Section 31.9

Supply and demand is a framework we use to explain and predict the equilibrium price and quantity of a good. A point on the market supply curve shows the quantity that suppliers are willing to sell for a given price. A point on the market demand curve shows the quantity that demanders are willing to buy for a given price. The intersection of supply and demand determines the equilibrium price and quantity that will prevail in the market.

The toolkit contains a presentation of supply and demand that you can use for reference purposes in this and the following chapters.

The supply-and-demand framework applies to the case that economists call a **competitive market**. A market is said to be competitive, or, more precisely, to exhibit perfect competition, under two conditions:

1. There are many buyers and many sellers, all of whom are small relative to the market.
2. The goods that sellers produce are perfect substitutes.

In a competitive market, buyers and sellers take the price as given; they think their actions have no effect on the price in the market.

Demand

The market demand for housing is shown in Figure 19.2.1. We call this the **market demand curve** because it reflects the choices of the many households in the economy. In macroeconomics, we typically look at markets at this level of aggregation and do not worry much about the individual decisions that underlie curves such as this one.

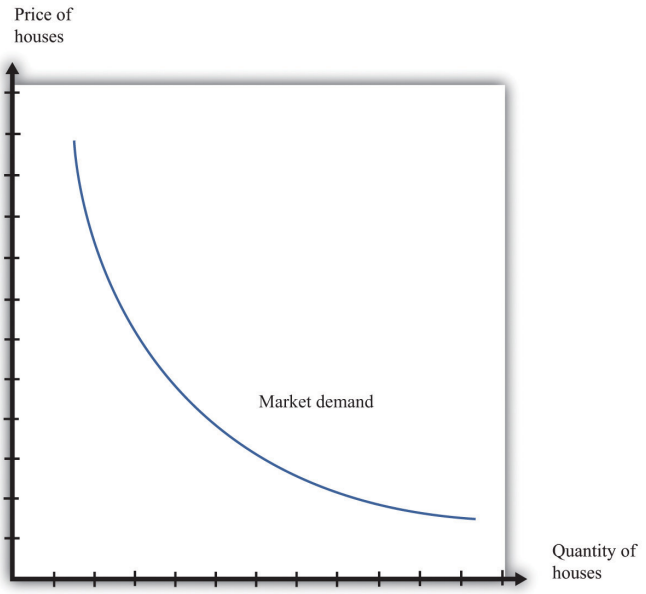


Figure 19.2.1: The Market Demand for Houses. The market demand curve shows the quantity of houses demanded at each price.

As the price of housing decreases, the quantity demanded increases. This is an example of the law of demand, which derives from two effects:

1. As the price of a good or service decreases, more individuals choose to buy a positive quantity rather than zero.
2. As the price of a good or a service decreases, individuals choose to buy a larger quantity.

In the case of the market for housing, the first of these is more important. Most people own either zero houses or one house. As houses become cheaper, more people decide that they can afford a house, so the quantity demanded increases. A few people might decide to buy an additional house, but they would presumably be in the rich minority. For other goods, such as chocolate bars or shoeshines, the second effect is more important: as price decreases, people increase the quantity that they buy.

Shifts in Demand

When we draw a demand curve, we are varying the price but *holding everything else fixed*. In particular, we hold fixed the level of income, the prices of other goods and services in the economy, and the tastes of households. If these other factors change, then the market demand curve will *shift*—that is, the quantity demanded will change at each price.

A leftward shift of the market demand curve for houses, as indicated in Figure 19.2.2, could be caused by many factors, including the following:

- A decrease in the incomes of households in the market
- Concerns about the future health of the economy
- A reduction in the price of a typical apartment rental
- An increase in the interest rates for mortgages
- A change in social tastes so that buying a house is no longer viewed as a status symbol

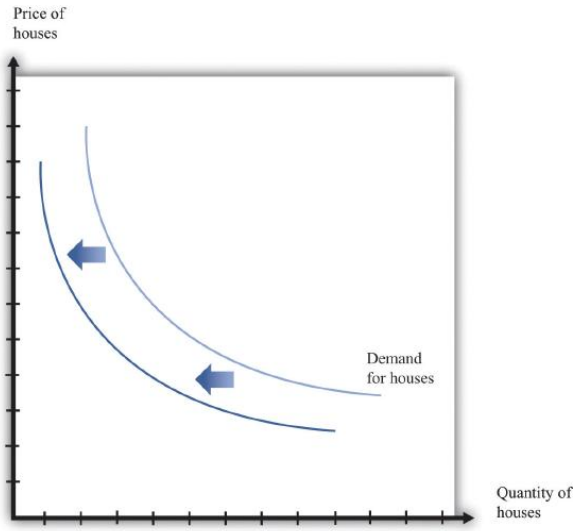


Figure 19.2.2: A Shift in the Market Demand Curve. If there is a decrease in demand for houses, then fewer houses are demanded at each price. The demand curve shifts leftward.

Supply

The counterpart to the market demand curve is the **market supply curve**, which is obtained by adding together the individual supply curves in the economy. The supply curve slopes upward: as price increases, the quantity supplied to the market increases. As with demand, there are two underlying effects.

1. As price increases, more firms decide to enter the market—that is, these firms produce some positive quantity rather than zero.
2. As price increases, firms increase the quantity that they wish to produce.

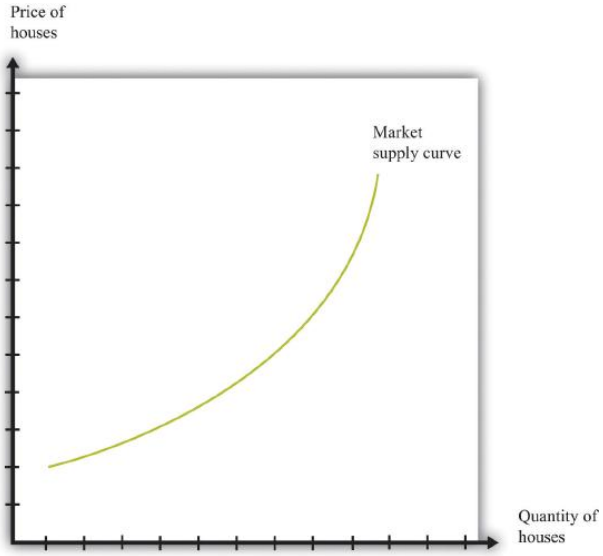


Figure 19.2.3: The Market Supply of Houses. The market supply curve shows the quantity of houses supplied at each price. It has a positive slope: as the price of houses increases, the number of houses supplied to the market increases as well.

Shifts in Supply

When we draw a supply curve, we again vary the price but hold everything else fixed. A change in any other factor will cause the market supply curve to shift. A leftward shift of the market supply curve for houses, as indicated in Figure 19.2.4, could be caused by many factors, including the following:

- Increases in the costs of production, such as wages, the cost of borrowing, or the price of oil
- Bad weather that delays or damages construction in process
- Changes in regulations that make it harder to build

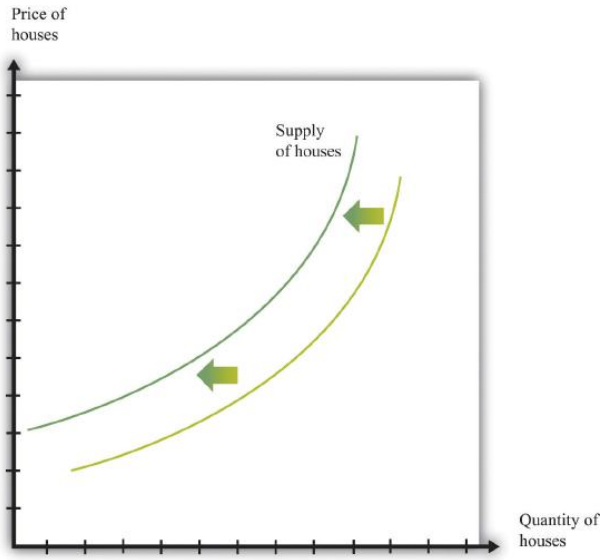


Figure 19.2.4: A Shift in Supply of Houses. If there is a decrease in supply of houses, then fewer houses are supplied at each price. The supply curve shifts leftward.

Market Equilibrium: What Determines the Price of Housing?

We now put the market demand and market supply curves together to give us the supply-and-demand picture in Figure 19.2.5. The point where supply and demand meet is the equilibrium in the market. At this point, there is a perfect match between the amount that buyers want to buy and the amount that sellers want to sell.

Toolkit: Section 31.9

Equilibrium in a market refers to an equilibrium price and an equilibrium quantity and has the following features:

- Given the equilibrium price, sellers supply the equilibrium quantity.
- Given the equilibrium price, buyers demand the equilibrium quantity.

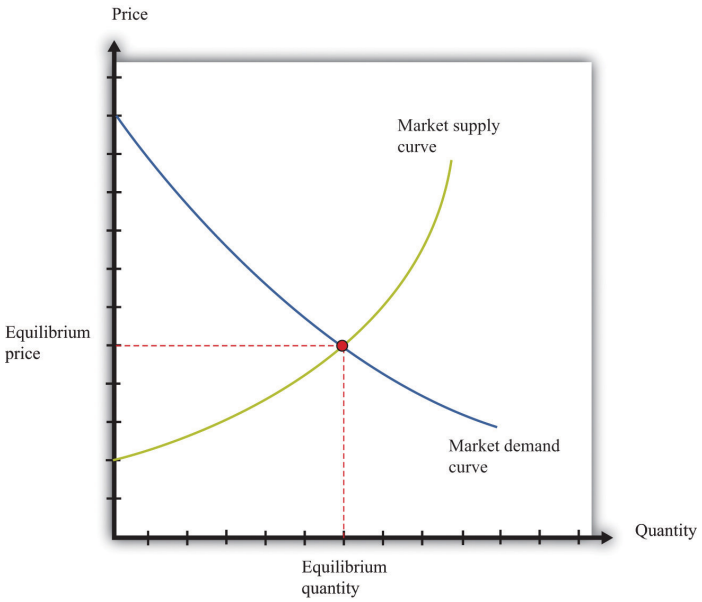


Figure 19.2.5: Market Equilibrium. In a competitive market, equilibrium price and quantity are determined by the intersection of the supply and demand curves.

We speak of equilibrium because there is a balancing of the forces of supply and demand in the market. At the **equilibrium price**, suppliers of the good can sell as much as they wish, and demanders of the good can buy as much of the good as they wish. There

are no disappointed buyers or sellers. Because the demand curve has a negative slope and the supply curve has a positive slope, supply and demand will cross once, and both equilibrium price and **equilibrium quantity** will be positive.

Table 19.2.1 "Market Equilibrium: An Example" provides an example of market equilibrium. It gives market supply and market demand for four different prices. Equilibrium occurs at a price of $100,000 and a quantity of 50 new houses.

Table 19.2.1: Market Equilibrium: An Example

Price ($)	Market Supply	Market Demand
10,000	5	95
50,000	25	75
100,000	50	50
200,000	100	0

Economists typically believe that a perfectly competitive market is likely to reach equilibrium. The reasons for this belief are as follows:

- If price is different from the equilibrium price, then there will be an imbalance between demand and supply. This gives buyers and sellers an incentive to behave differently. For example, if price is less than the equilibrium price, demand will exceed supply. Disappointed buyers might start bidding up the price, or sellers might realize they could charge a higher price. The opposite is true if the price is too high: suppliers might be tempted to try cutting prices, while buyers might look for better deals.
- There is strong support for market predictions in the evidence from experimental markets. When buyers and sellers meet individually and bargain over prices, we typically see an outcome very similar to the market outcome in Figure 19.2.5.
- The supply-and-demand framework generally provides reliable predictions about the movement of prices.

Pictures like Figure 19.2.5 are useful to help understand how the market works. Keep in mind, however, that firms and households in the market do not need any of this information. This is one of the beauties of the market. All an individual firm or household needs to know is the prevailing market price. All the coordination occurs through the workings of the market.

Key Takeaways

- The primary factor influencing demand for housing is the price of housing. By the law of demand, as price decreases, the quantity of housing demanded increases. The demand for housing also depends on the wealth of households, their current income, and interest rates.
- The primary factor influencing supply of housing is the price of housing. As price increases, the quantity supplied also increases. The supply of housing is shifted by changes in the price of inputs and changes in technology.
- The quantity and price of housing traded is determined by the equilibrium of the housing market.

Exercises

1. What would be the impact of a decrease in the cost of borrowing on the market supply curve of housing? What would be the impact of a decrease in the cost of borrowing on the market demand curve?
2. Name two events that would cause the housing market supply curve to shift rightward. Name two events that would cause the housing market demand curve to shift rightward.

This page titled 19.2: Housing Supply and Demand is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

19.3: Comparative Statics- Changes in the Price of Housing 19.3: Comparative Statics- Changes in the Price of HousingEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What are exogenous and endogenous events?
2. How does the equilibrium of a market respond to changes in exogenous variables?
3. What is comparative statics, and how is it used?

A driving factor in the crisis of 2008 was a decrease in the price of new housing. We can use our supply-and-demand tool to help us understand that. We use the framework to make predictions about the effects of events on economic outcomes. More precisely, economists predict the effects of exogenous events on equilibrium prices and quantities.

Toolkit: Section 31.16

An **exogenous** variable is something that comes from outside a model and is not explained in our analysis. An **endogenous** variable is one that is explained within our analysis. When using the supply-and-demand framework, price and quantity are endogenous variables; everything else is exogenous.

A Shift in Demand for Housing

The following is a typical account of the housing market crisis in 2008:

The immediate cause or trigger of the crisis was the bursting of the United States housing bubble which peaked in approximately 2005–2006. High default rates on "subprime" and adjustable rate mortgages (ARM), began to increase quickly thereafter. An increase in loan incentives such as easy initial terms and a long-term trend of rising housing prices had encouraged borrowers to assume difficult mortgages in the belief they would be able to quickly refinance at more favorable terms. […] However, once interest rates began to rise and housing prices started to drop moderately in 2006–2007 in many parts of the U.S., refinancing became more difficult. Defaults and foreclosure activity increased dramatically as easy initial terms expired, home prices failed to go up as anticipated, and ARM interest rates reset higher. Falling prices also resulted in 23% of U.S. homes worth less than the mortgage loan by September 2010, providing a financial incentive for borrowers to enter foreclosure."Subprime Mortgage Crisis," *Wikipedia*, accessed June 28, 2011, http://en.Wikipedia.org/wiki/Subprime_mortgage_crisis.

This quote identifies two forces that influenced the demand for housing in 2007–8. The first was expectations of future home prices. One of the gains from owning a house is the possibility that you can sell it at a higher price in the future. Prior to 2007, there had been a fairly consistent tendency for house prices to increase, but the quote seems to indicate that people began to doubt that this trend would continue. As a consequence, the demand for new homes decreased. The second force in the market for new housing was the availability of credit. Most households buy a new home by obtaining a loan (a mortgage) to cover some of the price of the house. During 2007 and 2008, it became increasingly difficult to obtain a mortgage. This was in contrast to a few years earlier when lending standards were easier, and many households easily qualified for mortgages.

These forces affect market demand. The anticipation of lower home prices in the future implies that fewer individuals will choose to buy a home now. Further, if financing is more expensive, then less housing will be purchased. These effects operate *given* the current price of housing. That is, at any given current price of houses, a smaller quantity of houses is demanded. The market demand curve shifts leftward: at each given price, market demand is lower.

The shift in demand is shown in Figure 19.3.1. Once the demand curve shifts, the market for new houses is no longer in equilibrium. At the original price, there is now an imbalance between supply and demand: at that price, buyers want to purchase fewer homes than sellers wish to sell. To restore equilibrium in the market, there needs to be a reduction in housing prices and a reduction in the quantity of new houses produced. The decrease in production comes about because the lower price of houses makes suppliers less willing to produce houses for the market. The shift in the demand curve leads to a *movement along* the supply curve.

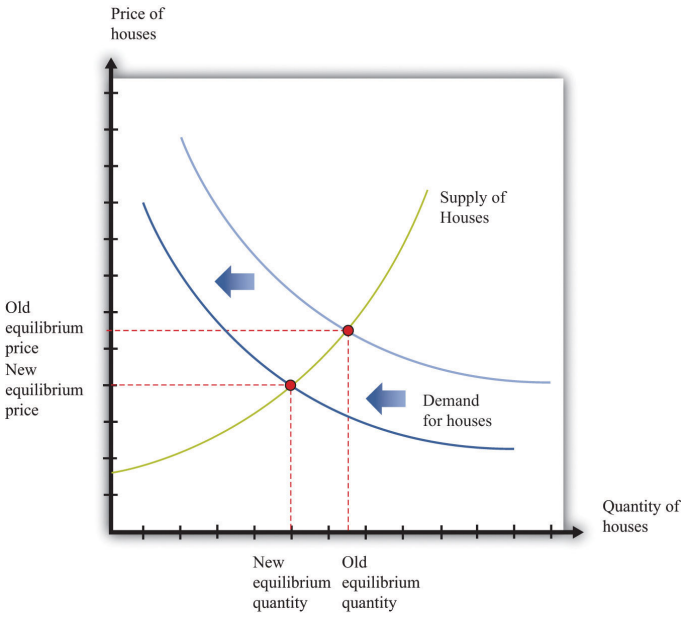


Figure 19.3.1: A Decrease in Demand for Housing. A decrease in demand for houses means that the demand curve shifts leftward, leading to a decrease in both the price of houses and the quantity of houses that are produced and sold.

Shifts in a Curve versus Movements along a Curve

Understanding the distinction between moving along a curve (either supply or demand) and shifting the curve is the hardest part about learning to use the supply-and-demand framework. Journalists and others frequently get confused about this—and no wonder, for it requires practice to learn how to use supply and demand properly.

First, consider the market demand curve. As the price of houses increases, the quantity demanded will decrease. This is a *movement along the market demand curve*. Changes in anything else—anything other than price—that affects the quantity demanded appears as a *shift in the market demand curve*. That is, at each given price, the quantity demanded changes.

Analogously, as the price of houses increases, the quantity supplied will increase. This is a *movement along the market supply curve*. If a change in anything else leads to a change in the quantity supplied, this appears as a *shift in the market supply curve*. That is, at each given price, the quantity supplied changes.

Comparative Statics

The example that we just discussed is an illustration of a general technique used by economists for two purposes. First, we use it to *explain* changes in prices and quantities that we have observed in the past. Second, we use it to *predict* what will happen to market prices and quantities in the future. The technique is called **comparative statics**.

Toolkit: Section 31.16

Comparative statics is a technique that allows us to describe how market equilibrium prices and quantities depend on exogenous events. As such, much of economics consists of exercises in comparative statics. In a comparative statics exercise, you must do the following:

1. Begin at an equilibrium point where the quantity supplied equals the quantity demanded.
2. Based on a description of an event, determine whether the change in the exogenous factor shifts the market supply curve or the market demand curve.
3. Determine the direction of this shift.
4. After shifting the curve, find the new equilibrium point.
5. Compare the new and old equilibrium points to predict how the exogenous event affects the market.

The most difficult part of a comparative statics exercise is to determine, from a description of the economic problem, whether it is the supply or demand curve (or both) that shifts. Once you conquer the economics of determining which curve is shifting, then it is

a matter of mechanically using the framework to find the new equilibrium. A comparison of the old and new equilibrium points allows you to predict what will happen to equilibrium prices and quantities following an exogenous change.

Key Takeaways

- Exogenous variables are determined from outside a framework, while endogenous variables are determined within the framework.
- Changes in exogenous variables lead to shifts in market supply and/or market demand curves. These shifts in supply and demand then lead to changes in quantities and prices.
- Comparative statics is a technique that describes how changes in exogenous variables influence equilibrium quantities and prices. It is used to answer questions about how markets respond to changes in exogenous variables.

Exercises

1. Name two exogenous variables that might affect the equilibrium outcome in the market for used cars.
2. Draw the market for housing when there is a decrease in supply *and* a decrease in demand. What happens to the price? Why can you not say for sure what happens to the quantity of houses bought and sold?

This page titled 19.3: Comparative Statics- Changes in the Price of Housing is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

19.4: Three Important Markets 19.4: Three Important MarketsEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What is the credit market, and what determines the interest rate?
2. What is the labor market, and what determines the real wage?
3. What is the foreign currency market, and what determines the exchange rate?

The financial crisis of 2008 began in the housing market. But as the excerpts at the beginning of this chapter make clear, its effects rapidly spread beyond that market. Those excerpts talked of credit, jobs, and the impact of the crisis on foreign countries. We now look at the knock-on effects of the crisis and, in the process, describe three key macroeconomic markets: the credit market, the labor market, and the foreign exchange market.These markets are used in several places in the book. In particular, we look at labor in Chapter 23, and credit and foreign exchange in Chapter 24.

The Credit Market

A **credit market (or loan market)** is a market in which credit is extended by lenders to borrowers. These credit arrangements, also called loans, are a specific kind of contract. A simple credit contract specifies three things: (1) the amount being borrowed, (2) the date(s) at which repayment must be made, and (3) the amount that must be repaid.Of course, since credit contracts are legal documents, lots of other details will be written into the contract as well. Here we focus on the most important features of the contract.

To be specific, suppose you go to your bank to inquire about a loan for $1,000, to be repaid in one year. In this case the lender—the bank—is a supplier of credit, and the borrower—you—is a demander of credit. The higher is the repayment amount, the more attractive this loan contract will look to the bank. Conversely, the lower is the repayment amount, the more attractive this loan contract looks to you. The relationship between the current price and the future repayment can be summarized in a single number, known as the **nominal interest rate**.

Toolkit: Section 31.24

The nominal interest rate is the number of additional dollars that must be repaid for every dollar that is borrowed. It is generally specified in annual terms; that is, it is the amount that must be paid per year.

For the one-year loan we are considering,

For example, suppose the repayment amount is $1,050. Then the left-hand side of this expression is 1,050/1,000 = 1.05. It follows that the nominal interest rate is 0.05, or 5 percent.

Financial markets are typically good examples of competitive markets. Loans are homogeneous, and there are potentially many buyers and sellers. So if we imagine that there are lots of banks that might be willing to supply credit, and lots of people like you who might demand credit, then we could draw supply and demand curves as in Figure 19.4.1. In this case, the units on the quantity axis are one-year $1,000 loans. The price on the vertical axis is the interest rate, which tells us the amount of the repayment per dollar loaned. The higher the repayment is, the more willing are banks to supply credit, so the supply curve slopes upward. The higher the repayment, the less willing are people to take out these loans, and so the demand curve slopes downward. If the repayment price were acceptable to you, you would “buy” one of these $1,000 loans. The equilibrium nominal interest rate is shown at the crossing of supply and demand.

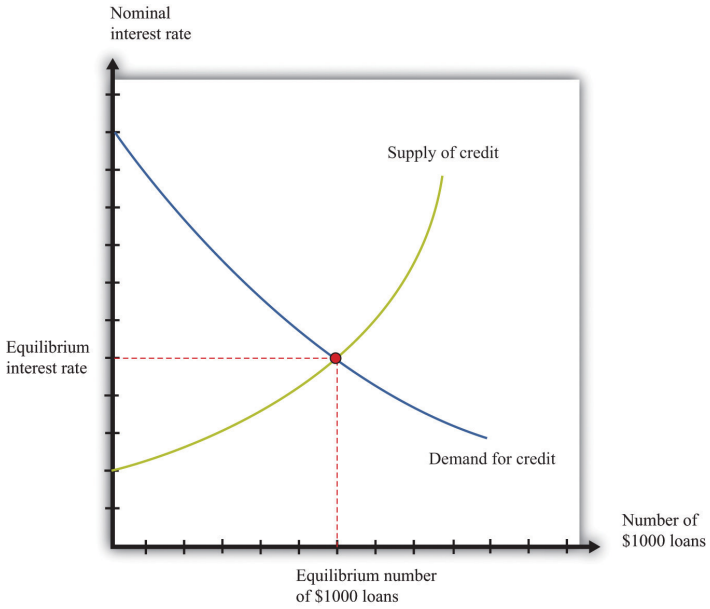


Figure 19.4.1: A Market for $1,000 Loans. In this credit market, lenders offer $1,000 loans to borrowers. The equilibrium nominal interest rate is where the quantity of credit supplied equals the quantity of credit demanded.

The Credit Market in the 2008 Crisis

At the height of the financial crisis of 2008, credit became much more expensive—that is, interest rates increased. Why? As housing prices collapsed in the United States and elsewhere, a substantial number of mortgage loans became *nonperforming*. This means that borrowers were unable or unwilling to repay these loans and defaulted on them instead. In addition, because banks had sold and resold some of these mortgage loans, it was hard to identify which loans would be repaid and which would not. Some financial institutions that were holding a lot of bad loans went bankrupt, and others were in danger of going under as well.

As a consequence, lenders became much more cautious about the types of loans they made—not only in mortgage markets but also throughout the economy. They were more careful about evaluating the likelihood that borrowers would repay their loans. This led to a reduction in the market supply of credit. The reduced supply of loans in the mortgage market was particularly acute. This appears as a leftward shift of the supply curve in Figure 19.4.1. Nominal interest rates increased, and the quantity of mortgages extended decreased. (The full story of what happened in credit markets is more complicated because central banks around the world also took actions to offset these changes and keep interest rates low.)

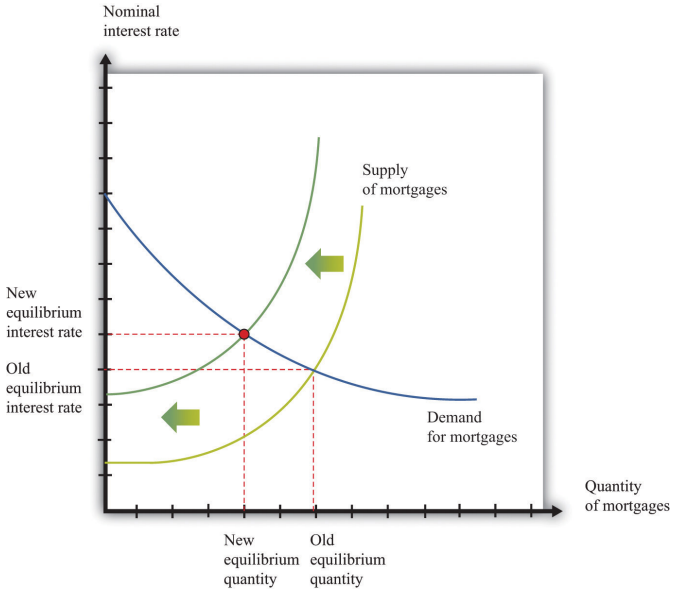


Figure 19.4.2: A Reduction in Supply in the Mortgage Market. As lenders became more cautious about making loans, the supply of mortgage loans shifted leftward. Interest rates in the economy increased, and the quantity of mortgages decreased.

Nominal Interest Rates and Real Interest Rates

Mortgage rates and other interest rates are based on underlying dollar amounts; the interest rate tells you how many dollars borrowers must pay to lenders for each dollar that they borrow. Because they are based on dollar amounts, they are called nominal interest rates. When you see a mortgage rate quoted by a bank or a rate on a credit card, it is a nominal rate.

The nominal rate does not tell us the true cost of borrowing, or return on lending, when there is inflation in an economy. For example, suppose that the nominal interest rate is 5 percent, but inflation is also 5 percent. If you took out a $1,000 loan, you would have to pay back $1,050 next year. But that $1,050 would buy exactly the same amount of **real gross domestic product (real GDP)** next year as $1,000 does this year—that is what it means to have 5 percent inflation. So, in terms of actual goods and services, you have to pay back the same amount that you borrowed. The **real interest rate**—that is, the interest rate corrected for inflation—is zero.

Toolkit: Section 31.8

The **Fisher equation** is a formula for converting from nominal interest rates to real interest rates, as follows:

$$real\ interest\ rate\ \approx\ nominal\ interest\ rate\ -\ inflation\ rate.$$

The real interest rate gives the true cost of borrowing and lending; it is the real interest rate that actually matters for the decisions of savers and borrowers. We derive the Fisher equation more fully in Chapter 24. That doesn't mean, by the way, that our previous two diagrams were incorrect because they used the nominal interest rate. Provided that the inflation rate doesn't change, a comparative static exercise using the nominal interest rate will give you exactly the same conclusion as one using the real interest rate.

Individual Credit Markets and the Aggregate Credit Market

We have described a market for a particular kind of loan, but more generally we know that there are all kinds of different ways in which credit is offered in an economy. Households borrow from banks to buy houses or cars. Households and firms make purchases using credit cards. Firms borrow from financial institutions to buy new equipment. The government borrows to finance its spending, and so on. There is a very large number of credit markets in the economy, each offering a different kind of credit, and each with its own equilibrium interest rate.

These different credit markets are linked because most households and firms buy or sell in more than one market. Financial institutions in particular trade in large numbers of different credit markets. For much of what we do in macroeconomics, however, the distinctions among different kinds of credit are not critical, and it is sufficient to imagine a single aggregate credit market and a single real interest rate. In Chapter 24, we look in more detail at the different kinds of credit—and the associated different interest rates—that we see in an economy. We also investigate in more detail how these markets are linked together. Figure 19.4.3 shows the credit market for an entire economy. This is the market where all the savers in the economy bring funds to financial intermediaries, who then lend those funds to firms, households, and governments. The supply of credit increases as the interest rate increases. As the interest rate increases, other things being equal, households will generally save more and thus supply more to the credit market. The quantity of credit demanded decreases as the interest rate increases. When it is expensive to borrow, households and firms will borrow less.

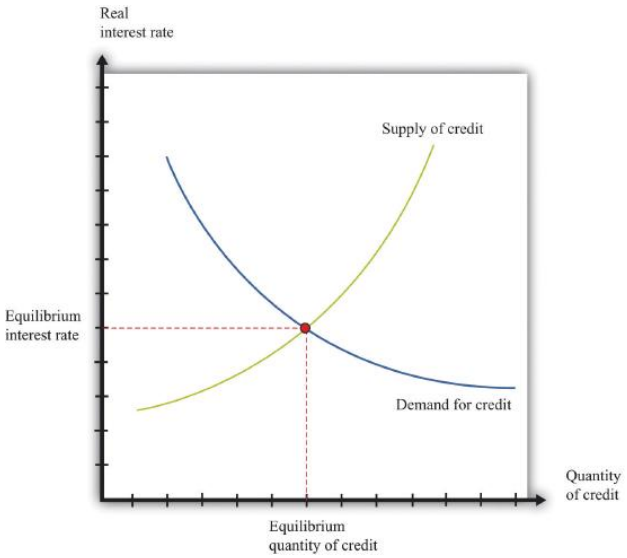


Figure 19.4.3: The Aggregate Credit Market. In the credit market, the equilibrium real interest rate is where the quantity of credit supplied equals the quantity of credit demanded.

Toolkit: Section 31.24

The credit market brings together suppliers of credit, such as households who are saving, and demanders of credit, such as businesses and households who need to borrow. The real interest rate is the price that brings demand and supply into balance. At the equilibrium interest rate, the amount of credit supplied and the amount of credit demanded are equal.

Two of the most important players in the credit market are the government and the monetary authority. If the US federal government borrows more, this shifts the demand for credit outward and increases the interest rate. (Notice that the government is a big player in this market, so its actions affect the interest rate.) The monetary authority, meanwhile, buys and sells in credit markets to influence interest rates in the economy.We study the actions of the Federal Reserve and other monetary authorities in Chapter 25. In the 2008 crisis, the Federal Reserve Bank, which is the monetary authority in the United States, took many actions to increase the supply of credit and ease the problems in the credit market.

The Labor Market

The story about the housing market in the United Kingdom at the beginning of this chapter contained some dire predictions about employment:

The lack of spending in these areas will hit employment, with some analysts forecasting that the construction sector alone could see a loss of up to 350,000 jobs within the next five years.

To understand this prediction, we need to look at another market—the **labor market**.

In the markets for goods and services, the supply side usually comes from firms, and the demand side comes from households. In the labor market, by contrast, firms and households switch roles: firms demand labor, and households supply labor. Supply and demand curves for construction workers are shown in Figure 19.4.4. Here the price of labor is the hourly **real wage** that is paid to workers in this industry.

Toolkit: Section 31.3

The real wage is the wage corrected for inflation. To obtain the real wage, simply divide the wage in dollars—the nominal wage—by the price level:

The individual demand for labor by firms comes from the fact that workers' time is an input into the production process. This demand curve obeys the law of demand: as the real wage increases, the quantity of labor demanded decreases. At a higher real wage, a firm will demand less labor services (by hiring fewer workers and/or reducing the hours of workers) and will respond to the higher labor cost by reducing production.

Workers care about the real wage because it tells them how much they can obtain in terms of goods and services if they give up some of their time. The supply of labor comes from households who allocate their time between work and leisure activities. In Figure 19.4.4, the supply of labor is upward sloping. As the real wage increases, households supply more labor because (1) higher wages induce people to work longer hours, and (2) higher wages induce more people to enter the labor force and look for a job.

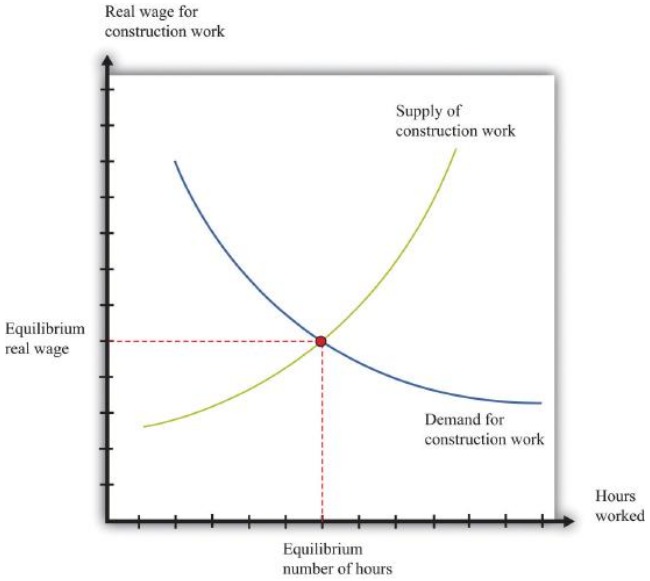


Figure 19.4.4: Equilibrium in the Market for Construction Workers. This picture shows the supply of and demand for hours of work in the construction industry.

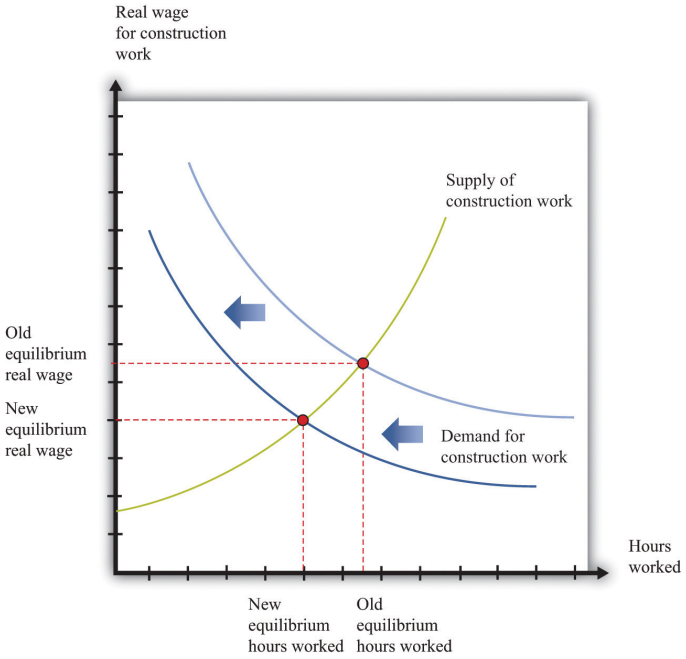


Figure 19.4.5: A Decrease in Demand for Construction Workers. Because builders are building fewer houses, they hire fewer construction workers, causing the labor demand curve to shift leftward.

The Labor Market in the 2008 Crisis

In the United Kingdom, there was a leftward shift in demand for housing (just like we showed in Figure 19.4.1). The response of homebuilders to such a shift is to build fewer homes and, therefore, demand less labor. As a result, there is a leftward shift in the demand curve for construction workers. Based on the supply-and-demand framework, we predict both lower wages and a reduction in employment in the construction sector of the economy, as shown in Figure 19.4.5.

Similar reductions in demand for labor occurred in the United States and many other countries around the world. There was a consequent reduction in employment and an increase in unemployment. The crisis was not restricted just to financial markets, in

other words. It had consequences for the “real” economy as well.

Individual Labor Markets and the Aggregate Labor Market

Because there are many different jobs and many different kinds of workers, there is no single labor market and no single wage. Instead, you can think of there being many different labor markets just as there are many different credit markets. Like different credit markets, different labor markets are linked: households may participate in more than one labor market, and most firms purchase many different kinds of labor. As with the credit market, we sometimes look at the market for a particular kind of labor and the economy as a whole. Most of the time in macroeconomics, it is sufficient to think about an aggregate labor market, as shown in Figure 19.4.6. In Chapter 23, we pay more attention to the fact that workers and jobs are not all identical. As the real wage increases, households supply more hours, and more households participate in the labor market. For both of these reasons, as the real wage increases, the quantity of labor supplied also increases. Labor demand comes from firms. As the real wage increases, the cost of hiring extra labor increases, and firms demand fewer labor hours. That is, the firm’s labor demand curve is downward sloping.

Toolkit: Section 31.3

The labor market is the market in which labor services are traded. The supply of labor comes from households. At the equilibrium real wage, the number of hours supplied and the number of hours demanded are equal.

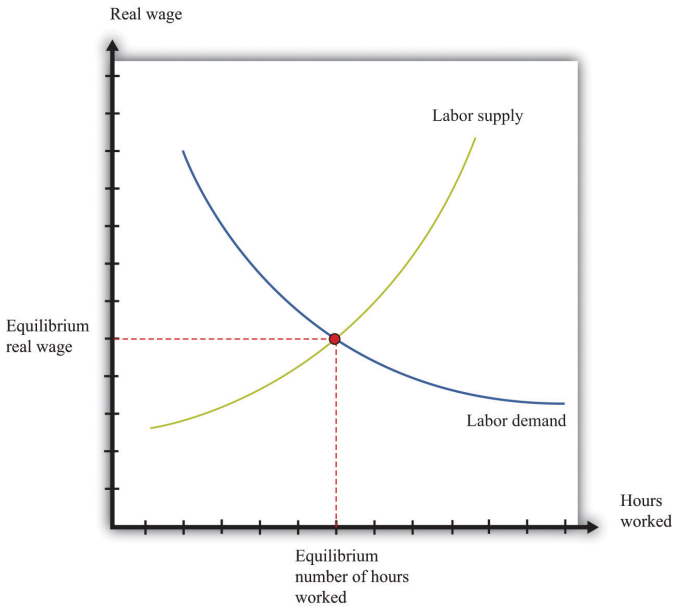


Figure 19.4.6: Equilibrium in the Labor Market. The labor market is the market in which firms hire workers. The equilibrium real wage is the price where the quantity of labor supplied equals the quantity of labor demanded.

The Foreign Exchange Market

The excerpts at the beginning of this chapter reveal that the financial crisis also impacted other countries. For example, we included an excerpt about the effects of the crisis on the value of a dollar and also an excerpt about exports from China. We could have also cited effects of the crisis on other countries: for example, India’s information technology sector and Canada’s lumber industry were both affected. To understand the transmission of the crisis to other countries, we have to learn about another market—the market where different currencies are bought and sold.

If you travel abroad, you must acquire the currency used in that region of the world. For example, if you take a trip to Finland, Russia, and China, you will buy euros, rubles, and yuan along the way. To do so, you need to participate in various **foreign exchange markets**.

Toolkit: Section 31.20

The foreign exchange market is the market where currencies are traded. The price in this market is the price of one currency in terms of another and is called the **nominal exchange rate**.

Dollars are supplied to foreign exchange markets by US households, firms, and governments who wish to purchase goods, services, or financial assets that are denominated in the currency of another economy. For example, if a US auto importer wants to buy a German car, it must sell dollars and buy euros. As the price of a dollar increases, the quantity supplied of that currency will increase.

Foreign currencies are supplied by foreign households, firms, and governments that wish to purchase goods, services, or financial assets (such as stocks or bonds) denominated in the domestic currency. For example, if a Canadian bank wants to buy a US government bond, it must sell Canadian dollars and buy US dollars. The law of demand holds: as the price of a dollar increases, the quantity of that currency demanded decreases.

Figure 19.4.7 shows an example of a foreign exchange market: the market in which euros are bought with and sold for US dollars. The horizontal axis shows the number of euros bought and sold on a particular day. The vertical axis shows the exchange rate—the price of a euro in dollars. This market determines the dollar price of euros just like the gasoline market determines the dollar price of gasoline.

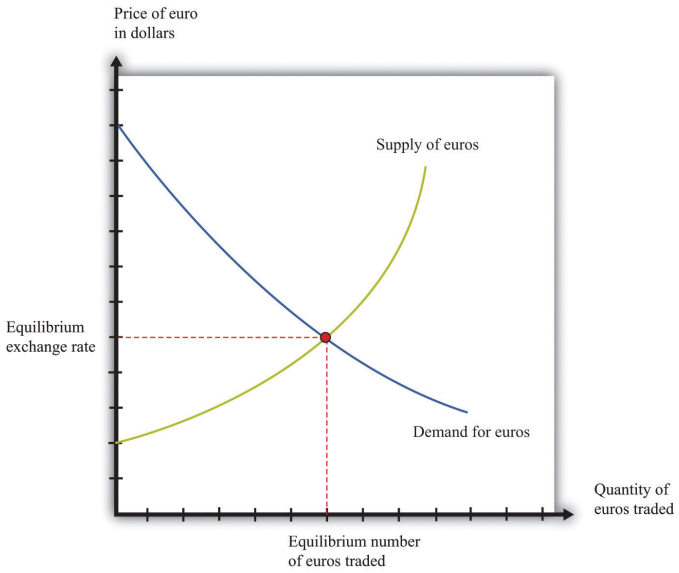


Figure 19.4.7: Equilibrium in the Foreign Exchange Market Where Dollars and Euros Are Exchanged. Currencies are traded in foreign exchange markets, such as the market shown here in which dollars and euros are exchanged. The equilibrium exchange rate is the price where the quantity of euros supplied equals the quantity of euros demanded.

On the supply side, there are households and firms in Europe who want to buy US goods and services. To do so, they need to buy dollars and, therefore, must supply euros to the market. This supply of euros need not come only from European households and firms. Anyone holding euros is free to sell them in this market. On the demand side, there are households and firms who are holding dollars and who wish to buy European goods and services. They need to buy euros.

There is another source of the demand for and the supply of different currencies. Households and, more importantly, firms often hold assets denominated in different currencies. You could, if you wish, hold some of your wealth in Israeli government bonds, in shares of a South African firm, or in Argentine real estate. But to do so, you would need to buy Israeli shekels, South African rand, or Argentine pesos. Likewise, many foreign investors hold US assets, such as shares in Dell Inc. or debt issued by the US government. Thus the demand for and the supply of currencies are also influenced by the asset choices of households and firms. In practice, banks and other financial institutions conduct the vast majority of trades in foreign exchange markets.

As well as households and firms, monetary authorities also participate in foreign exchange markets. For example, the US Federal Reserve Bank monitors the value of the dollar and may even intervene in the market, buying or selling dollars in order to influence the exchange rate.

Foreign Exchange Markets in the 2008 Crisis

One of the articles we used to open this chapter dealt with changes in the value of the dollar in the fall of 2008. The article pointed out that the dollar was getting stronger relative to other currencies, such as the euro. This means that the price of a dollar in euros was increasing or, equivalently, the price of a euro in dollars was decreasing. In fact, the euro price of a dollar was about 0.67 in late September 2008; the price increased to nearly 0.81 by late October and then decreased again through December 2008.

We can use the foreign exchange market to understand these events. Figure 19.4.8 shows the dollar market for euros once again. The increase in the value of the dollar discussed in the article is seen here as a rightward shift in the supply of euros, which decreases the value of the euro and—equivalently—increases the value of the dollar.

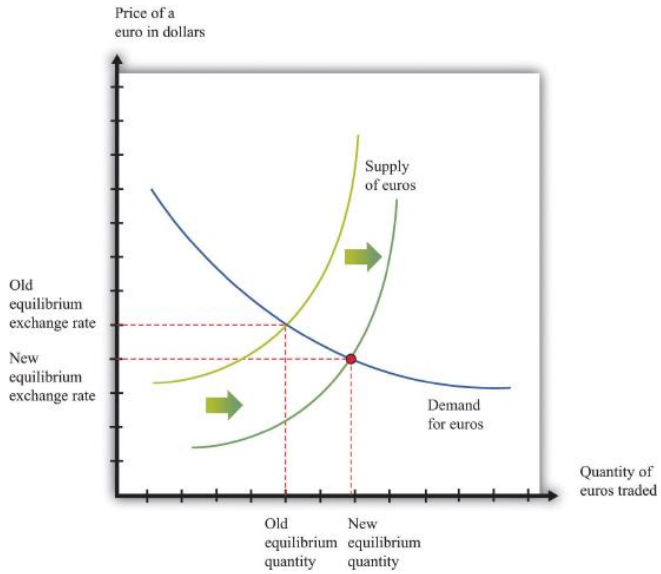


Figure 19.4.8: Comparative Statics in the Euro Market. The rightward shift in the supply of euros leads to a decrease in the price of a euro in terms of dollars.

There are two consequences of this shift in the supply curve. First, the shift in supply decreases the dollar price of the euro. So people in the United States who are planning to visit, say, France will find that they can obtain more euros for a given amount of dollars. Second, the quantity of euros actually bought and sold is higher. This is not inconsistent with the lower dollar price of a euro since the supply curve shifts along the demand curve for euros.

Individual Foreign Exchange Markets and the Aggregate Foreign Exchange Market

We sometimes look at an individual exchange rate (e.g., dollar-euro) by thinking of the market where dollars are exchanged for euros. However, there are many different currencies that are exchanged for the US dollar. There are markets where dollars are exchanged for British pounds, Japanese yen, and so on. We can combine these into an aggregate foreign exchange market. Think of this as being the market where US dollars are bought with and sold for all the other currencies in the world. In this market, there is an aggregate exchange rate, which you can think of as an average of the exchange rates in all the individual markets.More precisely, you should think of a weighted average. Because the United States trades much more with Canada than with, say, South Africa, movements in the US dollar–Canadian dollar exchange rate matter more than movements in the US dollar–South African rand exchange rate. Chapter 24 has more on exchange rates. We show this market in Figure 19.4.9.

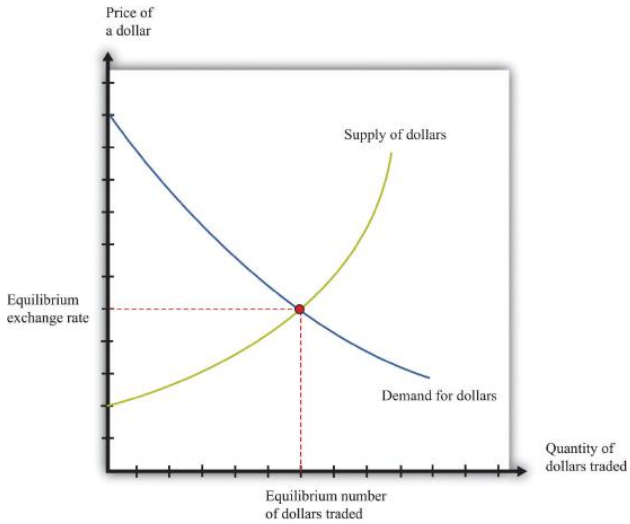


Figure 19.4.9: Foreign Exchange Market Equilibrium. Currencies are traded in foreign exchange markets, such as the market shown here in which dollars are bought and sold.

Key Takeaways

- The credit market brings together the suppliers of credit (households) with those who are demanding credit (other households, firms, and the government). The interest rate adjusts to attain a market equilibrium.
- The labor market is where labor services are traded. Households supply labor, and firms demand labor. The real wage adjusts to attain a market equilibrium.
- The foreign exchange market brings together demanders and suppliers of foreign currency. The exchange rate, which is the price of one currency in terms of another, adjusts to attain a market equilibrium.

Exercises

Figure 19.4.7 shows the market where euros are bought and sold using dollars. We could equivalently think about this as the market where dollars are bought and sold using euros. Draw the graph for this market. How are the supply and demand curves in the two markets related to each other?

This page titled 19.4: Three Important Markets is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

19.5: Linkages across Markets 19.5: Linkages across MarketsEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. How are the markets for goods, labor, credit, and foreign currency linked?
2. How do we use those links to understand the crisis that began in 2008?

In 19.3 Section "Three Important Markets", we talked about the markets for credit, labor, and foreign exchange. We explained that we sometimes look at individual examples of these markets and sometimes at versions of these markets that apply to the economy as a whole.

But the story of the economic crisis in 2008 was not about a single market. Instead, what started as a problem in the US mortgage market was felt in the housing market in England, the labor market in China, the foreign exchange market in Europe, and many other markets. These different markets are connected; in this section, we explore these linkages. We do so through the **circular flow of income**, shown in Figure 19.5.1. That model of the economy reveals the linkages across markets that the global financial crisis made so evident.

Toolkit: Section 31.27

You can review the circular flow of income in the toolkit.

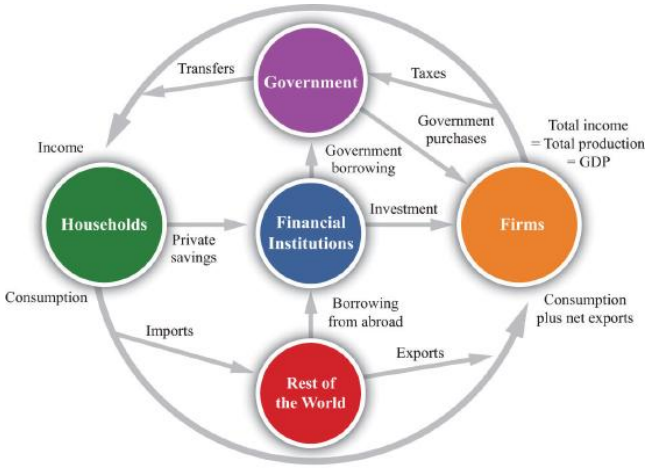


Figure 19.5.1: The Circular Flow of Income

We know from the circular flow that the production of goods and services generates income in an economy. Some of that income is paid to the government in the form of taxes, but the rest finds its way to households. Much of the flow of dollars from firms to households takes place through the labor market because firms demand labor to produce goods. If firms are producing large quantities of goods and services, then they demand lots of labor, and income from the sale of labor services in the economy is high.

Some of the income that households earn from selling labor services is saved. There is therefore a link, through the household sector, between the labor market and the credit market. So we can follow a connection from the production of goods and services to the supply of credit: if firms produce more, they generate more labor income, so there is more saving supplied by households to the credit market. There is also a link from the markets for goods and services to the *demand* for credit: firms borrow to purchase investment goods.

These markets are also linked—directly or indirectly—to foreign exchange markets. Whenever firms purchase imported goods, such as oil, this generates a demand for foreign exchange. When firms expand output, demand more labor, and so generate additional household income, households spend some of this income on imports, again generating a demand for foreign exchange. When households and firms in other countries want to buy our goods and services, that generates a supply of foreign exchange. And many transactions in credit markets also generate a demand for or supply of foreign exchange.

Comparative Statics in an Interconnected World

We could go on, but the point should be clear: the markets in every economy are intimately interconnected. This has a critical implication for our study of macroeconomics, which is that it both complicates and enriches our comparative static analyses. When a shift in supply or demand in one market affects the equilibrium price and quantity in that market, there are changes in other markets as well.There is a second, more abstract implication: we have to worry about whether *all the markets in the economy are in equilibrium at the same time*. In our analyses, we have looked at only one market at a time. But we now know that the outcome in one market (for example, the real wage) can affect supply and demand in other markets (for example, the supply of credit). In advanced studies in economics, we use complicated mathematics to see if there are prices that are consistent with all the markets being in equilibrium at once. The bottom line is good news: we can usually be confident that there is an equilibrium for all markets. But because this is such an advanced area of economics, we do not worry about it further in this book. In this section, we show how these interactions across markets help us understand the propagation of the 2008 crisis from the US housing market to the economies of the world. We have already hinted at some of these linkages, but now we make them more explicit.

Housing and Credit Markets in the 2008 Crisis

The story began with the first comparative static example that we looked at: a leftward shift in demand for housing. Potential buyers of houses started worrying that the *future* price of houses would decrease. This made people more reluctant to buy houses. Meanwhile, a tightening of lending standards made it harder for people to obtain loans. Both of these caused the demand for housing to shift leftward. Part (a) of Figure 19.5.2, which we already saw earlier in the chapter, shows us that this led to a decrease in both the price and the quantity of houses.

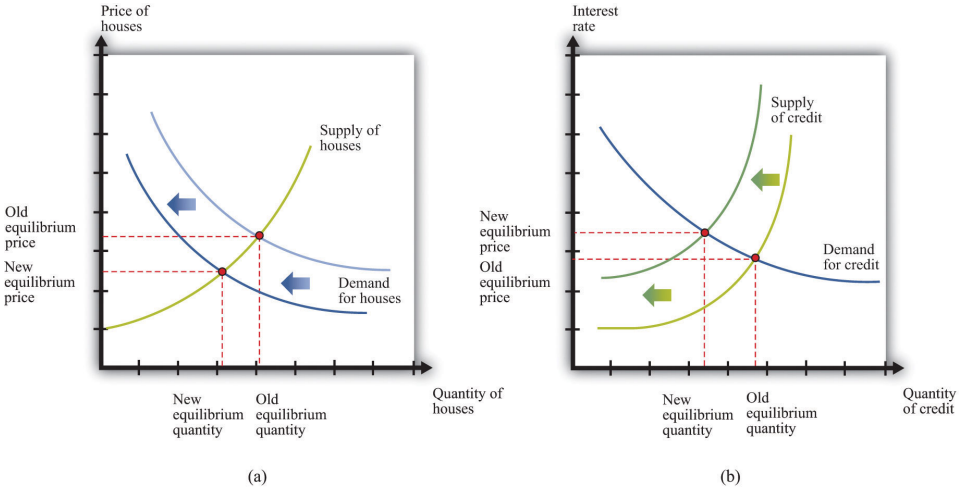


Figure 19.5.2. A decrease in demand for housing led to a decrease in supply of credit. (a) Worsening expectations about *future house prices, together with tighter lending conditions, led to a decrease in demand for housing. (b) In the credit market, banks and other lending institutions found themselves with bad debt, so the supply of credit decreased.*

Part (b) of Figure 19.5.2 also appeared earlier in the chapter. The decrease in housing prices, combined with the complicated way in which mortgages had been sold and resold by financial institutions, meant that many financial institutions found themselves in trouble. Some went bankrupt. This made financial institutions cautious about lending to each other, so the supply of credit shifted to the left. Interest rates rose. (Interest rates in the crisis were also affected by the actions of the US Federal Reserve and other monetary authorities around the world.We discuss such policies in detail when we return to the crisis in Chapter 30.)

A Shift in the Supply of Goods

If you run a business, you often have to rely on credit (loans) to finance the purchase of your inputs into the production process. For example, suppose you run a boutique clothing store. You have to buy the clothes to put on display first, and you get your revenues only when you sell the clothes. Weeks or even months may pass between the time you incur your costs and the time you get your revenues. Unless you have the funds available to buy all your stock up front, you will need to borrow. The same is true in many other businesses. Firms regularly take out short-term loans to pay for some of their costs of operation.

When interest rates increase, businesses see their costs increase. Higher costs make it less profitable to produce at any given price, so most businesses cut back on their production. Some may even leave the market altogether. As a consequence, the supply curve for most goods and services shifts leftward, as shown in Figure 19.5.3. We see that the equilibrium price increases, and the

equilibrium quantity decreases. Going back to an individual producer, what does this mean? The producer sees costs increase. In the new equilibrium, the producer also obtains a higher price. However, the increase in price is not as big as the increase in cost.

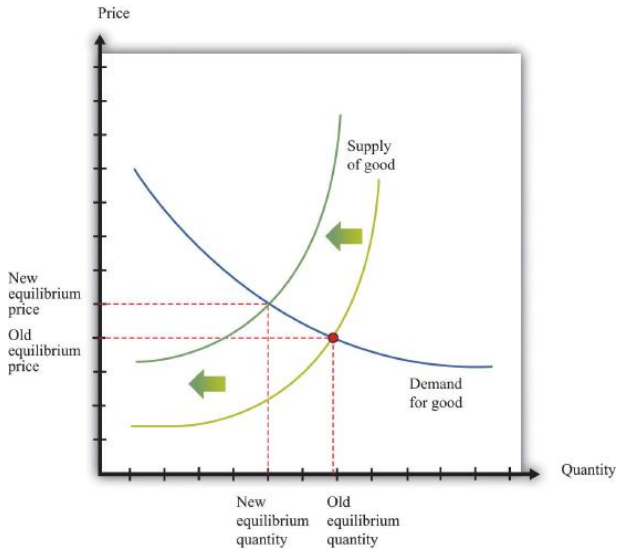


Figure 19.5.3. Higher interest rates lead to higher prices and lower quantities for most goods and services. Higher interest rates increase the cost of doing business, so the supply curve for a typical good or service shifts leftward.

A Shift in Demand for Labor

The effect of the higher interest rates on the output decisions of firms also leads them to demand less of all their inputs, including labor. Decreases in production lead to decreases in labor demand, as shown in Figure 19.5.4. In turn, decreases in wages and employment (more generally, a decrease in income) lead to decreased demand for goods.

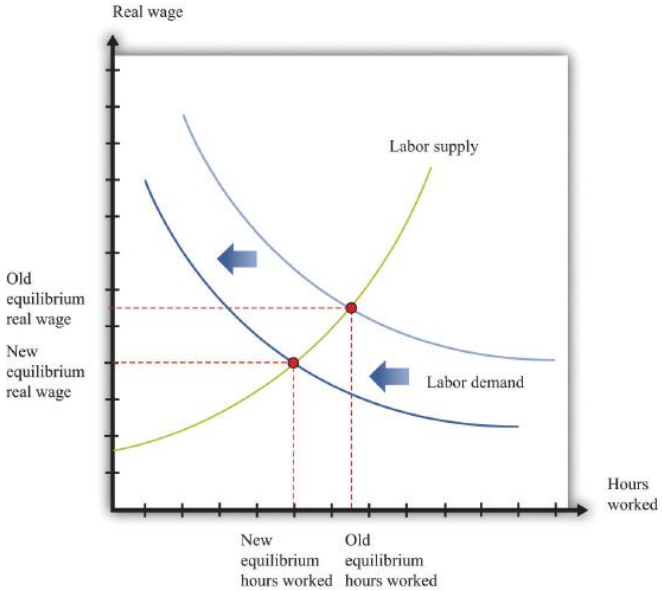


Figure 19.5.4: A Decrease in Demand for Labor. A decrease in demand for labor causes the labor demand curve to shift leftward.

A Shift in Demand for Goods

Notice the connection back and forth between households and firms. As firms reduce their demand for labor services, less income flows to households. This reduction in income leads to a reduction in the demand for goods and services, leading firms to reduce output and employment even further. The interaction between income and spending on goods and services can lead to much larger reductions in output and employment than the original shift in demand in the original market (in this case, the housing market). This means that Figure 19.5.3 does not tell the whole story of goods markets. That figure shows the effects of interest rates on the

supply of goods but does not include the reduction in demand stemming from the interaction of income and spending in the circular flow. Figure 19.5.5 completes the story by adding the shift in demand.

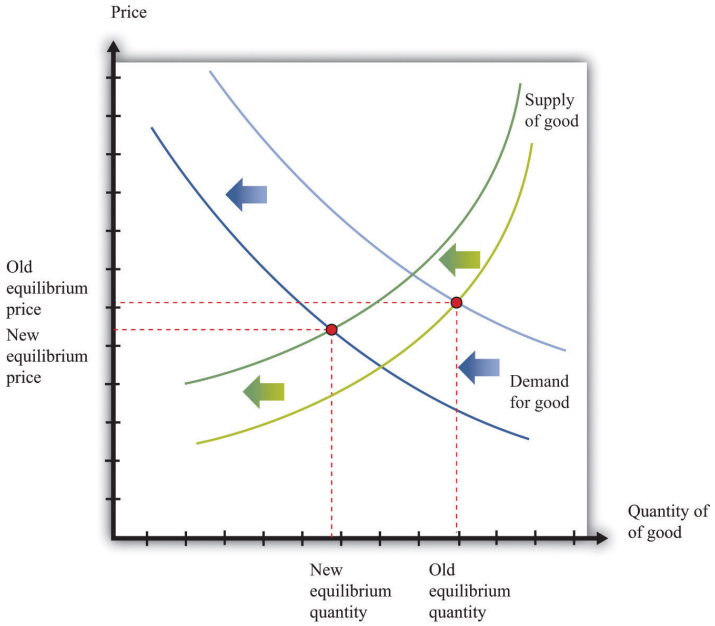


Figure 19.5.5. Higher interest rates lead to a leftward shift in supply, and lower income leads to a leftward shift of demand, resulting in lower quantities for most goods and services. Higher interest rates increase the cost of doing business. Lower income decreases the demand for goods and services.

The following from 2008 story illustrates such a connection across markets.

How to Tell Business Is Cutting Back

From fewer shoe shines to a slowdown in corporate art purchases, subtle bellwethers can help take the temperature of business activity.

Nelson Villanova doesn't need to watch the stock market indexes…or gross domestic product to gauge the health of the economy. He just has to look down. If he sees scuffed shoes, then he knows things are bad.

Villanova, general manager of Eddie's Shoe Repair in New York's Grand Central Terminal, has seen business drop 25% to 30% since August. The 15-year-old company employs 40 people across five locations in the sprawling train station, shining and repairing shoes and luggage. But lately, selling $4 shines seems to be as hard as unloading mortgage-backed securities.

[…]John Tozzi, "How to Tell Business Is Cutting Back," *Bloomberg BusinessWeek*, October 21, 2008, accessed June 28, 2011, http://www.businessweek.com/smallbiz/content/oct2008/sb20081020_372369.htm?chan=top$+$news_top+news+index+-+temp_small+ business.

Figure 19.5.6 shows the shoeshine market. Traders working on Wall Street started purchasing fewer shoeshines. This was not because shoeshines became more expensive. Rather, it was a shift in the demand for shoeshines because these traders saw that their incomes were decreasing.

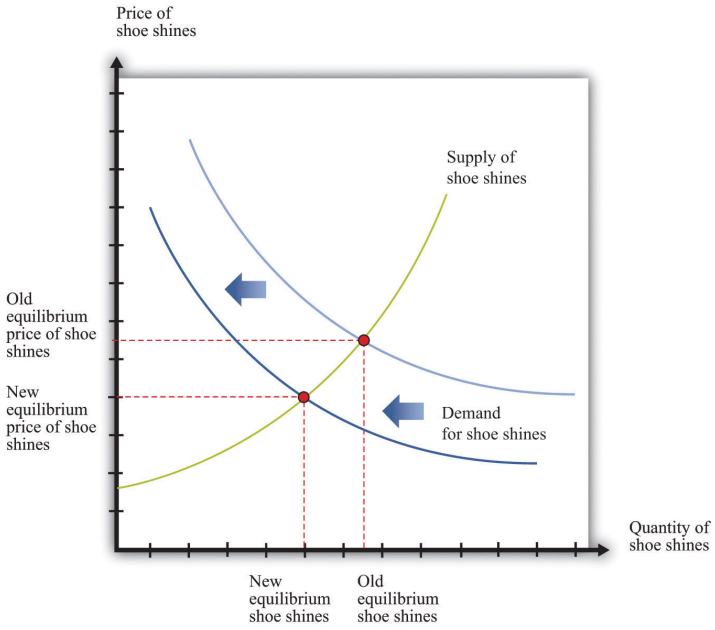


Figure 19.5.6: *A Decrease in Demand for Shoeshines*. A decrease in income leads to a decrease in demand (a leftward shift) for shoeshines.

Trade Flows and a Shift in the Demand for Foreign Exchange

One of the excerpts we used to introduce this chapter touched on the effects of the crisis on exports from China. We now broaden our discussion to include those effects as well. Looking back at Figure 19.5.4, recall that part of household spending goes toward the purchase of goods and services produced in other countries. A significant fraction of imports to the United States come from China. China also sells goods and services to Japan, Europe, and most of the world.

When demand from these economies slumps, as it did in 2008, exports from China also decrease. Since exports are a part of overall spending, this leads firms in China to cut back their production and employment. Thus the Chinese economy was also slowed down by the effects of the financial crisis.

The reduced demand for imports has another effect. Because the demand for foreign currency is partly motivated by the desire to buy goods from that country, a decrease in the import of Chinese goods to the United States and other countries leads to a decrease in demand for the Chinese yuan. There is a leftward shift in the demand for that currency and thus a lower price in dollars. (As with all comparative static exercises, this assumes that nothing else is changing to offset these effects on the demand for the yuan.)

The **current account balance** is (roughly speaking) the difference between the value of exports and imports of goods and services. A country has a current account surplus if the value of exports of goods and services exceeds the value of its imports. A country has a current account deficit if the value of imports of goods and services exceeds the value of its exports. Looking at the United States and China, one sees very different behavior for the current account.This discussion draws on data from the International Monetary Fund. See Stephan Danninger and Florence Jaumotte, "Divergence of Current Account Balances across Emerging Economies," World Economic Outlook, Chapter 6, accessed June 28, 2011, http://www.imf.org/external/pubs/ft/weo/2008/02/pdf/c6.pdf. In recent years, the United States has run a current account deficit of nearly 5 percent of its gross domestic product (GDP). China, in contrast, has run a current account surplus of about 6.1 percent of its GDP since 2002.

The reduced demand for imports from China has an effect on the current account balance of China. We would expect to see a reduction in the current account surplus of China due to the reduction in economic activity of its trading partners.

You might also wonder how the persistent deficits of the United States are paid for. When a country runs a current account deficit, it is borrowing from other countries. This is just like a household that pays for consumption above its income by means of borrowing. The rules of national income accounting tell us that the flows in and out of each sector must always be in balance. If we look at the flows in and out of the foreign sector we see that

$$borrowing\ from\ abroad\ =\ imports\ -\ exports$$

or

$$lending\ to\ abroad\ =\ exports\ -\ imports.$$

Net exports (sometimes called the trade surplus) equal exports minus imports. So lending to other countries equals net exports.

The circular flow of income tells us something powerful: whenever we import more than we export, we *must*, on net, be borrowing from abroad. On reflection, this is not so surprising. Other countries are giving us more goods and services than we are giving to them. This is not done out of generosity; they do so because they expect to be repaid at some point in the future. If we export more than we import, then this flow goes in the other direction, and we are lending to abroad.

Both China and the United States trade with many other countries, so this pattern of trade holds true bilaterally (that is, between them) as well. China has run systematic current account surpluses with the United States, meaning that China is lending to the United States. Those loans take many forms, with commentators highlighting Chinese purchases of US government debt. US Secretary of State Hilary Clinton alluded to this connection between the two economies during a visit to China in early 2009.

US Secretary of State Hillary Clinton yesterday urged China to keep buying US debt as she wrapped up her first overseas trip, during which she agreed to work closely with Beijing on the financial crisis.

[…]

By continuing to support American Treasury instruments the Chinese are recognizing our interconnection…"Keep Buying US Treasury Bills, Clinton Urges China," *Taipai Times*, February 23, 2009, accessed June 28, 2011, www.taipeitimes.com/News/front/print/2009/02/23/2003436802.

The Crisis of 2008: A Brief Summary

The crisis began with a reduction in the demand for houses and a consequent decrease in the value of houses. This reduced the value of assets, particularly mortgage-backed securities, and meant that the supply of credit in the economy shifted inward. The consequence was higher interest rates and reduced credit. Since many firms in the economy borrow to finance production, the increased interest rates increased their marginal costs of production. Supply curves throughout the economy shifted inward, leading to lower output. Firms needed fewer workers, so there was a reduction in employment.

The spread to other countries came through a couple of avenues. First, households and firms in other countries were one source of credit to the US economy. When asset prices decreased, the portfolios of foreign banks were also adversely affected. This led to higher interest rates and lower output in those countries. In addition, as the US economy went into recession, it purchased fewer imports from other countries. This led to lower production in those countries.

Our description of the crisis is of necessity a simple one. We have neglected many details, and we have not discussed how government policies also affected interest rates and the demand for goods and services. Later chapters in the book provide more tools for understanding these aspects of the crisis, so when we return to the topic in Chapter 30, we can provide a more complete analysis of the crisis.

Key Takeaways

- Markets are linked because supply and demand in one market generally depend on the outcomes in other markets. The circular flow of income illustrates some of these connections across markets.
- Although the crisis in 2008 may have started in the housing market, it did not end there. Instead, the crisis impacted markets for labor, credit, and foreign exchange.

Exercises

1. We have explained that increases in interest rates shift the supply of goods leftward, and decreases in incomes shift the demand for goods leftward. Draw diagrams with both shifts at once and show that the quantity definitely decreases, but the price may increase or decrease.
2. Can you think of a good for which the demand curve might shift rightward when incomes decrease?

This page titled 19.5: Linkages across Markets is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

- **Current page** by Anonymous is licensed CC BY-NC-SA 3.0. Original source: https://2012books.lardbucket.org/books/theory-and-applications-of-economics.
- **23.3: Job and Worker Flows** by Anonymous is licensed CC BY-NC-SA 3.0. Original source: https://2012books.lardbucket.org/books/theory-and-applications-of-economics.

19.6: End-of-Chapter Material 19.6: End-of-Chapter MaterialEdit Update titlecancel

In Conclusion

The supply-and-demand framework is almost certainly the most powerful model in the economist's toolkit. Armed with an understanding of this framework, you can make sense of much economic news, and you can make intelligent predictions about future changes in prices.

A true understanding of this framework is more than just an ability to shift curves around, however. It is an understanding of how markets and prices are one of the main ways in which the world is interlinked. Markets are, quite simply, at the heart of economic life. Markets are the means by which suppliers and demanders of goods and services can meet and exchange their wares. Since exchange creates value—because it makes both buyers and sellers better off—markets are the means by which our economy can prosper. Markets are the means by which economic activity is coordinated in our economy, allowing us to specialize in what we do best and to buy other goods and services.

Economists regularly point to these features of markets, but this should not blind us to the fact that markets can go wrong. There are many ways in which market outcomes may not be the most desirable or efficient, as the global financial crisis revealed. In the remainder of this book, we look in considerable detail at all the ways that markets can fail us as well as help us.

Key Links

- International Monetary Fund (IMF): www.imf.org/external/pubs/ft/weo/2008/01/pdf/c3.pdf
- IMF video on its response to the crisis: http://www.youtube.com/watch?v=f0z6nWQfvuA&feature=channel_page
- Federal Reserve Bank of San Francisco: http://www.frbsf.org/publications/economics/letter/2008/el2008-21.html
- New York University Stern blog on the financial crisis: http://sternfinance.blogspot.com/search/label/overview/

exercises

1. What would the impact be on the market demand curve for new homes if there were an increase in the price of old homes?
2. Name two factors that cause market demand curves to shift outward.
3. Fill in the blanks in the following table. What can you say about the missing price in the table?

Price of Chocolate Bar	Household 1's Demand	Household 2's Demand	Market Demand
1	7		22
2		11	16
10	.5	3	3.5
	.75	4	4.75

TABLE 19.6.1: INDIVIDUAL AND MARKET DEMAND

4. If the income levels of all households increase, what happens to the individual demand curves? What happens to market demand?
5. Suppose the price of coffee increases. Household 1 always eats a chocolate bar while drinking coffee. What will happen to Household 1's demand for chocolate bars when the price of coffee increases? Household 2 has either coffee or a chocolate bar for dessert. What happens to Household 2's demand for chocolate bars when the price of coffee increases? What happens to the market demand for chocolate bars when the price of coffee increases?
6. (Advanced) In Figure 19.6.3 we showed the market supply curve for new houses. Suppose that a change in government regulations makes it easier for people to become qualified electricians. What will happen to the supply curve for houses?
7. We said that the equilibrium price and quantity in a market is always positive. More precisely, this is true as long as the vertical intercept of the demand curve is bigger than the vertical intercept of the supply curve. If this is not the case, then the most that any buyer is willing to pay is less than the least any seller is willing to accept. Draw a version of Figure 19.6.5 to illustrate this possibility. How much trade do you expect in this market?
8. Suppose that households become worried about losing their jobs and decide to save more. What happens in the credit market? Do you expect interest rates to increase or decrease?

9. When interest rates decrease, firms find it cheaper to borrow. What do you think happens to the demand for labor? What happens to the real wage?
10. What happens to the value of the US dollar if
    1. foreign investors decide they want to buy more US assets.
    2. there is a recession in other countries that buy goods produced in the United States.
11. What do you think will be the effect on the markets for used homes and apartments if there is a reduction in expected capital gains from owning a new home? The shift in the supply curve came from an increase in the cost of credit. Where might the increase in the cost of credit come from?
12. Think about your hometown as an economy. What does it import (i.e., what goods and services does it purchase from outside the town)? What does it export (i.e., what goods and services are produced in the town and sold outside it)? What about the street you live on—what are its imports and exports?
13. Using supply and demand, explain how an increase in Chinese demand for Australian butter might be one of the factors causing the Australian dollar to appreciate. How might this affect the labor markets in Australia?
14. If oil prices increase, what will this do to the demand for apartments and houses in warm climates? What will happen to housing prices in cold climates? Use supply and demand to illustrate.

Economics Detective

1. Find three news articles that discuss the financial crisis. Which markets are discussed in these articles? Can you use a supply-and-demand picture to help you make sense of anything that is discussed in the articles you find?
2. Find one example of another country where there was a major decrease in housing prices, as in the United States and England. Find another country where housing prices did not seem to be affected.

Spreadsheet Exercise

1. Using a spreadsheet, construct a version of Table 19.2.1 "Market Equilibrium: An Example" assuming that

$$market\ demand\ =\ 50 - 0.005 \times price.$$

   Fill in all the prices (in thousands) from 1,000 to 100,000. What is the equilibrium price and quantity in the market? How would you explain the difference between this equilibrium and the one displayed in Table 19.2.1 "Market Equilibrium: An Example"?

This page titled 19.6: End-of-Chapter Material is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

20: Globalization and Competitiveness

20.1: Five Stories
20.2: The Production of Real GDP
20.3: Labor in the Aggregate Production Function
20.4: Physical Capital in the Aggregate Production Function
20.5: Other Inputs in the Aggregate Production Function
20.6: Accounting for Changes in GDP
20.7: Globalization and Competitiveness Revisited
20.8: End-of-Chapter Material

This page titled 20: Globalization and Competitiveness is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

20.1: Five Stories 20.1: Five StoriesEdit Update titlecancel

We begin this chapter with five stories from around the world.

The United Kingdom

The following is a BBC report on Polish immigration to the United Kingdom.

So You're Polish and Want a Job…

If there was ever any doubt that the UK is in the grips of an extraordinary revolution, then hunt out the migrant worker recruitment fairs that are starting to spring up.

Last month, thousands of young Polish workers turned up at the third recruitment fair hosted by Polish Express, the London-based newspaper for the diaspora, […]

As they queued to enter the hall that was filled to its legal safety capacity, they scribbled away at resumes, going over their pitch time and time again.

Most were in their mid-20s. Some had only recently arrived, having stuffed a few belongings into a backpack, bought a one-way no-frills airline ticket. […]

[A] willingness to do jobs that employers say British workers don't want, was at the heart of the boom, said Bob Owen of Polcat, a Doncaster safety training firm targeting the Polish employees market.

"I must admit it, I have never seen a workforce like the Poles," said Mr Owen. "They want to work, you can see it in their eyes. But here's the thing—they're not in competition with the British workforce—they are finding ways of fulfilling a need that just wasn't being met and that's why they are being welcomed."

[…]See Dominic Casiani, "So You're Polish and You Want a Job," *BBC News*, September 25, 2006, accessed June 28, 2011, http://news.bbc.co.uk/1/hi/uk/5376602.stm.

United Arab Emirates

Figure 20.1.1 is a screenshot from a Dubai government website that promotes business and tourism in Dubai.See "Dubai for Business," Government of Dubai: Department of Tourism and Commerce Marketing, accessed July 27, 2011, http://www.dubaitourism.ae/definitely-dubai/dubai-business. It details many different ways in which Dubai is a desirable place for businesses to locate. For example, the website contains the following:



Figure 20.1.1

Pro-Business Environment

Dubai offers incoming business all the advantages of a highly developed economy. Its infrastructure and services match the highest international standards, facilitating efficiency, quality, and service. Among the benefits are:

- Free enterprise system.
- Highly developed transport infrastructure.
- State-of-the-art telecommunications.
- Sophisticated financial and services sector.
- Top international exhibition and conference venue.
- High quality office and residential accommodation.
- Reliable power, utilities, etc.
- First class hotels, hospitals, schools, and shops.
- Cosmopolitan lifestyle.

The website goes on to talk about benefits such as the absence of corporate or income taxes, the absence of trade barriers, competitive labor and energy costs, and so on.

Vietnam

The following is an extract from the *Taipei Times*, April 9, 2007.

Compal Eyes Vietnam for Factory

Compal Electronics Inc, the world's second-largest laptop contract computer maker, is considering building a new factory in Vietnam.

Compal could join the growing number of Taiwanese electronic companies investing in Vietnam—such as component maker Hon Hai Precision Industry Co—in pursuit of more cost-effective manufacturing sites outside China.

[…]

Compal forecast last month that its shipments of notebook computers would expand around 38 percent to 20 million units this year, from 14.5 million units last year. The company currently makes 24 million computers a year at its factories in Kunshan, China.

Compal, which supplies computers to Dell Inc and other big brands, could lack the capacity to match customers' demand next year if its shipments increase any faster,…

Lower wages and better preferential tax breaks promised by the Vietnamese government could be prime factors for choosing Vietnam, Compal chairman Rock Hsu said earlier this year.

[…]See "Compal Eyes Vietnam for Factory, *Taipei Times*, April 9, 2007, accessed June 28, 2011, www.taipeitimes.com/News/biz/print/2007/04/09/2003355949. We have corrected a minor grammatical error in the article.

Niger

In Niger, West Africa, the World Bank is funding a $300 million project to improve education: "The Basic Education Project for Niger's objectives are: (i) to increase enrollment and completion in basic education programs and (ii) to improve management at all levels by improving the use of existing resources, focusing on rural areas to achieve greater equity and poverty reduction in the medium to long term."For more details, see World Bank, "Basic Education Project" World Bank, accessed June 28, 2011, web.worldbank.org/external/projects/main?page PK=64283627&piPK=73230&theSitePK=40941&menuPK=228424&Projectid=P061209. The World Bank website explains that the goals of the project are to improve access to primary education (including adult literacy), improve the quality of primary and secondary education, and improve the management capability of the Ministry of Education.

United States

President Obama recently established the President's Council on Jobs and Competitiveness, which is charged, among other things, with reporting "directly to the President on the design, implementation, and evaluation of policies to promote the growth of the American economy, enhance the skills and education of Americans, maintain a stable and sound financial and banking system, create stable jobs for American workers, and improve the long term prosperity and competitiveness of the American people."See "President's Council on Jobs and Competitiveness: About the Council," accessed July 27, 2011, www.whitehouse.gov/administration/advisory-boards/jobs-council/about. In his concern with competitiveness, President Obama follows directly in the footsteps of President George W. Bush, who, in 2006, established the American Competitiveness Initiative to Encourage American Innovation and Strengthen Our Nation's Ability to Compete in the Global Economy.George W. Bush, "State of the Union: American Competitiveness Initiative," White House Office of Communications, January 31, 2006, accessed June 28, 2011, www2.ed.gov/about/inits/ed/competitiveness/sou-competitiveness.pdf.

At first reading, these five stories seem to have little to do with each other. There is no obvious connection between the actions of the World Bank in Niger and Taiwanese computer manufacturers in Vietnam or between the marketing of Dubai and the arrival of Polish migrants in the United Kingdom. Yet they are indeed all connected. Think for a moment about the *consequences* of the following:

- An influx of workers to the United Kingdom
- A superior business environment in Dubai
- Improved education in Niger
- A new factory opening in Vietnam
- An improved banking system in the United States

Of course, each story has many different implications. But they have something fundamental in common: every single one of them will increase the real gross domestic product (real GDP) of the country in question. They all therefore shed light on one of the most fundamental questions in macroeconomics:

What determines a country's real GDP?

As we tackle this question, we will see that it is indeed connected to our stories of Dubai, the United Kingdom, Niger, Vietnam, and the United States.

Our stories have something else in common as well. In each case, they concern not only the country in isolation but also how it interacts with the rest of the world. The funds for Niger's education program are coming from other countries (via the World Bank). The US policy is designed to ensure that America is "leading the global competition that will determine our success in the 21st century.""Obama Presses for an Economy in 'Overdrive': Will Jobs Soon Follow?," PBS NewsHour, January 21, 2011, accessed August 22, 2011, www.pbs.org/newshour/bb/business/jan-june11/obamabusiness_01-21.html. Dubai is trying to attract investment from other countries. The workers in the United Kingdom are coming from Poland. The factory in Vietnam is being built so that a Taiwanese company can supply other manufacturers throughout the world.

Road Map

Real GDP is the broadest measure that we have of the amount of economic activity in an economy. In this chapter, we investigate the *supply* of real GDP in an economy. Firms in an economy create goods and services by transforming inputs into outputs. For example, think about the manufacture of a pizza. It begins with a recipe—a set of instructions. A chef following this recipe might require 30 minutes of labor time to make the dough and assemble the toppings and then might need 15 minutes use of a pizza oven to cook the pizza. The inputs here are as follows: the pizza oven, the labor time, the skills of the chef, and the recipe. Given 15 minutes of capital time, 30 minutes of labor time, a skilled chef, and the instructions, we can make one pizza.

In macroeconomics, we work with the analogous idea that explains how the total production in an economy depends on the available inputs. We first explain the relationship between the available inputs in the economy and the amount of real GDP that the economy can produce. Then we look at all the individual inputs in turn. If we can explain what determines the amount of each input in an economy and if we know the link from inputs to real GDP, then we can determine the level of real GDP. Finally, we look at a technique that allows us to quantify the relationship between inputs and output. Specifically, we look at how increases in different inputs translate into increases in overall GDP. Using this technique, we can see which inputs are particularly important.

This page titled 20.1: Five Stories is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

- **Current page** by Anonymous is licensed CC BY-NC-SA 3.0. Original source: https://2012books.lardbucket.org/books/theory-and-applications-of-economics.
- **23.3: Job and Worker Flows** by Anonymous is licensed CC BY-NC-SA 3.0. Original source: https://2012books.lardbucket.org/books/theory-and-applications-of-economics.

20.2: The Production of Real GDP 20.2: The Production of Real GDPEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What determines the production capabilities of an economy?
2. What is the marginal product of an input?
3. How is competitiveness related to the aggregate production function?

Economists analyze production in an economy by analogy to the production of output by a firm. Just as a firm takes inputs and transforms them into output, so also does the economy as a whole. We summarize the production capabilities of an economy with an **aggregate production function**.

The Aggregate Production Function

Physical capital refers to goods—such as factory buildings, machinery, and 18-wheel trucks—that have two essential features. First, *capital goods are used in the production of other goods*. The production of physical capital does not increase our well-being in and of itself. It allows us to produce more goods in the future, which permits us to enjoy more consumption at some future date. Second, *capital goods are long lasting*, which means we accumulate a capital stock over time. Capital goods are thus distinct from intermediate goods, which are fully used up in the production process.

The **capital stock** of an economy is the total amount of physical capital in the economy. As well as factories and machines, the capital stock includes physical infrastructure—road networks, airports, telecommunications networks, and the like. These are capital goods that are available for multiple firms to use. Sometimes these goods are supplied by governments (roads, for example); sometimes they are provided by private firms (cellular telephone networks are an example). For brevity, we often simply refer to "capital" rather than "physical capital." When you see the word *capital* appearing on its own in this book you should always understand it to mean physical capital.

Labor hours are the total number of hours worked in an economy. This depends on the size of the workforce and on how many hours are worked by each individual.We use the term *workforce* rather than labor force deliberately because the term *labor force* has a precise definition—those who are unemployed as well as those who are working. We want to include only those who are working because they are the ones supplying the labor hours that go into the production function. Chapter 23 discusses this distinction in more detail.

Human capital is the term that economists use for the skills and training of an economy's workforce. It includes both formal education and on-the-job training. It likewise includes technical skills, such as those of a plumber, an electrician, or a software designer, and managerial skills, such as leadership and people management.

Knowledge is the information that is contained in books, software, or blueprints. It encompasses basic mathematics, such as calculus and the Pythagorean theorem, as well as more specific pieces of knowledge, such as the map of the human genome, the formula for Coca-Cola, or the instructions for building a space shuttle.

Natural resources include land; oil and coal reserves; and other valuable resources, such as precious metals.

Toolkit: Section 31.26

The aggregate production function describes how aggregate output (**real gross domestic product [real GDP]**) in an economy depends on available inputs. The most important inputs are as follows:

- Physical capital: machines, production facilities, and so forth used in production
- Labor: the number of hours that are worked in the entire economy
- Human capital: the skills and education embodied in the work force of the economy
- Knowledge: the blueprints that describe the production process
- Natural resources: oil, coal, and other mineral deposits; agricultural and forest lands; and other resources
- Social infrastructure: the general business climate, the legal environment, and any relevant features of the culture

> Output increases whenever there is an increase in one of these inputs, all else being the same.

Social infrastructure refers to the legal, political, social, and cultural frameworks that exist in an economy. An economy with good social infrastructure is relatively free of corruption, has a functional and reliable legal system, and so on. Also included in social infrastructure are any relevant cultural variables. For example, it is sometimes argued that some societies are—for whatever reason—more entrepreneurial than others. As another example, the number of different languages that are spoken in a country influences GDP.

We show the production function schematically in Figure 20.2.1.

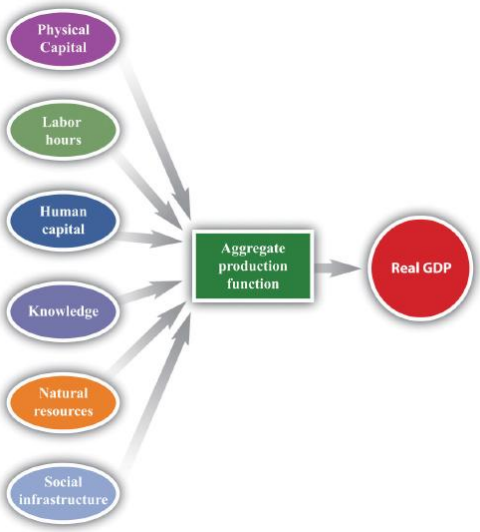


Figure 20.2.1: The Aggregate Production Function. The aggregate production function combines an economy's physical capital stock, labor hours, human capital, knowledge, natural resources, and social infrastructure to produce output (real GDP).

The idea of the production function is simple: if we put more in, we get more out.

- With more physical capital, we can produce more output. If you want to dig a foundation for a house, you will be more productive with a backhoe than a shovel; if you want to deliver documents from Chicago to St. Louis, you will be more productive using a truck than a bicycle.
- With more labor hours, we can produce more output. If there are more workers in an economy, or if they work longer hours, the economy will produce more real GDP.
- With more education and skills, we can produce more output. Skilled workers can produce more from an hour's work than unskilled workers can produce.
- With more knowledge, we can produce more output. Inventions and innovations make an economy more productive.
- With more natural resources, we can produce more output. For example, if an economy discovers additional oil reserves, it can produce more with given labor and capital than can economies without such resources. Of course, this input more often decreases rather than increases over time, as economies use up their existing stocks of natural resources.
- With better institutions, we can produce more output. Economies in which it is easy to establish businesses, where corruption is limited, and where the laws are reliable get more out of their workers and capital.

We call the extra output that we get from one more unit of an input, holding all other inputs fixed, the **marginal product** of that input. For example, the extra output we obtain from one more unit of capital is the **marginal product of capital**, the extra output we get from one more unit of labor is the **marginal product of labor**, and so on.

Physical capital and labor hours are relatively straightforward to understand and measure. To measure labor hours, we simply count the number of workers and the number of hours worked by an average worker. Output increases if we have more workers or if they work longer hours. For simplicity, we imagine that all workers are identical. Aggregate differences in the type and the quality of labor are captured in our human capital variable. For physical capital, we similarly imagine that there are a number of identical machines (pizza ovens). Then, just as we measure labor as the number of worker hours, so also we could measure capital by the total number of machine hours.If this were literally true, we could measure capital stock by simply counting the number of machines in an economy. In reality, however, the measurement of capital stock is trickier. Researchers must add together the *value*

of all the different pieces of capital in an economy. In practice, capital stock is usually measured indirectly by looking at the flow of additions to capital stock. We can produce more output by having more machines or by using each machine more intensively.

The other inputs that we listed—human capital, knowledge, social infrastructure, and natural resources—are trickier to define and much harder to quantify. Economists have used measures of educational attainment (e.g., the fraction of the population that completes high school) to compare human capital across countries.We use an index of human capital in Chapter 21. There are likewise some data that provide some indication of knowledge and social infrastructure—such as spending on research and development (R&D) and survey measures of perceived corruption.

The measurement of natural resources is problematic for different reasons. Land is evidently an input to production: factories must be put somewhere, and agriculture requires fields and orchards, so the value of land can be measured in principle. But what about reserves of oil or underground stocks of coal, uranium, or gold? First, such reserves or stocks contribute to real GDP only if they are extracted from the earth. An untapped oil field is part of a nation's wealth but makes no contribution to current production. Second, it is very hard to measure such stocks, even in principle. For example, the amount of available oil reserves in an economy depends on mining and drilling technologies. Oil that could not have been extracted two decades ago is now available; it is likely that future advances in drilling techniques will further increase available reserves in the economy.

We simply accept that, as a practical matter, we cannot directly measure an economy's knowledge, social infrastructure, and natural resources. As we see later in this chapter, however, there is a technique for indirectly measuring the combined influence of these inputs.

One thing might strike you as odd. Our description of production does *not* include as inputs the raw materials that go into production. The production process for a typical firm takes raw materials and transforms them into something more valuable. For example, a pizza restaurant buys flour, tomatoes, pepperoni, electricity, and so on, and transforms them into pizzas. The aggregate production function measures not the total value of these pizzas but the *extra value* that is added through the process of production. This equals the value of the pizzas *minus* the value of the raw materials. We take this approach to avoid double counting and be consistent with the way real GDP is actually measured.Reserves of natural resources are not counted as raw materials. The output of the mining sector is the value of the resources that have been extracted from the earth.

A Numerical Example of a Production Function

Table 20.2.1 "A Numerical Example of a Production Function" gives a numerical example of a production function. The first column lists the amount of output that can be produced from the inputs listed in the following columns.

Table 20.2.1: A Numerical Example of a Production Function

Row	Output	Capital	Labor	Other Inputs
Increasing Capital				
A	100	1	1	100
B	126	2	1	100
C	144	3	1	100
D	159	4	1	100
Increasing Labor				
E	100	1	1	100
F	159	1	2	100
G	208	1	3	100
H	252	1	4	100

Row	Output	Capital	Labor	Other Inputs
Increasing Other Inputs				
I	100	1	1	100
J	110	1	1	110
K	120	1	1	120
L	130	1	1	130

If you compare row A and row B of Table 20.2.1 "A Numerical Example of a Production Function", you can see that an increase in capital (from 1 unit to 2 units) leads to an increase in output (from 100 units to 126 units). Notice that, in these two rows, all other inputs are unchanged. Going from row B to row C, capital increases by another unit, and output increases from 126 to 144. And going from row C to row D, capital increases from 3 to 4 and output increases from 144 to 159. We see that increases in the amount of capital lead to increases in output. In other words, the marginal product of capital is positive.

Similarly, if you compare rows E–H of Table 20.2.1 "A Numerical Example of a Production Function", you can see that the marginal product of labor is positive. As labor increases from 1 to 4 units, and we hold all other inputs fixed, output increases from 100 to 252 units. Finally, rows I to L show that increases in other inputs, holding fixed the amount of capital and labor, likewise leads to an increase in output.

Figure 20.2.2 illustrates the production function from Table 20.2.1 "A Numerical Example of a Production Function". Part (a) shows what happens when we increase capital, holding all other inputs fixed. That is, it illustrates rows A–D of Table 20.2.1 "A Numerical Example of a Production Function". Part (b) shows what happens when we increase labor, holding all other inputs fixed. That is, it illustrates rows E–H of Table 20.2.1 "A Numerical Example of a Production Function".

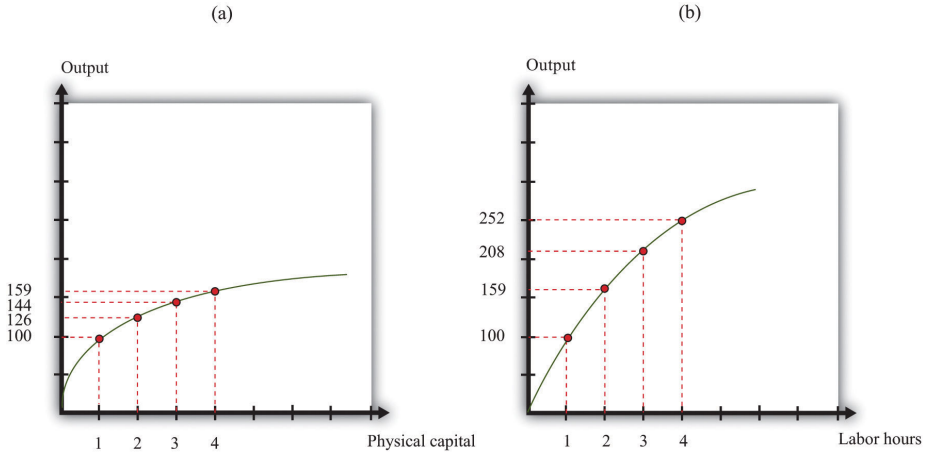


Figure 20.2.2: A Graphical Illustration of the Aggregate Production Function. The aggregate production function shows how the amount of output depends on different inputs. Increases in the amount of physical capital (a) or the number of labor hours (b)—all else being the same—lead to increases in output.

Diminishing Marginal Product

You may have noticed another feature of the production function from Figure 20.2.2 and Table 20.2.1 "A Numerical Example of a Production Function". Look at what happens as the amount of capital increases. Output increases, as we already noted—but by smaller and smaller amounts. Going from 1 unit of capital to 2 yields 26 extra units of output (= 126 − 100). Going from 2 to 3 units of capital yields 18 extra units of output (= 144 − 126). And going from 3 to 4 yields 15 extra units of output (= 159 − 144). The same is true of labor: each *additional* unit of labor yields less and less additional output. Graphically, we can see that the production function becomes more and more flat as we increase either capital or labor. Economists say that the production function we have drawn exhibits **diminishing marginal product**.

The more physical capital we have, the less additional output we obtain from additional physical capital. As we have more and more capital, other things being equal, additions to our capital stock contribute less and less to output. Economists call this idea **diminishing marginal product of capital**.

The more labor we have, the less additional output we obtain from additional labor. Analogously, this is called **diminishing marginal product of labor**. As we have more and more labor, we find that additions to our workforce contribute less and less to output.

Diminishing marginal products are a plausible feature for our production function. They are easiest to understand at the level of an individual firm. Suppose you are gradually introducing new state-of-the-art computers into a business. To start, you would want to give these new machines to the people who could get the most benefit from them—perhaps the scientists and engineers who are working in R&D. Then you might want to give computers to those working on production and logistics. These people would see a smaller increase in productivity. After that, you might give them to those working in the accounting department, who would see a still smaller increase in productivity. Only after those people have been equipped with new computers would you want to start supplying secretarial and administrative staff. And you might save the chief executive officer (CEO) until last.

The best order in which to supply people would, of course, depend on the business. The important point is that you should at all times give computers to those who would benefit from them the most in terms of increased productivity. As the technology penetrates the business, there is less and less additional gain from each new computer.

Diminishing marginal product of labor is also plausible. As firms hire more and more labor—holding fixed the amount of capital and other inputs—we expect that each hour of work will yield less in terms of output. Think of a production process—say, the manufacture of pizzas. Imagine that we have a fixed capital stock (a restaurant with a fixed number of pizza ovens). If we have only a few workers, then we get a lot of extra pizza from a little bit of extra work. As we increase the number of workers, however, we start to find that they begin to get in each others' way. Moreover, we realize that the amount of pizza we can produce is also limited by the number of pizza ovens we have. Both of these mean that as we increase the hours worked, we should expect to see each *additional* hour contributing less and less in terms of additional output.

In contrast to capital and labor, we do not necessarily assume that there are diminishing returns to human capital, knowledge, natural resources, or social infrastructure. One reason is that we do not have a natural or obvious measure for human capital or technology, whereas we do for labor and capital (hours of work and capital usage).

Globalization and Competitiveness: A First Look

Over the last several decades, a host of technological developments has reduced the cost of moving both physical things and intangible information around the world. The lettuce on a sandwich sold in London may well have been flown in from Kenya. A banker in Zurich can transfer funds to a bank in Pretoria with a click of a mouse. People routinely travel to foreign countries for vacation or work. A lawyer in New York can provide advice to a client in Beijing without leaving her office. These are examples of **globalization**—the increasing ability of goods, capital, labor, and information to flow among countries.

One consequence of globalization is that firms in different countries compete with each other to a much greater degree than in the past. In the 1920s and 1930s, the automobiles produced by Ford Motor Company were almost exclusively sold in the United States, while those produced by Daimler-Benz were sold in Europe. Today, Ford and DaimlerChrysler (formed after the merger of Chrysler and Daimler-Benz in 1998) compete directly for customers in both Europe and the United States—and, of course, they also compete with Japanese manufacturers, Korean manufacturers, and others.

Competition between firms is a familiar idea. Key to this idea of competition is that one firm typically gains at the expense of another. If you buy a hamburger from Burger King instead of McDonald's, then Burger King is gaining at McDonald's expense: it is getting the dollars that would instead have gone to McDonald's. The more successful firm will typically see its production, revenues, and profits all growing.

It is tempting to think that, in a globalized world, nations compete in much the same way that firms compete—to think that one nation's success must come at another's expense. Such a view is superficially appealing but incorrect. Suppose, for example, that South Korea becomes better at producing computers. What does this imply for the United States? It does make life harder for US computer manufacturers like Dell Inc. But, at the same time, it means that there is more real income being generated in South Korea, some of which will be spent on US goods. It also means that the cheaper and/or better computers produced in South Korea will be available for US consumers and producers. In fact, we expect growth in South Korea to be beneficial for the United States. We should welcome the success of other countries, not worry about it.

What is the difference between our McDonald's-Burger King example and our computer example? If lots of people switched to McDonald's from Burger King, then McDonald's would become less profitable. It would, in the end, become a smaller company: it would lay off workers, close restaurants, and so on. A company that is unable to compete at all will eventually go bankrupt. But if South Korea becomes better at making computers, the United States doesn't go bankrupt or even become a significantly smaller economy. It has the same resources (labor, capital, human capital, and technology) as before. Even if Dell closes factories and lays off workers, those workers will then be available for other firms in the economy to hire instead. Other areas of the economy will expand even as Dell contracts.

In that case, do countries compete at all? And if so, then how? The Dubai government's website that we showed at the beginning of this chapter provides a clue. The website sings the praises of the Emirate as a place for international firms to establish businesses. Dubai is trying to entice firms to set up operations there: in economic language, it wants to attract capital and skilled labor. Dubai is not alone. Many countries engage in similar advertising to attract business. And it is not only countries: regions, such as US states or even cities, deliberately enact policies to influence business location.

Dubai is trying to gain more resources to put into its aggregate production function. If Dubai can attract more capital and skilled labor, then it can produce more output. If it is successful, the extra physical and human capital will lead to Dubai becoming a more prosperous economy.

In the era of globalization, inputs can move from country to country. Labor can move from Poland to the United Kingdom or from Mexico to the United States, for example. Capital can also move. At the beginning of this chapter, we quoted from an article explaining that a Taiwanese manufacturer was planning to open a factory in Vietnam, drawn by low wages and preferential tax treatment. This, then, is the sense in which countries compete with each other—they compete to attract inputs, particularly capital. **Competitiveness** refers to the ability of an economy to attract physical capital.

We have more to say about this later. But we should clear up one common misconception from the beginning. Competing for capital does *not* mean "competing for jobs." People worried about globalization often think that if a Taiwanese factory opens a factory in Vietnam instead of at home, there will be higher unemployment in Taiwan. But the number of jobs—and, more generally, the level of employment and unemployment—in an economy does not depend on the amount of available capital. Chapter 23 and Chapter 25 explain what determines these variables. This does not mean that factory closures are benign. They can be very bad news for the individual workers who are laid off and must seek other jobs. And movements of capital across borders can—as we explain later—have implications for the *quality* of available jobs and the wages that they pay. But they do not determine the number of jobs available.

Key Takeaways

- The production capabilities of an economy are described by the aggregate production function, characterizing how the factors of production, such as capital, labor, and technology, are combined to produce real GDP.
- In the aggregate production function, the marginal product is the extra amount of real GDP obtained by adding an extra unit of an input.
- One measure of competitiveness is the ability of an economy to attract inputs for the production function, particularly capital.

Exercises

1. Earlier, we observed that our news stories were about the following:
   - Improved education in Niger
   - A new factory opening in Vietnam
   - A superior business environment in Dubai
   - An influx of workers to the United Kingdom
   - A better banking system in the United States

   Which input to the production function is being increased in each case?
2. Building on part (b) of Figure 20.2.2, draw an aggregate production function that does not exhibit diminishing marginal product of labor.

This page titled 20.2: The Production of Real GDP is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

20.3: Labor in the Aggregate Production Function 20.3: Labor in the Aggregate Production FunctionEdit Update titlecancel

learing objectives

After you have read this section, you should be able to answer the following questions:

1. What determines the amount of labor in the aggregate production function?
2. What determines the patterns of labor migration?
3. Why do real wages differ across countries?

The aggregate production function tells us how much output we get from the inputs that we have available. Our next task is to *explain* how much of each input goes into this production function. When we have done this, we will have explained the level of real gross domestic product (real GDP). We begin with labor because it is the most familiar—almost everyone has had the experience of selling labor services.

The Labor Market

Figure 20.3.1 shows a diagram for the **labor market**. In this picture, we draw the supply of labor by households and the demand for labor by firms. The price on the vertical axis is the **real wage**. The real wage is just the nominal wage (the wage in dollars) divided by the price level. It tells us the amount that you can consume (measured as the number of units of real GDP that you get) if you sell one hour of your time.

Toolkit: Section 31.3 and Section 31.8

When we adjust the nominal wage in this way, we are "correcting for inflation." The toolkit gives more information. You can also review the labor market in the toolkit.

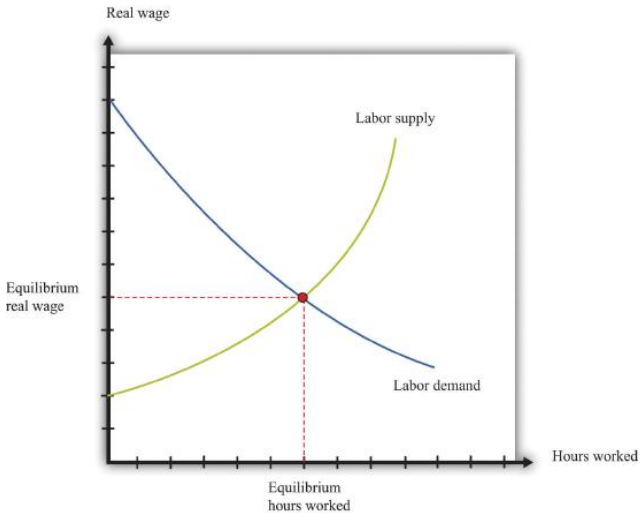


Figure 20.3.1: Equilibrium in the Labor Market. Equilibrium in the labor market occurs where the number of hours of labor supplied by households equals the number of hours of labor demanded by firms.

The upward-sloping **labor supply** curve comes from both an increase in hours worked by each employed worker and an increase in the number of employed workers.We discuss labor supply in more detail in Chapter 27. The downward-sloping **labor demand** curve comes from the decision rule of firms: each firm purchases additional hours of labor up to the point where *the extra output that it obtains from that labor equals the cost of that labor*. The extra output that can be produced from one more hour of work is—by definition—the marginal product of labor, and the cost of labor, measured in terms of output, is the real wage. Therefore firms hire labor up to the point where the marginal product of labor equals the real wage.

The marginal product of labor also depends on the other inputs available in an economy. An economy with more physical or human capital, for example, is one in which workers will be more productive. Increases in other inputs shift the labor demand curve

rightward.

The point where the labor supply and labor demand curves meet is the point of equilibrium in the labor market. At the equilibrium real wage, the number of hours that workers want to work exactly matches the number of hours that firms wish to use. Figure 20.3.1 shows that equilibrium in the labor market tells us two things: the real wage in the economy and how many hours of work go into the aggregate production function.

The Mobility of Labor

In November 2004, the median hourly wage in Florida was $12.50. In Washington State, it was $16.07. On average, in other words, wages were almost 30 percent higher in the Northwest compared to the Southeast. To take a more specific example, the median wage for health-care support occupations (dental assistants, pharmacy aides, hospital orderlies, etc.) was $8.14 in Mississippi and $12.81 in Massachusetts. Dental assistants who moved from Baton Rouge to Boston could expect to see about a 50 percent increase in their hourly wage."Occupational Employment Statistics," Bureau of Labor Statistics, accessed June 29, 2011, http://www.bls.gov/oes/current/oessrcst.htm.

People in the United States are free to move from state to state, and many people do indeed move from one state to another every year. People move for many reasons: to go to college, join a girlfriend or boyfriend, or move to the place where they have always dreamed of living (such as New York; Los Angeles; or Burr Ridge, Illinois). People also move to take up new jobs, and one of the things that induces them to take one job rather than another is the wage that it pays. Different wages in different places therefore affect the patterns of migration across the United States.

Figure 20.3.2 shows the labor markets in Florida and Washington State for November 2004. The cost of living was different in those two states but, to keep our story simple, we ignore these differences. That is, we assume that there is no difference in the price level in the two states. If we set 2004 as the base year, the **price level** is 1. This means that the real wage is the same as the nominal wage. A more careful analysis would correct for differences in state taxes and the cost of living.

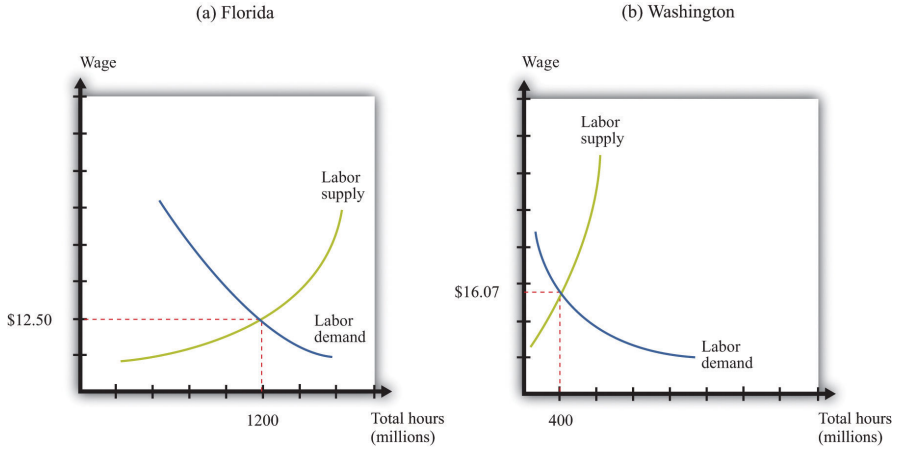


Figure 20.3.2: Labor Markets in Florida and Washington State. These diagrams show the labor markets in (a) Florida and (b) Washington State. Real wages are higher in Washington State than in Florida.

Part (a) of Figure 20.3.2 shows the labor market in Florida. The equilibrium wage is $12.50, and the equilibrium level of employment is 1,200 million hours. This is roughly calibrated to the actual experience in Florida, where total employment in 2004 was just under 7.5 million individuals. Part (b) of Figure 20.3.2 shows Washington State, where the equilibrium wage is $16.07, and employment is 400 million hours.

We expect that the higher wages in Washington State would attract people to move from Florida to Washington State. Workers would migrate from Florida to Washington State, causing the labor supply curve to shift leftward in Florida and rightward in Washington State. As a consequence, wages would increase in Florida and decrease in Washington State. Figure 20.3.3 shows what would happen if the only thing people cared about was wages: migration would stop only when wages were equal in both states. Employment would be lower in Florida and higher in Washington State. (The exact number of people who moved and the new equilibrium wage would depend on the slopes of the supply and demand curves in both labor markets.)

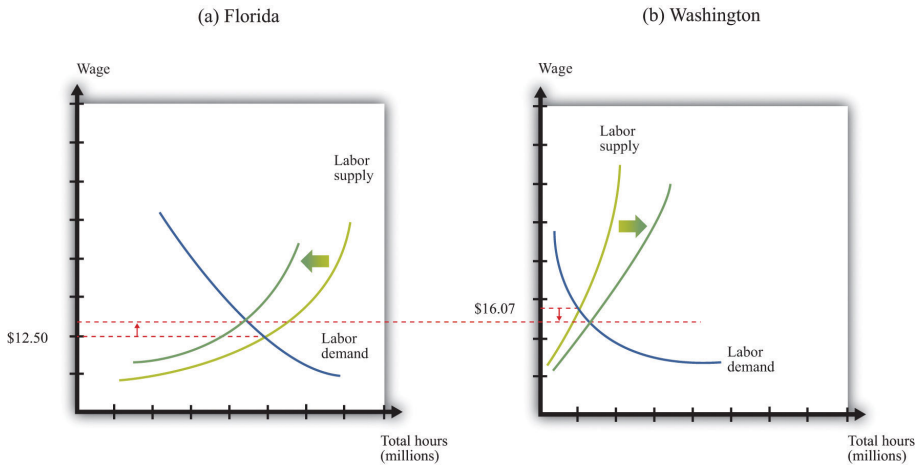


Figure 20.3.3: Migration from Florida to Washington State. Workers move from Florida to Washington State in search of higher wages. Labor supply decreases in Florida and increases in Washington State.

If wages were the only factor affecting people's decisions, migration would completely equalize real wages across the different state economies. In fact, we do not expect wages to become exactly equal in Florida, Washington State, and the other 48 states of the Union. Differences in both state taxes and the cost of living in different states and cities lead to persistent differences in wages. Some places are less attractive to live than others, so people will need to be paid more to induce them to live there. Our example nevertheless illustrates a key economic principle: people respond to incentives. Individual decisions about where to live respond to differences in real wages. *Labor tends to migrate to where it can earn the highest return.*

International Migration

People migrate between different US states because of wage differences. In China and other developing economies, many workers migrate from rural areas to urban areas, again in search of better wages. The same forces operate across international borders. Workers seek to emigrate from countries where their wages are low and move to countries that pay higher wages. Sometimes, this movement is actively encouraged. Some countries attract immigrant workers—particularly rich economies that want to attract relatively unskilled workers to perform low-paying and unattractive jobs.

However, there are many more barriers to movement among countries compared to movement within countries. Some are legal barriers. Most countries strictly limit the immigration that they permit. In the United States, a physical barrier has been constructed along some of the US-Mexican border to prevent illegal immigration from Mexico to the United States. Some countries also make emigration very difficult.

Even when legal impediments to migration are absent, there are cultural and language barriers. European Union citizens are legally free to live and work anywhere in the countries of the Union, and we saw at the beginning of this chapter that many young Polish workers take advantage of this by moving to the United Kingdom in search of work. But such examples notwithstanding, most European workers remain in the country of their birth. Migration from Portugal to Finland is very limited, for example, despite the higher wages paid in Finland. A Portuguese worker who wants to move to Finland must learn to cope with a completely different language and culture, not to mention a much colder climate.

To summarize, while we do see some movement of labor across national borders, people remain, for the most part, in the country in which they were born. When we are analyzing national economies, the main determinant of labor hours is, in the end, the number of people in the economy and the number of hours that they choose to work. International migration plays a limited role.

We can also turn this argument on its head. We observe huge differences in real wages in different economies. If people were truly able and willing to migrate across economies, we would expect most of those differences to disappear. So we can conclude that there must be substantial barriers to migration.

Population Growth and Other Demographic Changes

Over long periods of time, the amount of labor in the production function is affected by changes in population and other demographic changes. As a country's population increases, it has more workers to "plug in" to the aggregate production function. Changes in the age structure of the population also have an effect. Much of the developed world has an aging population, meaning that the fraction of the population that is working is decreasing.We discuss some implications of this in Chapter 28.

Changes in social norms can also affect the amount of labor that goes into the production function. For example, child labor is now uncommon, whereas a century ago it was much more usual. Another example is the increase in women's participation in the labor force over the last half century, both in the United States and other countries. Public health matters as well. In some countries of the world, particularly in Africa, the HIV/AIDS crisis is having devastating effects. Quite apart from the human misery that the disease causes, the epidemic means that there is less labor available. The problem is particularly acute because working-age individuals are disproportionately affected.

In an introductory economics textbook such as this one, we do not seek to explain such social changes. To be sure, these changes are studied by economists, as well as by sociologists and other researchers. But here we investigate the effects rather than the causes of such social changes.

Explaining International Differences in the Real Wage

Real wages differ markedly across countries: the typical worker in Australia is paid much more than the typical worker in Bolivia, for example. Suppose that we compare two countries, and we find that real wages are higher in one country (country A) than in the other (country B). This tells us that the marginal product of labor is higher in country A than in country B. There are two basic reasons why this might be true:

1. Hours worked are fewer in country A than in country B.
2. Other inputs are larger in country A than in country B.

Figure 20.3.4 illustrates these possibilities. Part (a) compares two countries that are identical except that less labor is supplied to the market in country A. In country A, the real wage is higher, and the equilibrium number of hours is lower. In part (b), the two countries have identical labor supplies, but one or more of the other inputs (physical capital, human capital, knowledge, social infrastructure, or natural resources) is higher in country A. This means that the labor demand curve in country A is further to the right, so the real wage is higher, and the equilibrium number of hours is also higher.

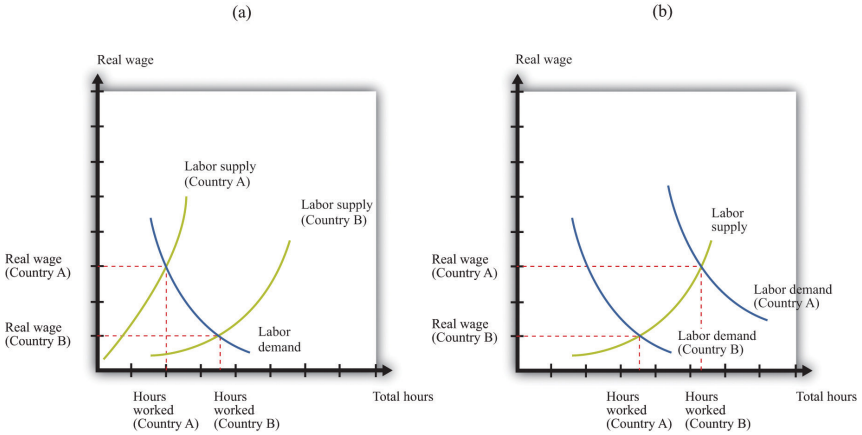


Figure 20.3.4: Why Real Wages May Be Different in Different Countries. Real wages are higher in country A than in country B either because of lower labor supply in country A (a) or greater labor demand in country A (b).

The real wage is an indicator of societal welfare because it tells us about the living standards of the typical worker. From the perspective of workers, increases in other inputs—such as capital stock or an economy's human capital—are desirable because they increase the marginal product of labor and hence the real wage.

Thus, when the World Bank helps to fund education in Niger, it is helping to increase GDP by increasing the amount of human capital in the production function. Furthermore, this increased GDP will appear in the form of higher wages and living standards in the economy. Conversely, if a food processing company decides to close a factory in England, capital stock in England decreases, and output and real wages decrease.

Key Takeaways

- The quantity of labor in the aggregate production function is determined in the labor market.
- All else being the same, labor will migrate to the place with the highest real wage.
- Differences in real wages across economies reflect differences in the marginal product of labor due to differences in the number of hours worked, technology, and capital stocks.

Exercises

1. To determine the patterns of labor migration, should we look at nominal or real wages? Should we look at wages before or after taxes?
2. Building on Figure 20.3.4, suppose that country A had fewer workers than country B but more capital. Would the real wage be higher or lower in country A than country B?

This page titled 20.3: Labor in the Aggregate Production Function is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

20.4: Physical Capital in the Aggregate Production Function 20.4: Physical Capital in the Aggregate Production FunctionEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What determines the movement of investment in a country?
2. How does the capital stock of a country change?
3. What determines the movement of capital across countries?

Many of the arguments that we have just made about labor have analogies when we think about capital. Just as the amount of labor in an economy depends on the size of the workforce, so the amount of capital depends on the capital stock. Just as the amount of labor depends on how many hours each individual works, so the amount of capital depends on the utilization rate of capital.

Capital utilization is the rate at which the existing capital stock is used. For example, if a manufacturing firm runs its production lines 24 hours per day, 7 days per week, then its capital utilization rate is very high.

Just as labor can migrate from country to country, so also capital may cross national borders. In the short run, the total amount of capital in an economy is more or less fixed. We cannot make a significant change to the capital stock in short periods of time. In the longer run, however, the capital stock changes because some of the real gross domestic product (real GDP) produced each year takes the form of new capital goods—new factories, machines, computers, and so on. Economists call these new capital goods **investment**.

Toolkit: Section 31.27

Investment is one of the components of overall GDP.

The Circular Flow: The Financial Sector

We can use the circular flow to help us understand how much investment there is in an economy. Figure 20.4.1 reviews the four flows of dollars in and out of the financial sector. The circular flow is introduced in Chapter 18. We elaborate on it in Chapter 19, Chapter 22, Chapter 27, and Chapter 29.

1. Households put their savings into the financial sector. Any income that households receive today but wish to put aside for the future is sent to the financial markets. Although individual households both save and borrow, there is almost always more saving than borrowing, so, on net, there is a flow of dollars from the household sector into the financial markets (**private savings**).
2. There is a flow of dollars between the financial sector and the government sector. This flow can go in either direction. Figure 20.4.1 is drawn for the case where the government is borrowing (there is a **government deficit**), so the financial markets send money to the government sector. In the case of a government surplus, the flow goes in the other direction. The **national savings** of an economy are the savings carried out by the private and government sectors taken together:

$$national\ savings\ =\ private\ savings\ +\ government\ surplus$$

or

$$national\ savings\ =\ private\ savings\ -\ government\ deficit.$$

1. There is a flow of dollars between the financial sector and the foreign sector. This flow can also go in either direction. When our economy exports more than it imports, we are sending more goods and services to other countries than they are sending to us. This means that there is a flow of dollars from the economy as foreigners buy dollars so that they can make these purchases. It also means that we are lending to other countries: we are sending more goods and services to other countries now in the understanding that we will receive goods and services from them at some point in the future. By contrast, when our economy imports more than it exports, we are receiving more goods and services from other countries than we are sending to them. We are then borrowing from other countries, and there is a flow of dollars into the economy. Figure 20.4.1 illustrates the case of borrowing from other countries.

2. There is a flow of dollars from the financial sector into the firm sector. These are the funds that are available to firms for investment purposes.

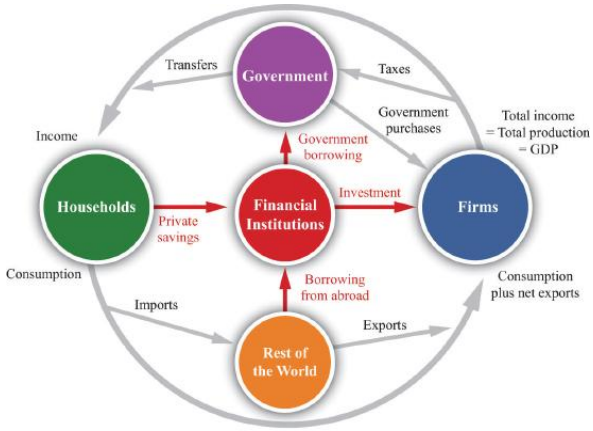


Figure 20.4.1: The Flows In and Out of the Financial Sector. The flows in and out of the financial sector must balance, which tells us that investment is financed by national savings plus borrowing from abroad.

The total flows in and out of the financial sector must balance. Because of this, as we see from Figure 20.4.1, there are two sources of funding for new physical capital: savings generated in the domestic economy and borrowing from abroad.

$$investment \ = \ national \ savings \ + \ borrowing \ from \ other \ countries.$$

Or, in the case where we are lending to other countries,

$$investment \ = \ national \ savings \ - \ lending \ to \ other \ countries.$$

Changes in the Capital Stock

Capital goods don't last forever. Machines break down and wear out. Technologies become obsolete: a personal computer (PC) built in 1988 might still work today, but it won't be much use to you unless you are willing to use badly outdated software and have access to old-fashioned 5.25-inch floppy disks. Buildings fall down—or at least require maintenance and repair.

Depreciation is the term economists give to the amount of the capital stock that an economy loses each year due to wear and tear. Different types of capital goods depreciate at different rates. Buildings might stay standing for 50 or 100 years; machine tools on a production line might last for 20 years; an 18-wheel truck might last for 10 years; a PC might be usable for 5 years. In macroeconomics, we do not worry too much about these differences and often just suppose that all capital goods are the same.

The overall capital stock increases if there is enough investment to replace the worn out capital and still contribute some extra. The overall change in the capital stock is equal to new investment minus depreciation:

change in capital stock = investment − depreciation of existing capital stock.

Investment and depreciation are the *flows* that lead to changes in the *stock* of physical capital over time. We show this schematically in Figure 20.4.2. Notice that capital stock could actually become smaller from one year to the next, if investment were insufficient to cover the depreciation of existing capital.

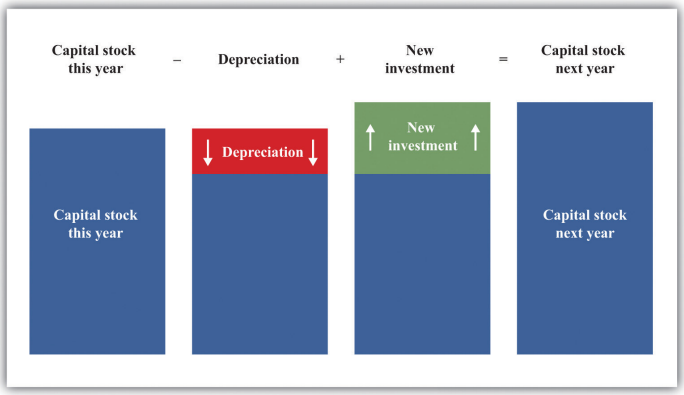


Figure 20.4.2: The Accumulation of Capital. Every year, some capital stock is lost to depreciation, as buildings fall down and machines break down. Each year there is also investment in new capital goods.

The Mobility of Capital

Can physical capital move from place to place? A first guess might be no. Although some capital goods, such as computers, can be transported, most capital goods are fixed in place. Factories are not easily moved from one place to another.

New capital, however, can be located anywhere. When Toyota decides to build a new factory, it could put it in Japan, the United States, Italy, Vietnam, or Brazil. Even if existing capital stocks are not very mobile, investment is. In the long run, firms can decide to close operations in one country and open in another. To understand how much capital a country has, therefore, we must recognize that investment in one country may come from elsewhere in the world.

Just as workers go in search of high wages, so the owners of capital seek to find the places where capital will have the highest return. We already know that the real wage is a measure of the marginal product of labor. Similarly, the real return on investment is the marginal product of capital (more precisely, the marginal product of capital adjusted for depreciation). Remember that the marginal product of capital is defined as the amount of *extra* output generated by an extra unit of capital. The owners of capital look to put their capital in countries where its marginal product is high.

Earlier, we saw two reasons why the marginal product of labor (and thus the real wage) might be higher in one country rather than another. There are likewise two reasons why the marginal product of capital might be higher in one country (A) rather than in another country (B). Holding all else the same, the marginal product of capital will be higher in country A if

- The capital stock is smaller in country A than in country B.
- The stock of other inputs is larger in country A than in country B.

These two factors determine the return on investment in a country. The benefits of acquiring more capital are higher in a country that has relatively little capital than in a country that has a lot of capital. This is because new capital can be allocated to projects that yield a lot of extra output, but as the country acquires more and more capital, such projects become harder and harder to find. Conversely, a country that has more of the other inputs in the production function will have a higher marginal product of capital.

Countries with a lot of labor, other things being equal, will be able to get more out of a given piece of machinery—because each piece of machinery can be combined with more labor time. As a simple example, think about taxis. In a capital-rich country, there may be only one driver for every taxi. In a poorer country, two or three drivers often share a single vehicle, so that vehicle spends much more time on the road. The return on capital—other things being equal—is higher in countries with a lot of labor and not very much capital to share around. Such countries are typically relatively poor, suggesting that poor countries should attract investment funds from elsewhere. In other words, basic economics suggests that if the return on investment is indeed higher in poor countries, investment funds should flow to those countries.

We certainly do see individual examples of such flows. The story at the beginning of this chapter about a Taiwanese company establishing a factory in Vietnam is one example. The following quotation from a British trade publication describes another.

Less than two months into 2006 and the UK's grocery manufacturing industry is already notching up a growing list of casualties: Leaf UK is considering whether to close its factory in Stockport; Elizabeth Shaw is shutting a plant in Bristol; Arla Foods UK is pulling out of a site at Uckfield; Richmond Foods is ending production in Bude; and Hill Station is shutting a site in Cheadle.

[…]

The stories behind these closures are all very different. But two common trends emerge. First, suppliers are being forced to step up the pace of consolidation as retailer power grows and that means more facilities are being rationalised. Second, production is being shifted offshore as grocery suppliers take advantage of lower-cost facilities.“Shutting Up Shop,” *The Grocer*, February 25, 2006, accessed June 28, 2011, www.coadc.com/grt_article_6.htm. *The Grocer* is a trade publication for the grocery industry in the United Kingdom.

This excerpt observes that food processing that used to be carried out in Britain is being shifted to poorer Eastern European countries, such as Poland. When factories close in Britain and open in Poland, it is as if physical capital—factories and machines—is moving from Britain to other countries.

If the amount of capital (relative to labor) were the only factor determining investment, we would expect to see massive amounts of lending going from rich countries to poor countries. Yet we do not see this. The rich United States, in fact, *borrows* substantially from other countries. The stock of other inputs—human capital, knowledge, social infrastructure, and natural resources—also matters. If workers are more skilled (possess more human capital) or if an economy has superior social infrastructure, it can obtain more output from a given amount of physical capital. The fact that the United States has more of these inputs helps to explain why investors perceive the marginal product of capital to be high in the United States.

Earlier we explained that even though migration could in principle even out wages in different economies, labor is, in fact, not very mobile across national boundaries. Capital *is* relatively mobile, however, and the *mobility of capital will also tend to equalize wages*. If young Polish workers move from Poland to England, real wages will tend to increase in Poland and decrease in England. If grocery manufacturers move production from England to Poland, then real wages will likewise tend to increase in Poland and decrease in England.

In fact, imagine that two countries have different amounts of physical capital and labor, but the same amount of all other inputs. If physical capital moves freely to where it earns the highest return, then both countries will end up with the same marginal product of capital *and the same marginal product of labor*. The movement of capital substitutes for labor migration and leads to the same result of equal real wages. This is a striking result.

The result is only this stark if the two countries have identical human capital, knowledge, social infrastructure, and natural resources.There are, not surprisingly, other, more technical, assumptions that matter as well. Perhaps the most important is that the production function should indeed display diminishing marginal product of capital, as we have assumed in this chapter. If other inputs differ, then the mobility of capital will still affect wages, but wages will remain higher in the economy with more of other inputs. If workers in one country have higher human capital, then they will earn higher wages even if capital can flow freely between countries. But the underlying message is the same: globalization, be it in the form of people migrating from one country to another or capital moving across national borders, should tend to make the world a more equal place.

Key Takeaways

- As an accounting identity, the amount of investment is equal to the national savings of a country plus the amount it borrows from abroad.
- The capital stock of a country changes over time due to investment and depreciation of the existing capital stock.
- Differences in the marginal product of capital lead to movements of capital across countries.

Exercises

1. Can investment ever be negative at a factory? In a country?
2. Explain why the movement of capital across two countries will have an effect on the real wages of workers in the two countries.

This page titled 20.4: Physical Capital in the Aggregate Production Function is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

20.5: Other Inputs in the Aggregate Production Function 20.5: Other Inputs in the Aggregate Production FunctionEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. How does the amount of human capital in a country change over time?
2. How is knowledge created?
3. How do property rights influence the aggregate production function?

We have less to say about the other inputs into the aggregate production function, so we group them together.

Human Capital

Education makes the most important contribution to human capital in an economy. Kindergarteners learning to count are acquiring human capital, as are high-school students learning algebra, undergraduate students learning calculus, and experienced workers studying for an MBA. People also acquire human capital on the job—either as a result of explicit company training programs or simply because of practice and experience (sometimes called “learning by doing”).

The education policy of national governments therefore plays a big part in determining how much human capital there is in a country. In the United States and Europe, education is typically compulsory up to age 15 or 16. In other countries, the school-leaving age is lower: 10 in Bangladesh, 11 in Iran, and 13 in Honduras, for example. In still other countries, education is not compulsory at all.“At What Age…? Comparative Table,” *Right to Education Project*, accessed June 28, 2011, http://www.right-to-education.org/node/279. One of the aims of the American Competitiveness Initiative, mentioned at the beginning of this chapter, was to “provide American children with a strong foundation in math and science.”

There are many similarities between human capital and physical capital. Human capital, like physical capital, is accumulated through a process of investment. Basic education is an investment made by parents and governments. University education is an investment made by individuals and households. When you go to college, you give up time that you could have spent working or having fun. This is one cost of education. The other cost is the expense of tuition. The gain from education—the return on your investment—is that sometime in the future you will be more productive and earn more income. An individual decision to go to college is based on an evaluation of the costs (such as tuition and foregone time) and the benefits (such as higher salary after graduation and the joy of studying fascinating subjects like economics).

Firms also invest in human capital. They seek to increase the productivity of their workers by in-house training or by sending workers to external training courses. Large firms typically devote substantial resources to the training and development of their employees. Some of the skills that workers acquire are transferable to other firms if the worker moves to another job. For example, workers who have attended a training course on accounting would be able to use the knowledge they acquired from that course at many different firms. Other skills are specific to a particular firm (such as knowing exactly where to hit a particular machine with a hammer when it jams).

Human capital, like physical capital, can depreciate. People forget things that they learned, or their knowledge becomes obsolete. VisiCalc was once a leading spreadsheet software, so people skilled in its use had valuable human capital; yet knowledge of this program is of little use today. Human capital that is specific to a particular firm is particularly prone to depreciation because it becomes worthless if the worker leaves or if the firm goes out of business. One reason why factory closures—such as those in the food retailing sector in the United Kingdom—arouse such concern is that laid-off workers may see their useful human capital decline and end up with lower paying jobs as a result.

While there are similarities between physical and human capital, there are also differences. Most importantly, human capital is trapped inside people. Economists say that such skills are “embodied” in the labor force. You cannot sell the human capital that you own without selling your own labor time as well. The implication for government policy is that importing human capital means importing people. Dubai is trying to attract human capital—so it advertises the things that make the country attractive to individuals who own that human capital. Thus their website speaks of the “cosmopolitan lifestyle” in Dubai, together with the quality of the hospitals, schools, shops, and so on.

Knowledge

Many large firms contain research and development (R&D) divisions. Employees in these divisions engage in product development and process development. **Product development** consists of developing new products and improving a firm's existing products. **Process development** consists of finding improvements in a firm's operations and methods of manufacture to reduce the costs of production.

An example of product development is the development and testing of a new pharmaceutical compound to treat cancer. An example of process development is the way in which transportation firms now use global positioning systems (GPSs) to better manage the movements of their trucks. In either case, firms invest today in the hope of gains in the future from lower production costs and better products.

Knowledge of this kind is also created by independent research laboratories, universities, think tanks, and other such institutions. In many cases, governments subsidize these institutions: policymakers actively intervene to encourage the production of new knowledge. Governments get involved because new knowledge can benefit lots of different firms in an economy. Think of how the invention of electric power, the internal combustion engine, the microchip, or the Internet benefits almost every firm in the economy today.

Economists say that basic knowledge is a **nonrival**. A good is nonrival if one person's consumption of that good does not prevent others from also consuming it. A good is rival if one person's consumption prevents others from also consuming it. The fact that one marketing manager is using economic theory to set a profit-maximizing price doesn't prevent another manager in a different firm from using the same piece of knowledge. (Contrast this with, say, a can of Coca-Cola: if one person drinks it, no one else can drink it.)

Knowledge is also often nonexcludable. A **nonexcludable good** is one for which it is impossible to selectively deny access. In other words, it is not possible to let some people consume a good while preventing others from consuming it. An excludable good is one to which we can selectively allow or deny access. Once a piece of knowledge is out in the world, it is difficult to prevent others from obtaining access to it. Nobody has patents on basic economic principles of price setting.

Together, these two properties of knowledge mean that a discoverer or inventor of new knowledge may not get all, or even most, of the benefits of that knowledge. As a result, there is insufficient incentive for individuals and firms to try to create new knowledge.

Social Infrastructure

Social infrastructure is a catchall term for the general business environment within a country. Is the country relatively free of corruption? Does it possess a good legal system that protects property rights? In general, is the economy conducive to the establishment and operation of business?

Economists have found that social infrastructure is a critical input into the aggregate production function. Why does it matter so much? When a firm in the United States or another advanced country builds a factory, there is an expectation of revenues generated by this investment that will make the investment profitable. The owners of the firm expect to obtain the profits generated by the activities in that plant. They also expect that the firm has the right to sell the plant should it wish to do so. The firm's owners may confront uncertainty over the *profitability* of the plant—the product manufactured there might not sell, or the firm's managers might miscalculate the costs of production. But it is clear who owns the plant and has the rights to the profits that it generates.

If the owners of firms are unsure if they will obtain these profits, however, they have less incentive to ensure that firms are well managed, and indeed they have less incentive to establish firms in the first place. Output in an economy is then lower. Governments take many actions that influence whether owners will indeed receive the profits from their firms. First, in most countries, governments tax the profits of firms. High tax rates reduce the return on investment. Uncertainty in tax rates also matters because it effectively lowers the return on investment activities. Economists have found that countries with high political turnover tend to be relatively slow growing. One key reason is that frequent changes in political power lead to uncertainty about tax rates.

Governments can also enact more drastic policies. The most extreme example of a policy that affects the return on investment is called *expropriation*—the taking of property by the government without adequate compensation. Although both domestically owned and foreign-owned firms could be subject to expropriation, expropriation is more often about the confiscation of the assets of foreign investors. The World Bank has an entire division dedicated to settling disputes over expropriation.It is called the International Centre for Settlement of Investment Disputes (icsid.worldbank.org/ICSID/Index.jsp). For example, it is arbitrating on a $10 million dispute between a Cypriot investment firm and the government of Turkey: in 2003 the Turkish government seized

without compensation the assets of two hydroelectric utilities that were majority owned by the Cypriot firm. Such settlements can take a long time; at the time of this writing (mid-2011), the dispute has not yet been settled.

There are also more subtle challenges to the rights of foreign investors. Governments may limit the amount of profits that foreign companies can distribute to their shareholders. Governments may limit currency exchanges so that profits cannot be converted from local currencies into dollars or euros. Or governments may establish regulations on foreign-owned firms that increase the cost of doing business. All such actions reduce the attractiveness of countries as places for foreign investors to put their funds.

Economists group these examples under the heading of **property rights**. An individual (or institution) has **property rights** over a resource if, by law, that individual can make all decisions regarding the use of the resource. The return on investment is higher when property rights are protected. In economies without well-established property rights, the anticipated rate of return on investment must be higher to induce firms and households to absorb the investment risks they face.

As a consequence, countries with superior social infrastructure are places where firms will prefer to do business. Conversely, countries that have worse infrastructure are less attractive and will tend to have a lower output. On the website for Dubai at the beginning of the chapter, we see that Dubai touts its free enterprise system, for example. (Dubai's website reveals that physical infrastructure, which is part of the Emirate's capital stock, also plays a critical role: the website touts the superior transport, financial, and telecommunications infrastructure to be found in Dubai.) As another example, Singapore has a project known as Intelligent Nation 2015 that aims to "fuel creativity and innovation among businesses and individuals""Singapore: An Intelligent Nation, A Global City, powered by Infocomm," iN2015, accessed June 29, 2011, www.ida.gov.sg/About%20us/20100611122436.aspx. through improved information technology, including making the entire country Wi-Fi enabled.

An illustration of the importance of social infrastructure comes from the vastly different economic performance of artificially divided economies. At the time that North Korea and South Korea were divided, the two countries were in very similar economic circumstances. Obviously, they did not differ markedly in terms of culture or language. Yet South Korea went on to be one of the big economic success stories of the past few decades, while North Korea is now one of the poorest countries in the world. The experience of East Germany and West Germany is similar: East Germany stagnated under communism, while West Germany prospered.

Natural Resources

There is less to say about what determines the amount of natural resources in the production function. The natural resources available to a country are largely accidents of geography. The United States is fortunate to have high-quality agricultural land, as well as valuable deposits of oil, coal, natural gas, and other minerals. South Africa has deposits of gold and diamonds. Saudi Arabia, Iraq, Kuwait, the United Arab Emirates, and other Middle Eastern countries have large reserves of oil. The United Kingdom and Norway have access to oil and natural gas from the North Sea. For every country, we can list its valuable natural resources.

Natural resources are divided into those that are **renewable** and those that are **nonrenewable**. A **renewable resource** is a resource that regenerates over time. A **nonrenewable (exhaustible) resource** is one that does not regenerate over time. Forests are an example of a renewable resource: with proper management, forests can be maintained over time by judicious logging and replanting. Solar and wind energy are renewable resources. Coal, oil, and minerals are nonrenewable; diamonds taken from the ground can never be replaced.

It is difficult to measure the natural resources that are available to an economy. The availability of oil and mineral reserves is dependent on the technologies for extraction. These technologies have developed rapidly over time. The economic value of these resources, meanwhile, depends on their price in the marketplace. If the price of oil decreases, the value of untapped oil fields decreases as well.

Economists and others sometimes use real gross domestic product (real GDP) as an indicator of economic welfare. One problem with real GDP as an indicator of economic welfare is that it fails to take into account declines in the stock of natural resources. In Chapter 18, we note several of these. If the stock of natural resources is viewed—as it should be—as part of the wealth of a country, then depreciation of that stock should be viewed as a loss in income. (The same argument, incidentally, applies to depreciation of a country's physical capital stock. Real GDP also does not take this into account. However, national accounts do report other statistics that adjust for the depreciation of physical capital, whereas they do not report any adjustment for natural resource depletion.)

Key Takeaways

- The human capital of a country can be accumulated by education, the training of workers, and immigration of workers into that country.
- Knowledge about new products and new processes is created by R&D activities within firms, universities, and government agencies.
- Property rights influence the amount of capital in the aggregate production function. In an economy where property rights are not well defined, there is a lower incentive to invest and hence less capital.

Exercises

1. How does on-the-job experience affect the human capital of an economy?
2. Why is it difficult to measure the natural resources available in an economy?

This page titled 20.5: Other Inputs in the Aggregate Production Function is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

20.6: Accounting for Changes in GDP 20.6: Accounting for Changes in GDPEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What is growth accounting?
2. What are the different time horizons that we use in economics?

We have inventoried the factors that contribute to gross domestic product (GDP). The next step is to understand *how much* each factor contributes. If an economy wants to increase its GDP, is it better off trying to boost domestic savings, attract more capital from other countries, improve its infrastructure, or what? To answer such questions, we introduce a new tool that links the **growth rate** of output to the growth rate of the different inputs to the production function.

Toolkit: Section 31.21

A growth rate is the percentage change in a variable from one year to the next. For example, the growth rate of real GDP is defined as

You can learn more about growth rates in the toolkit.

Some of the inputs to the production function—most notably knowledge, social infrastructure, and natural resources—are very difficult to measure individually. Economists typically group these inputs together into **technology**, as shown in Figure 20.6.1. The term is something of a misnomer because it includes not only technological factors but also social infrastructure, natural resources, and indeed anything that affects real GDP but is not captured by other inputs.

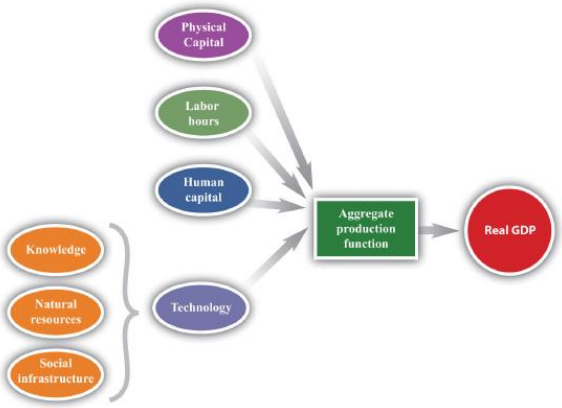


Figure 20.6.1: The Aggregate Production Function

The aggregate production function combines an economy's physical capital stock, labor hours, human capital, and technology (knowledge, natural resources, and social infrastructure) to produce output (real GDP).

The technique for explaining output growth in terms of the growth of inputs is called **growth accounting**.

Toolkit: Section 31.28

Growth accounting tells us how *changes* in real GDP in an economy are due to *changes* in available inputs. Under reasonably general circumstances, the change in output in an economy can be written as follows:

$$output\ growth\ rate = a \times\ capital\ stock\ growth\ rate + [(1-a)times\ labor\ hours\ growth\ rate] + [(1-a)times\ human\ capital\ growth\ rate] + technology\ growth\ rate.$$

In this equation, *a* is just a number. Growth rates can be positive or negative, so we can use the equation to analyze decreases and increases in GDP.

We can measure the growth in output, capital stock, and labor hours using easily available economic data. The growth rate of human capital is trickier to measure, although we can use information on schooling and literacy rates to estimate this number. We also have a way of measuring *a*. The technical details are not important here, but a good measure of $(1 - a)$ is simply total payments to labor in the economy (that is, the total of wages and other compensation) as a fraction of overall GDP. For most economies, *a* is in the range of about 1/3 to 1/2.

For the United States, the number *a* is about 1/3. The growth rate of output is therefore given as follows:

Because we can measure everything in this equation except growth in technology, we can use the equation to determine what the growth rate of technology must be. If we rearrange, we get the following:

To emphasize again, the powerful part of this equation is that we can use observed growth in labor, capital, human capital, and output to infer the growth rate of technology—something that is impossible to measure directly.

Growth Accounting in Action

Table 20.6.1 "Some Examples of Growth Accounting Calculations*" provides information on output growth, capital growth, labor growth, and technology growth. The calculations assume that *a* = 1/3. In the first row, for example, we see that

$$growth\ in\ technology\ =\ 5.5 - [(1/3) \times 6.0] - [(2/3) \times (2.0 + 1.0)] = 5.5 - 2.0 - 2.0 = 1.5.$$

Table 20.6.1: Some Examples of Growth Accounting Calculations *

Year	Output Growth	Capital Growth	Labor Growth	Human Capital Growth	Technology Growth
2010	5.5	6.0	2.0	1.0	1.0
2011	2.0	3.0	1.5	0	0
2012	6.5	4.5	1.0	0.5	
2013	1.5	0	0	2.2	
2014	4.4	3.3		0	1.3

*The figures in each column are percentage growth rates.

Growth accounting is an extremely useful tool because it helps us diagnose the causes of economic success and failure. We can look at successful growing economies and find out if they are growing because they have more capital, labor, or skills or because they have improved their technological know-how. Likewise, we can look at economies in which output has fallen and find out whether declines in capital, labor, or technology are responsible.We use this tool in Chapter 22 to study the behavior of the US economy in the 1920s and 1930s.

Researchers have found that different countries and regions of the world have vastly varying experiences when viewed through the lens of growth accounting. A World Bank study found that, in developing regions of the world, capital accumulation was a key contributor to output growth, accounting for almost two-thirds of total growth in Africa, Latin America, East Asia, and Southeast Asia.The study covered the period 1960–1987. See World Bank, *World Development Report 1991: The Challenge of Development*, vol. 1, p. 45, June 30, 1991, accessed August 22, 2011, econ.worldbank.org/external/default/main?pagePK=64165259&theSitePK=478060&piPK=64165421&menuPK=64166093&entityID=000009265_3981005112648. Technology and human capital growth played a surprisingly small role in these regions, contributing nothing at all to economic growth in Africa and Latin America, for example.

The Short Run, the Long Run, and the Very Long Run

Growth accounting focuses on how inputs—and hence output—change over time. We use the tool both to look at changes in an economy over short time periods—say, from one month to the next—and also over very long time periods—say, over decades. We are limited only by the data that we have available to us. It is sometimes useful to distinguish three different time horizons.

1. The short run refers to a period of time that we would typically measure in months. If something has only a short-run effect on an economy, the effect will vanish within months or a few years at most.
2. The long run refers to periods of time that are better measured in years. If something will happen in the long run, we might have to wait for two, three, or more years before it happens.
3. The very long run refers to periods of time that are best measured in decades.

These definitions of the short, long, and very long runs are not and cannot be very exact. In the context of particular chapters, however, we give more precise definitions to these ideas.For example, Chapter 25 explains the adjustment of prices in an economy. In that chapter, we define the short run as the time horizon in which prices are "sticky"—not all prices have adjusted fully—whereas the long run refers to a period where all prices have fully adjusted. Meanwhile, Chapter 21 uses the very long run to refer to a situation where output and the physical capital stock grow at the same rate. Figure 20.6.2 summarizes the main influences on the inputs to the production function in the short run, the long run, and the very long run.

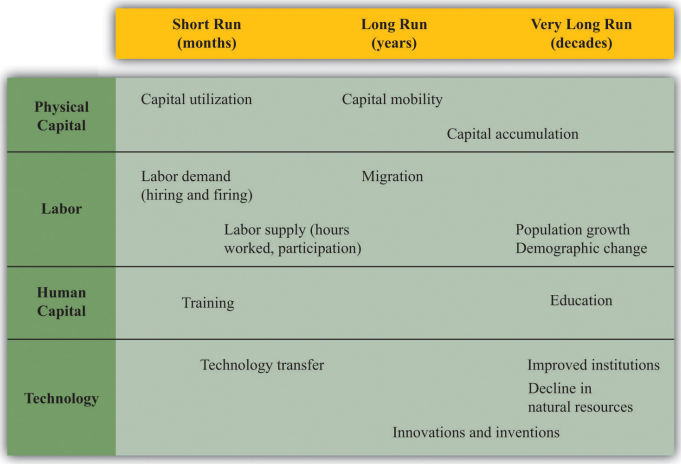


Figure 20.6.2: The Different Time Horizons in Economics

Look first at physical capital—the first row in Figure 20.6.2. In the short run, the amount of physical capital in the economy is more or less fixed. There are a certain number of machines, buildings, and so on, and we cannot make big changes in this capital stock. One thing that firms can do in the short run is to change capital utilization—shutting down a production line if they want to produce less output or running extra shifts if they want more output. Once we move to the long run and very long run, capital mobility and capital accumulation become important.

Look next at labor. In the short run, the amount of labor in the production function depends primarily on how much labor firms want to hire (labor demand) and how much people want to work (labor supply). As we move to the long run, migration of labor becomes significant as well: workers sometimes move from one country to another in search of better jobs. And, in the very long run, population growth and other demographic changes (the aging of the population, the increased entry of women into the labor force, etc.) start to matter.

Human capital can be increased in the long run (and also in the short run to some extent) by training. The most important changes in human capital come in the very long run, however, through improved education.

There is not very much that can be done to change a country's technology in the short run. In the long run, less technologically advanced countries can import better technologies from other countries. In practice, this often happens as a result of a multinational firm establishing operations in a developing country. For example, if Dell Inc. establishes a factory in Mexico, then it effectively transfers some know-how to the Mexican economy. This is known as **technology transfer**, the movement of knowledge and advanced production techniques across national borders.

In the long run and very long run, technology advances through innovation and the hard work of research and development (R&D) that—hopefully—gives us new inventions. In the very long run, countries may be able to improve their institutions and thus create better social infrastructure. In the very long run, declines in natural resources also become significant.

Key Takeaways

- Growth accounting is a tool to decompose economic growth into components of input growth and technological progress.
- In economics, we study changes in GDP over very different time horizons. We look at short-run changes due mainly to changes in hours worked and the utilization of capital stock. We look at long-run changes due to changes in the amount of available labor and capital in an economy. And we look at very-long-run changes due to the accumulation of physical and human capital and changes in social infrastructure and other aspects of technology.

Exercises

1. Rewrite the growth accounting equation for the case where $a = 1/4$.
2. Using the growth accounting equation, fill in the missing numbers in Table 20.6.1 "Some Examples of Growth Accounting Calculations*".

This page titled 20.6: Accounting for Changes in GDP is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

20.7: Globalization and Competitiveness Revisited 20.7: Globalization and Competitiveness RevisitedEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. How is competitiveness measured?
2. What are some of the policies governments use to influence their competitiveness?

At the beginning of this chapter, we noted that both President George W. Bush and President Obama have emphasized policies to improve the competitiveness of the United States. Such interest in national competitiveness is not restricted to the United States.It is perhaps more pronounced in the United States than in other countries. A Google search on October 17, 2011 reveals that the string "Keep America Competitive" has almost twice as many hits as the string "Keep Canada Competitive" and more than three times as many hits as "Keep Britain Competitive." In their "Lisbon Agenda" of 2000, the heads of European countries stated an aim of making the European Union "the most competitive and dynamic knowledge-driven economy by 2010."See "Lisbon Agenda," EurActive, May 21, 2007, accessed July 27, 2011, www.euractiv.com/en/future-eu/lisbon-agenda/article-117510.

Competitiveness: Another Look

Various organizations produce rankings of the competitiveness of countries. For example, IMD, a business school in Switzerland, produces a World Competitiveness Yearbook (WCY) every year.See "World Competitiveness Center," IMD, accessed August 22, 2011, http://www.imd.org/research/centers/wcc/index.cfm. The World Economic Forum (WEF) produces an annual Global Competitiveness Report.See "Global Competitiveness Report," World Economic Forum, accessed June 29, 2011, www.weforum.org/s?s=global+competitiveness. In 2010, the WEF judged Switzerland to be the most competitive economy in the world, followed by the United States and Singapore. According to IMD, the top three were Hong Kong, the United States, and Singapore. These rankings are covered extensively in the business press, and there is also a market for them—WCY resources cost over $1,000. Business and governments purchase these reports each year. National competitiveness is big business.

In their bid to measure competitiveness, the WEF and the WCY look at a combination of "hard" economic data and surveys of businesspeople. Each looks at hundreds of measures in their respective attempts to measure national competitiveness. If these two institutions are to be believed, national competitiveness is a very complicated animal indeed. Although we do not want to go through their measures in detail, a few themes emerge.

- Both the WEF and the WCY look at measures of human capital, such as the number of people enrolled in tertiary education.
- Technology and technological infrastructure feature prominently in both lists of data. The WEF and the WCY look at measures such as the penetration of computers, the Internet, and mobile phones and the granting of patents.
- The quality of public institutions and the prevalence of corruption feature prominently in both lists. Here, the WEF relies on survey data on corruption, bribes, and the extent to which the legal system is fair and transparent. The WCY includes survey information on management practices and "attitudes and values."

Thus the items that we have identified as components of social infrastructure and human capital are included as key determinants of competitiveness. (Technological infrastructure is difficult to classify and measure. In part, it is captured by measures of capital stock because knowledge can be embodied in the capital stock.)

Countries that do better in terms of these rankings will tend to have higher levels of output because these are all inputs into the aggregate production function. The competitiveness of a country is not a matter of how much output it produces, however; we already have a perfectly good measure of that, called real gross domestic product (real GDP). Instead, competitiveness is the ability to attract foreign capital. If countries do not have enough domestic savings to fund investment, then they need to obtain capital from other countries. The amount of capital in the world is limited, so countries compete for this capital by trying to make their economies attractive places to invest. More human capital, better knowledge, or superior social infrastructure all serve to increase the return on investment. If workers are more skilled, then extra capital will generate more output. If firms have better processes in place, then extra capital will generate more output. If a country is free of corruption, then extra capital will generate more output.

This suggests that one good yet simple indicator of national competitiveness is the marginal product of capital. Country A is more competitive than country B if capital investment in country A is more productive than in country B. More exactly, a country is

more competitive if it has a higher marginal product of capital.

Globalization: Another Look

Whenever a good or service is produced and sold, economic value is created. The amount of value is given by the difference between the value to the buyer and the value to the seller. For example, suppose a toy car is produced in a factory in Kansas at a cost of $5. Imagine that a potential buyer in California values the car at $20—that is, she is willing to pay up to this amount for the toy. Then the value created if the buyer and seller trade is $\$20 - \$5 = \$15$.

In a globalized world, toy cars can be transported around the world. This means two things. First, goods can go to where buyers value them the most. There might be a buyer in Germany who values the car at $25. If he buys the car, then the trade creates $20 worth of value ($= \$25 - \$5$). Second, goods can be manufactured where production costs are lowest. Perhaps the toy car can be manufactured in China for $2. If the toy is produced in China and sold in Germany, then the total value created by the trade increases to $23 ($= \$25 - \$2$). Globalization thus contributes to a more efficient global economy because goods—and many services—can be shipped around the world to create more value. They can be produced where it is most efficient to produce them and sold where they are valued the most.

We have also seen that capital (and to a lesser extent labor) moves around the world. Capital moves to competitive economies—that is, to the places where its marginal product is highest. This again contributes to economic efficiency because it means that we (that is, the world as a whole) get more output from a given amount of capital input.

This brief description paints a rosy picture of globalization as a force that makes the world a more productive place. Yet globalization has vehement critics. Protesters have taken to the streets around the world to complain about it. And the recent era of globalization has seen mixed results in terms of economic success. Some economies—particularly in East Asia—have exploited the opportunities of globalization to their advantage. But other countries—most notably in sub-Saharan Africa—remain stuck in poverty.

So what is our story missing? What is wrong with the idea that the free movement of goods and capital can encourage prosperity everywhere? There are some reasons why we should temper our optimism about the process of globalization, including the following:

- **There are winners and losers.** There is a strong presumption from economic theory that globalization will increase overall economic efficiency, but there is no guarantee that everyone will gain. Investors are winners from globalization because they can send their funds to wherever capital earns the highest return. Workers in countries that attract capital will, in general, be winners because they will obtain higher real wages. However, workers in countries that lose capital lose from globalization: they see their real wages decrease. In our example, the buyers of toys in California and Germany benefit from the fact that toys are cheaper and available in greater variety. But the toy manufacturer in Kansas loses out because it cannot compete with the cheaper product from China. The factory may close, and its workers may be forced to look for other—perhaps less attractive—jobs.
- **The playing field is not level.** In an introductory economics book, we do not have room to review the details of trade agreements throughout the world. But one trenchant criticism of globalization is that developed countries have maintained high tariffs and subsidies even as they have encouraged poorer countries to eliminate such measures. As a result, the benefits of globalization have been almost entirely absent for some of the poorest countries in the world. Moreover, rich countries have disproportionate control over some of the key international institutions: the managing director of the International Monetary Fund (IMF) is traditionally a European; the president of the World Bank is appointed by the United States.
- **One size may not fit all.** International institutions such as the IMF and the World Bank typically advocate similar policies for all countries. In fact, different policies might be appropriate for different countries. For example, these organizations argued that countries should allow free movement of capital across their borders. We have seen that there is a strong argument for allowing capital to go in search of the highest return. But not all capital flows take the form of building new factories. Sometimes, the movement of capital consists of only very fast transfers of money in and out of countries, based on guesses about movements in interest rates and exchange rates. These flows of money can be a source of instability in a country. There is increasing recognition that, sometimes at least, it is better to place some limits on such speculative capital movements.

Most economists are convinced that the benefits of globalization are enough to outweigh these concerns. Many—perhaps most—are also convinced that, if globalization is to live up to its promise for the world, it needs to be managed better than it has been in the past.

Policies to Increase Competitiveness and Real Wages

We know that if an economy increases its labor input, other things being equal, the marginal product of labor (and hence the real wage) decreases. If an economy increases its physical capital stock, meanwhile, then the marginal product of capital (and hence the economy's competitiveness) decreases.

There is a critical tension between competitiveness and real wages. Suppose for a moment that human capital and technology are unchanging. Then an economy in which real wages are increasing must also be an economy that is becoming less competitive. Conversely, the only way in which an economy can become more competitive is by seeing its real wages decrease.

High real wages make a country less attractive for businesses—after all, firms choose where to locate in an attempt to make as much profit as possible, so, other things being equal, they prefer to be in low-wage economies. Indeed, the WEF and the WCY both use labor costs as one of their indicators of competitiveness. Our article about Compal locating in Vietnam likewise cited low wages as an attraction of the country.

But we must not be misled by this. High real wages signal prosperity in a country. Low real wages, even if they make an economy competitive and help to attract capital, are not in themselves desirable. After all, the point of attracting capital in the first place is to increase economic well-being. As an example, China has been quite successful at attracting capital, in large part because of low real wages. As the country has become more prosperous, real wages have risen. A *Business Week* article, commenting on the increasing wages in the country, observed the following: "The wage issue has started to affect how companies operate in China. U.S. corporations and their suppliers are starting to rethink where to locate facilities, whether deeper into the interior (where salaries and land values are smaller), or even farther afield, to lower-cost countries such as Vietnam or Indonesia. Already, higher labor costs are beginning to price some manufacturers out of more developed Chinese cities such as Shanghai and Suzhou.""How Rising Wages Are Changing the Game in China," *Bloomberg BusinessWeek*, March 27, 2006, accessed June 29, 2011, http://www.businessweek.com/magazine/content/06_13/b3977049.htm. In other words, increasing real wages are making China less competitive. But this tells us that China is getting richer, and workers in China are able to enjoy improvements in their standard of living. This is a good thing, not a problem.

What we really want are policies that will increase both competitiveness *and* real wages at the same time. The only way to do this is by increasing the stocks of human capital, knowledge, and social infrastructure (there is little a country can do to increase its stock of natural resources). There are no easy or quick ways to increase any of these. Still, important policy options include the following:

- **Invest in education and training.** Overall economic performance depends to a great degree on the education and skills of the workforce. This is one reason why countries throughout the world recognize the need to provide basic education to their citizens. It is worthwhile for countries to build up their stock of human capital just as it is worthwhile for them to build up their stocks of physical capital.
- **Invest in research and development (R&D).** The overall knowledge in an economy is advanced by new inventions and innovations. The romantic vision of invention is that some brilliant person comes up with a completely new idea. There are celebrated examples of this throughout human history, starting perhaps with the cave dweller who had the idea of cracking a nut with a stone and including the individual insights of scientists like Louis Pasteur, Marie Curie, and Albert Einstein. But the reality of invention in the modern economy is more mundane. Inventions and innovations today almost always originate from teams of researchers—sometimes in universities or think tanks or sometimes in the R&D departments of firms. Governments often judge it worthwhile to subsidize such research to help increase the stock of knowledge. R&D expenditures in the United States and other rich countries are substantial; in the United States they amount to about 2 percent of GDP.
- **Encourage technology transfer.** Firms in developed countries tend to have access to state-of-the-art knowledge and techniques. To increase their stock of knowledge, such countries must advance the overall knowledge of the world. For poorer countries in the world, however, there is another possibility. Factories in poor countries typically do not use the most advanced production techniques or have the most modern machinery. These countries can improve their stock of knowledge by importing the latest techniques from other countries. In practice, governments often do this by encouraging multinational firms from rich countries to build factories in their countries. Technology transfer within a country is also important. Researchers have found that, even with a country, there can be big differences in the productivity of different factories within an industry.See Chang-Tai Hsieh and Peter Klenow, "Misallocation and Manufacturing TFP in China and India," *The Quarterly Journal of Economics* CXXIV, no. 4, November 2009, accessed June 28, 2011, http://klenow.com/MMTFP.pdf. So countries may be able to increase real GDP by providing incentives for knowledge sharing across plants.

- **Invest in social infrastructure.** Improvements in social infrastructure are hard to implement. A government, no matter how well intentioned, cannot eliminate corruption overnight. Nor can it instantly establish a reliable legal system that will uphold contracts and protect property rights. (Even if a country could do so, it would still take considerable time for international investors to gain confidence in the system.) Improving social infrastructure is, for most countries, a struggle for the long haul.

We should ask whether government needs to play a role in any of this. After all, individuals have an incentive to invest in their own education. Many people find it worthwhile to pay for undergraduate or graduate degrees because they know they will get better, higher paying jobs afterward. Similarly, firms have a lot of incentive to carry out R&D because a successful invention will allow them to earn higher profits.

There is no doubt that these private incentives play a big role in encouraging the advancement of knowledge. Still, most economists agree that private incentives are not sufficient. Particularly in poor countries, people may not be able to afford to pay for their own education or be able to borrow for that purpose, even if it would eventually pay off for them to do so.

Because knowledge is nonrival and frequently nonexcludable, not all the benefits from R&D flow to the firms that make the investment. For example, suppose a firm comes up with some new software. Other firms may be able to imitate the idea and capture some of the benefits of the invention. (Although the United States and other countries have patent and copyright laws to help ensure that people and firms can enjoy the benefits of their own inventions, such laws are imperfect, and firms sometimes find that their ideas are copied or stolen.) Private markets will do a poor job of providing nonrival and nonexcludable goods, so there is a potential role for the government.

Similar arguments apply to much social infrastructure. The provision of roads is a classic function of government because they are again (most of the time, at least) nonrival and nonexcludable. And the establishment of a reliable legal system is one of the most basic functions of government.

Key Takeaways

- In some leading studies, the items that we have identified as components of social infrastructure and human capital are included as key determinants of competitiveness. Overall, the marginal product of capital is a good indicator of the competitiveness of a country.
- Governments take actions to increase their competitiveness and the real wages of their workers by encouraging the accumulation of human capital, knowledge, and the transfer of technology.

Exercises

1. Why is GDP not a good measure of competitiveness?
2. How could a policy to increase the inflow of capital lead to a decrease in competitiveness? What does this inflow of capital do to the real wage of workers?

This page titled 20.7: Globalization and Competitiveness Revisited is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

20.8: End-of-Chapter Material 20.8: End-of-Chapter MaterialEdit Update titlecancel

In Conclusion

We began the chapter with five stories from all around the world. Let us briefly review these stories, based on what we have learned in this chapter.

Niger

Niger is an extremely poor country. Life expectancy in Niger is 52, the infant mortality rate is over 10 percent, and less than 30 percent of the population can read and write. It is extremely poor because it lacks the key inputs to the production function. It is largely a subsistence agricultural economy: it has relatively little physical capital or human capital, little physical infrastructure, and poor social infrastructure as well. It is a natural target for World Bank help. The particular World Bank project that we cited is aimed at one particular input: its goal is to improve Niger's human capital.

Vietnam

In a globalized world, savings and investment do not have to be equal in any individual economy. Savers can send their funds almost anywhere in the world in search of a high return on capital. Countries that are competitive, in the sense that they have a high marginal product of capital, will tend to attract such funds. One manifestation of these flows of capital is that multinational companies establish factories where they can produce most cheaply. In the story, we see that Vietnam, a low-wage economy, is attracting capital investment from a Taiwanese company. Capital flows have a similar effect to the migration of labor: when capital flows into a country, it increases the real wage; when capital flows out of a country, real wages decrease. Globalization benefits the world as a whole, but many individual workers may lose out.

United Arab Emirates

The policies of Dubai are straightforward to understand in the framework of this chapter. Dubai is actively trying to import foreign physical capital and human capital. It is encouraging multinational firms to establish operations in the country. This makes sense because, as we now know, increased physical and human capital will both tend to increase the marginal product of labor in Dubai, leading to higher wages and higher prosperity. Dubai's claims of attractiveness rest largely on its social infrastructure.

The United Kingdom

Migrant workers are a global phenomenon, be they Poles traveling to England, Mexicans moving to the United States, or Filipinos moving to Saudi Arabia. Like the young Poles in this story, they move from country to country in search of higher wages. Worker migration across national boundaries tends to equalize wages in different countries. As workers leave Poland, for example, labor becomes scarcer there, so wages in Poland tend to increase. When they arrive in the United Kingdom, there is more labor supplied to the United Kingdom labor market, so wages there tend to decrease. However, labor migration is still quite limited because (1) countries restrict immigration and (2) most workers still do not want to suffer the upheaval of moving to a different country and culture.

United States

The competitiveness initiatives of President Obama and President George W. Bush are designed to increase both human capital and knowledge within the United States. They include measures to strengthen education (human capital), increase research and development (R&D; knowledge), and encourage entrepreneurship and innovation. We have seen that the idea of competitiveness is subtle: nations do not compete in the same way that countries do. Still, improvements in technology and human capital will tend to increase the marginal product of capital, making the United States a more attractive place for investment. In that sense, they do make the country more competitive.

Key Links

- World Economic Forum: http://www.weforum.org
- World Competitiveness Yearbook: http://www.imd.ch/wcc
- World Bank: http://www.worldbank.org
- Dubai government: http://www.dubaitourism.ae/node

exercises

Output	Capital	Labor	Human Capital	Technology
10	1	1	10	10
20	2	2	10	10
20	4	1	10	10
20	1	4	10	10
30	9	1	10	10
30	1	9	10	10
30	3	3	10	10
40	2	8	10	10
40	8	2	10	10
40	4	4	10	10
40	4	4	20	5
40	4	4	5	20
80	4	4	20	20

TABLE 20.8.1: AN EXAMPLE OF A PRODUCTION FUNCTION

1. By comparing two different rows in the preceding table, show that the marginal product of labor is positive. Make sure you keep all other inputs the same. In other words, find two rows that show that an increase in labor, *keeping all other inputs the same*, leads to an increase in output.
2. By comparing two different rows in the preceding table, show that the marginal product of human capital is positive. Again, make sure you keep all other inputs the same.
3. By comparing two different rows in the preceding table, show that the marginal product of technology is positive.
4. Does the production function exhibit diminishing marginal product of physical capital? [Hint: if more and more extra capital is needed to generate the same increase in output, then there is diminishing marginal product.]
5. Does the production function exhibit diminishing marginal product of labor?
6. (Difficult) Can you guess what mathematical function we used for the production function?
7. Why are electricians not paid the same amount in Topeka, Kansas, and New York City? Why are electricians not paid the same amount in North Korea and South Korea? Is the explanation the same in both cases?
8. Think about the production function for the university or college where you are studying. What are some of the different inputs that go into it? Classify these inputs as physical capital, human capital, labor, knowledge, natural resources, and social infrastructure. Try to come up with at least one example of each.
9. Suppose government spending is 30, government income from taxes (including transfers) is 50, private saving is 30, and lending to foreign countries is 20. What is national savings? What is investment?
10. Explain how it is possible for investment to be positive yet for the capital stock to fall from one year to the next.
11. Is a fireworks display nonrival? Nonexcludable?
12. Suppose that a country's capital stock growth rate is 8 percent, the labor hours growth rate is 4 percent, the human capital growth rate is 2 percent, and the technology growth rate is 3 percent. The parameter a is 0.25. What is the output growth rate?
13. Suppose that a country's capital stock growth rate is 4 percent, the labor hours growth rate is 3 percent, the human capital growth rate is 1 percent, and the output growth rate is 5 percent. The parameter a is 0.5. What is the technology growth rate?
14. Explain why a decrease in a country's competitiveness can be a sign that the country is becoming more prosperous.

15. Firms are sometimes willing to pay for training courses for their workers. Other things being equal, do you think a firm would prefer to pay for one of its employees to do a general management course or a course that trains the employee in the use of software designed specifically for the firm? Explain.

Economics Detective

1. Go to the website of the Bureau of Labor Statistics (http://www.bls.gov). Find the *median* hourly wage in the state in which you live. (If you do not live in the United States, pick a state at random.)
    1. How does it compare to the median hourly wage for the country as a whole?
    2. Which is higher in your state—the median wage or the mean wage? Can you explain why?
2. Find an example of a competitiveness initiative in some country other than the United States. How will the proposed policies help to attract capital?

This page titled 20.8: End-of-Chapter Material is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

21: Global Prosperity and Global Poverty

21.1: Life Around the World
21.2: The Single-Person Economy
21.3: Four Reasons Why GDP Varies across Countries
21.4: The Accumulation of Physical Capital
21.5: Balanced Growth
21.6: The Role of International Institutions in Promoting Growth
21.7: End-of-Chapter Material

This page titled 21: Global Prosperity and Global Poverty is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

21.1: Life Around the World 21.1: Life Around the WorldEdit Update titlecancel

[…]

I thought about my new friend Mariya, her life and daily routine. She was married before I had my driver's license. She pounded millet all day, sweating yet smiling. She hauled water from the well. She cooked. She birthed child after child. There was no end to the manual labor her life required. I liked to watch her. It was fascinating. But if hers was my life, I'd probably jump into that well.

I've always thought what life each soul is assigned to is a game of chance. I couldn't help but to wonder what would have become of me had the powers that be had shaken those dice one more time on March 16, 1982, before moving the game piece that sent me to DePaul Hospital in St. Louis, Missouri.

If I had been born into Mariya's life, would I have been able to hack it? If she had been born into my life, would she have been happier? […]Alexis Wolff, "Village Life—Niger," February 27, 2005, accessed June 28, 2011, http://www.bootsnall.com/articles/05-02/village-life-niger.html.

In Niger, where Mariya lives, about 1 in 9 children die before their first birthday. Life expectancy at birth is 53 years, and less than 30 percent of the population can read and write. About one-fifth of the population is nomadic. An Oxfam study in 2005 found that nomads had recently lost about 70 percent of their animals, and that "almost one in ten families is surviving on a diet of mainly wild plants, leaves, and grass."See "The World Factbook," Central Intelligence Agency, accessed June 28, 2011, www.cia.gov/library/publications/the-world-factbook/index.html; and "Nomadic Way of Life in Niger Threatened by Food Crisis," Oxfam America, August 16, 2005, accessed June 29, 2011, www.oxfamamerica.org/press/pressreleases/nomadic-way-of-life-in-niger-threatened-by-food-crisis. **Real gross domestic product (real GDP)** per person in Niger is the equivalent of about $700 per year.

Call centers are a phenomenon that has taken over the young crowd of metros in India by a storm. Its implications are social, cultural and economic. It is a new society of the young, rich and free, selling the new dream of an independent life to the regular desi.

[…]

[C]heap labor in India owes its origin to the high rate of unemployment here. Hundreds of thousands of graduates are jobless and desperate for work in India.

[…]

Most call center jobs require a basic understanding of computers and a good grasp over English. And the urban youth of India are computer literate graduates with a command over English language. This is the ideal unskilled labor that the call center industry is looking for.

[…]

With its operations mainly during the night, the call centers offer an opportunity for the young to live a perpetual nocturnal life, a saleable idea to the youth. The fascination of the dark and the forbidden, is tremendous for the Indian youth, recently unleashed from the chains of tradition and culture. Because of this fascination, the industry has developed an air of revolution about itself. Not only is it cool to work for call centers, it is radical and revolutionary.

Just like the bikers subculture of the 60s and the flower children of the 70s, these call centerites also have their own lingo and a unique style of existence. Most of them are happy in a well paying monotonous job, reaping the benefits of technology, enjoying a life away from rush hour traffic and local trains. The moolah is good, the work is easy and life is comfortable."The Indo-American Dream—Coming of Age with Call Center Jobs," Mumbai Travel & Living Guide, accessed June 28, 2011, www.mumbaisuburbs.com/articles/call-centers-mumbai.html.

Life expectancy in India is 67 years, and the infant mortality rate is about 1 in 20. Real GDP per person is about $3,500.

More Americans own pets than ever before, and they're spending more money to keep them healthy, according to a survey released today by the American Veterinary Medical Association.

The number of U.S. households with pets climbed 7.6 million, to 59.5% of all homes, up from 58.3% in 2001. By comparison, about 35% of U.S. households have children, the Census Bureau says.

[…]

Pet owners are spending more on medical care. Veterinary expenditures for all pets were estimated at $24.5 billion in 2006. In inflation-adjusted dollars, Americans spent $22.4 billion in 2001.

This represents “the high-tech care that pet owners are demanding and willing to pay for,” DeHaven says. “Diseases that once would have been difficult to treat—diabetes, heart disease, cancer—today are very treatable. We’re even putting pacemakers in dogs.”Elizabeth Weiss, “We Really Love—and Spend on—Our Pets,” *USA Today*, December 11, 2007, accessed July 29, 2011, www.usatoday.com/life/lifestyle/2007-12-10-pet-survey_N.htm#.

In the United States, where spending on veterinary care for pets is considerably more than twice the entire GDP of Niger, the infant mortality rate is about 1 in 170, and life expectancy is about 78. Real GDP per person is more than 10 times greater than in India and almost 70 times greater than in Niger.

These stories are more than anecdotes. They are, in a real sense, representative of these three countries, as we can see by looking at economic data. Figure 21.1.1 shows real GDP per person in India, the United States, and Niger over the 1960–2009 period.Alan Heston, Robert Summers, and Bettina Aten, “Penn World Table Version 6.2,” Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, September 2006, accessed June 29, 2011, pwt.econ.upenn.edu/php_site/pwt_index.php. The data in the Penn World Tables are constructed so that dollar figures for different countries can be legitimately compared. Specifically, the data are constructed on a purchasing power parity basis, meaning that they take into account the different prices of goods and services in different countries and are based on how much can actually be purchased. From part (a) of Figure 21.1.1, we can see that GDP per person in the United States has grown substantially. On average, real GDP per person grew at 2 percent per year. Perhaps this doesn’t sound like a lot. Economic growth cumulates over time, however. An annual growth rate of 2 percent means that real GDP per person is about 2.6 times higher than half a century ago. To put it another way, each generation is roughly twice as rich as the previous generation. Although there are periods of high and low (sometimes even negative) growth in GDP per person, these fluctuations are overwhelmed by the overall positive growth in our economy. With this growth come many benefits: higher consumption, more varieties of goods, higher quality goods, better medical care, more enjoyable leisure time, and so on.

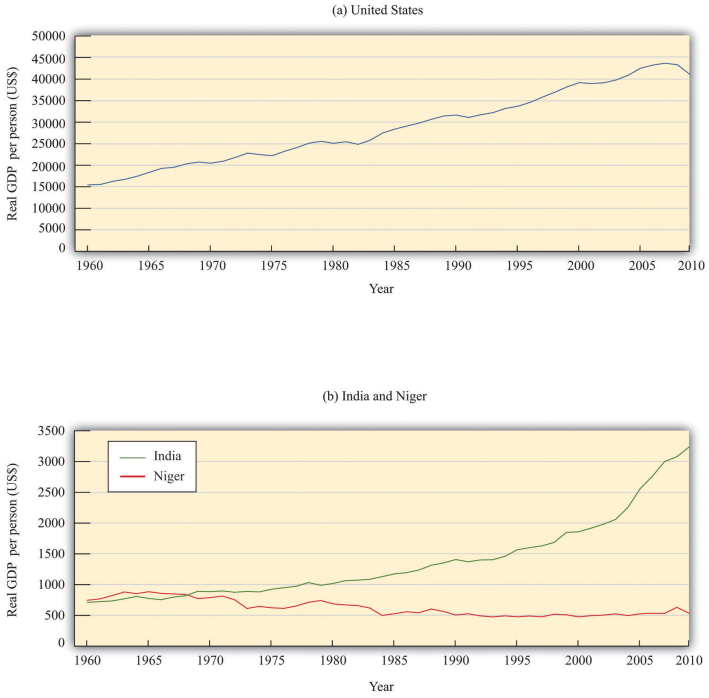


Figure 21.1.1: Real GDP per Person in the United States, India, and Niger. Real GDP per person in the United States (a) is substantially larger than (b) real GDP in India and Niger. The growth experiences of the three countries are also very different

Source: Alan Heston, Robert Summers and Bettina Aten, Penn World Table Version 7.0, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, May 2011.

Part (b) of Figure 21.1.1 shows real GDP per person for India and Niger. Notice first that the scale on this graph is very different. In 1960, real GDP per person in the United States was about $15,000 (measured in year 2005 dollars). In Niger and India, it was about 5 percent of the US figure—about $700 per person. The second striking feature of this graph is the very different performance of

India and Niger. India, like the United States, has grown: GDP per person is much higher at the end of the sample than at the beginning. Indeed, India has grown faster than the United States: the average growth rate over the period was 3.1 percent. Over the last two decades, the difference is even starker: India has grown at about 4.4 percent per year on average. Nevertheless, the United States is still a lot richer than India.

By world standards, India is a long way from being the poorest country. In 1960, Niger was richer than India on a per person basis. But in the following half century, Niger became poorer, not richer. GDP per person decreased by almost 30 percent. India in 2009 was six times richer than Niger. Statistics on GDP are just that—statistics—and it is easy to look at graphs like these and forget that they are telling us about the welfare of human beings. But imagine for a moment that Niger had managed to grow like India, instead of collapsing as it did. People would not be surviving by eating grass, infants would be more likely to grow up to be adults instead of dying of preventable diseases, and children would be learning to read and write.

This is why the study of economic growth matters. And this is why, in this chapter, we take on arguably the most important question in the entire book.

Why are some countries rich and other countries poor?

Along the way, we tackle two other closely related questions. We want to know if the differences in income that we see in the world are likely to persist over time. The experiences of the United States, India, and Niger suggest that this question may not have a simple answer: India has been tending to catch up with the United States, but Niger has been falling further behind. As we seek to answer that question, we will also investigate the ultimate sources of economic growth:

Will poorer countries catch up to richer countries?

Why do countries grow?

Road Map

The big mystery we investigate in this chapter is the vast variation in economic performance from country to country. We want to know why the experiences of the United States, India, and Niger are so different from one another.

We begin this chapter with an extended story. We think about how growth would work in a country with just a single inhabitant. Then we turn to a back-of-the-envelope calculation to understand why countries differ so much in terms of economic performance. To understand these differences, we focus attention on different inputs to the production function, first considering physical capital and then looking at human capital and technology. After that, we develop a complete framework for understanding how and why economies grow in the very long run. Finally, we look at policy and international institutions.

This page titled 21.1: Life Around the World is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

21.2: The Single-Person Economy 21.2: The Single-Person EconomyEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. How does the capital stock increase?
2. What are the factors that lead to output growth?
3. What are the differences between growth in a closed economy and growth in an open economy?

The macroeconomy is very complicated. Overall economic performance depends on billions of decisions made daily by millions of people. Economists have developed techniques to keep us from being overwhelmed by the sheer scale of the economy and the masses of data that are available to us. One of our favorite devices is to imagine what an economy would look like if it contained only one person. This fiction has two nice features: we do not have to worry about differences among individuals, and we can easily isolate the most important economic decisions. Thinking about the economy as if it were a single person is only a starting point, but it is an extremely useful trick for cutting through all the complexities of, say, a $12 trillion economy populated by 300 million individuals.

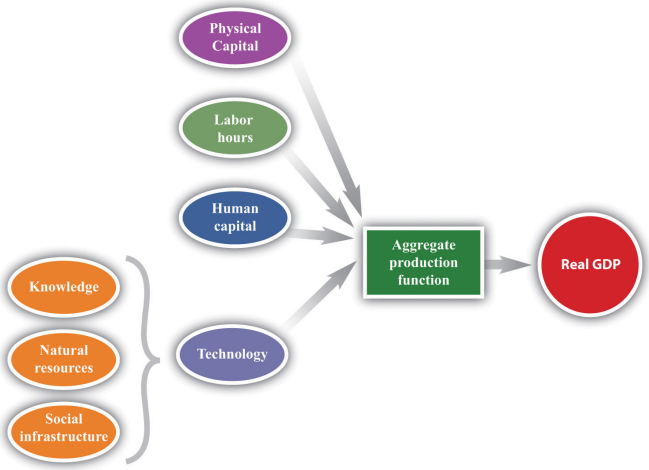


Figure 21.2.1: The Aggregate Production Function. The aggregate production function combines an economy's physical capital stock, labor hours, human capital, and technology to produce output (real gross domestic product [real GDP]).

Solovenia

Imagine, then, an economy called Solovenia. Solovenia is populated by one individual—we will call him Juan. Juan has access to an **aggregate production function**. The amount of output (real GDP) that he can produce depends on how large a physical **capital stock** he owns, how many hours he chooses to work, his human capital, and his technology (Figure 21.2.1). **Physical capital** is the stock of factories and machinery in the economy, while **human capital** refers to the skills and education of the workforce. **Technology** is a catchall term for everything else (other than capital, labor, or human capital) that affects output.Physical capital, human capital, and technology are discussed in more detail in Chapter 20. It includes the following:

- **Knowledge.** The technological know-how of the economy
- **Social infrastructure.** The institutions and social structures that allow a country to produce its real GDP
- **Natural resources.** The land and mineral resources in the country

Toolkit: Section 31.26

You can review the aggregate production function, including its inputs, in the toolkit.

Much of our focus in this chapter is on how economies build up their stock of physical capital. Figure 21.2.2 shows how output in the aggregate production function depends on the capital stock. Increases in the capital stock lead to more output. If Juan has more

tools to work with, then he can produce more goods. However, we usually think that the production function will exhibit **diminishing marginal product of capital**, which means that a given increase in the capital stock contributes more to output when the capital stock is low than when the capital stock is high. In Figure 21.2.2, we can see this from the fact that the production function gets flatter as the amount of physical capital increases.

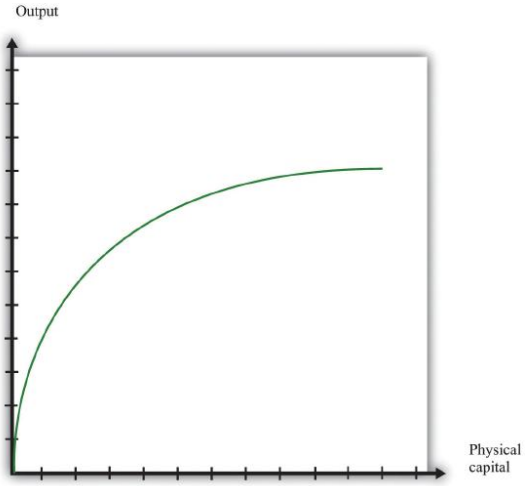


Figure 21.2.2: The Aggregate Production Function: Output as a Function of the Physical Capital Stock. As the amount of physical capital increases, output increases, but at a decreasing rate because of the diminishing marginal product of capital.

Each day Juan chooses how much time to work and how much time to spend in leisure. Other things being equal, we expect that Juan likes to have leisure time. This is not to say that Juan never gets any satisfaction from working. But like most people—even those who enjoy their jobs—he would prefer to work a little bit less and play a little bit more. He cannot spend all his time in leisure, however. He works because he likes to consume. The harder he works, the more real GDP he can produce and consume. Juan's decision about how many hours to work each day is determined in large part by how productive he can be—that is, how much real GDP he can produce for each hour of leisure time that he gives up.

Juan does not have to consume all the output that he produces; he might save some of it for the future. As well as deciding how much to work, he decides how much to consume and how much to save each day. You have probably made decisions like Juan's. At some time in your life, you may have worked at a job—perhaps in a fast-food restaurant, a grocery store, or a coffee shop. Perhaps you were paid weekly. Then each week you might have spent all the money you earned on movies, meals out, or clothes. Or—like Juan—you might have decided to spend only some of that money and save some for the future. When you save money instead of spending it, you are choosing to consume goods and services at some future date instead of right now. You may choose to forgo movies and clothes today to save for the purchase of a car or a vacation.

The choice we have just described—consuming versus saving—is one of the most fundamental decisions in macroeconomics. It comes up again and again when we study the macroeconomy. Just as you and Juan make this choice, so does the overall economy. Of course, the economy doesn't literally make its own decision about how much to save. Instead, the saving decisions of each individual household in the economy determine the overall amount of savings in the economy. And the economy as a whole doesn't save the way you do—by putting money in a bank. An economy saves by devoting some of its production to capital goods rather than consumer goods. If Juan chooses to produce capital goods, he will have a larger capital stock in the future, which will allow him to be more productive and enjoy higher consumption in the future.

Growth in a Closed Economy

At any given moment, Juan's ability to produce output is largely determined by his stock of physical capital, his human capital, and the state of technology in Solovenia. But, as time passes, the level of output in Solovenia can change through a variety of mechanisms.

First, the capital stock in Solovenia can grow over time, as shown in Figure 21.2.3. Juan builds up his capital stock by saving. Since Juan is the only inhabitant, the amount he saves is equal to the **national savings** of Solovenia. It is the difference between his output (real GDP of Solovenia) and the amount he consumes.In a real economy, national savings also include the savings of government: we must add in the government surplus or subtract the government deficit, as appropriate.

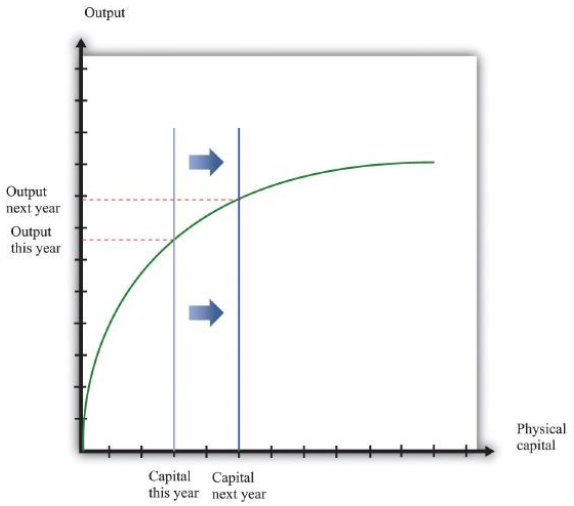


Figure 21.2.3. Increases in the capital stock lead to increases in output. If the capital stock in Solovenia increases between this year and next year, output also increases. Increases in the capital stock are one source of growth.

The more that Juan saves today, the more he can build up his capital stock, and the higher his future standard of living will be. If Juan chooses to consume less today, he will have a higher living standard in the future. If Juan chooses to consume more today, he must accept that this means less consumption in the future. Economies, like individuals, can choose between eating their cake now or saving it for the future.

In making this decision, Juan weighs the cost of giving up a little bit of consumption today against the benefit of having a little bit more consumption in the future. The higher the marginal product of capital, the more future benefit he gets from sacrificing consumption today. Other things being equal, a higher marginal product of capital induces Juan to save more. Juan's choice also depends on how patient or impatient he is. The more patient he is, the more he is willing to give up consumption today to enjoy more consumption in the future.

Increases in the amount of physical capital are one way in which an economy can grow. Another is through increases in human capital and technology. These shift the production function upward, as shown in Figure 21.2.4. Perhaps Juan sometimes has better ideas about how to do things. Perhaps he gets better with practice. Perhaps Juan spends some time trying to come up with better ways of producing things.

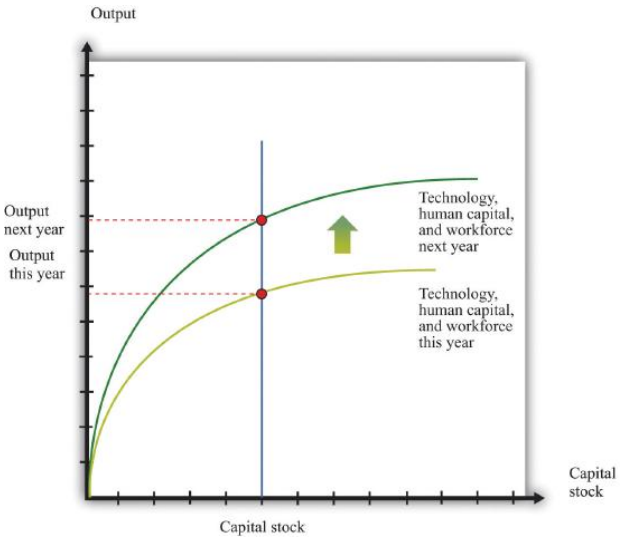


Figure 21.2.4. Increases in human capital or technology lead to increases in output. Increases in technology, human capital, and the workforce, like increases in the capital stock, are a source of output growth.

Through the accumulation of physical and human capital, and by improving the components of technology such as knowledge and social infrastructure, the output in Solovenia will grow over time. The combined effect of physical capital growth and improvements in technology is shown in Figure 21.2.5.

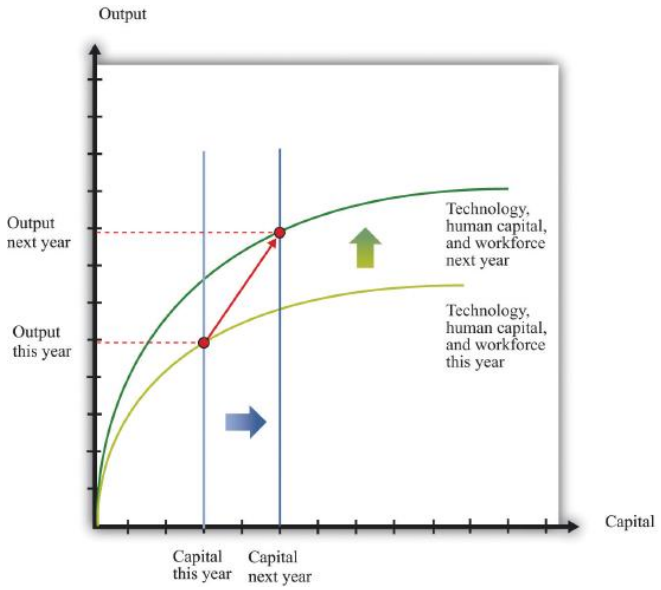


Figure 21.2.5. Increases in capital, human capital, and technology all lead to increases in output. In general, economies grow because of increases in capital, technology, human capital, and the workforce.

Growth in an Open Economy

If Juan does not trade with the rest of the world, his only way to save for the future is by building up his capital stock. In this case, national savings equal **investment**. An economy that does not trade with other countries is called a **closed economy**. An economy that trades with other countries is called an **open economy**. In the modern world, no economy is completely closed, although some economies (such as Belgium) are much more open than others (such as North Korea). The world as a whole is a closed economy, of course.

If Solovenia is an open economy, Juan has other options. He might decide that he can get a better return on his savings by investing in foreign assets (such as Italian real estate, shares of Australian firms, or Korean government bonds). Domestic investment would then be less than national savings. Juan is lending to the rest of the world.

Alternatively, Juan might think that the benefits of investment in his home economy are sufficiently high that he borrows from the rest of the world to finance investment above and beyond the amount of his savings. Domestic investment is then greater than national savings. Of course, if Juan lends to the rest of the world, then he will have extra resources in the future when those loans are repaid. If he borrows from the rest of the world, he will need to pay off that loan at some point in the future.

There may be very good opportunities in an economy that justify a lot of investment. In this case, it is worthwhile for an economy to borrow from other countries to supplement its own savings and build up the capital stock faster. Even though the economy will have to pay off those loans in the future, the benefits from the higher capital stock are worth it.

The **circular flow of income** shows us how these flows show up in the national accounts. If we are borrowing on net from other countries, there is another source of funds in additional to national savings that can be used for domestic investment. If we are lending on net to other countries, domestic investment is reduced.

Toolkit: Section 31.27

You can review the circular flow of income in the toolkit.

$$investment = national\ savings + borrowing\ from\ other\ countries$$

or

$$investment = national\ savings - lending\ to\ other\ countries.$$

Savings and investment in a country are linked, but they are not the same thing. The savings rate tells us how much an economy is setting aside for the future. But when studying the accumulation of capital in an economy, we look at the **investment rate** rather than the savings rate. Total investment as a fraction of GDP is called the investment rate:

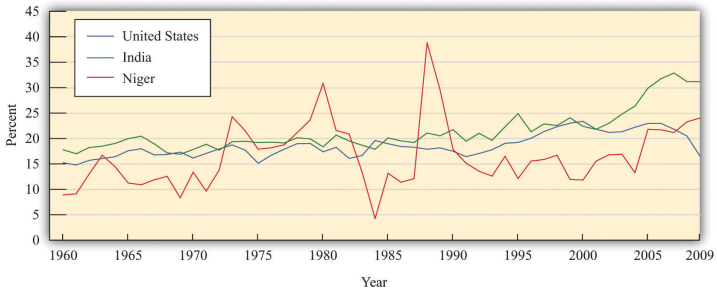


Figure 21.2.6: Investment Rates in the United States, India, and Niger. There are marked differences in investment rates in the United States, India, and Niger.

Source: Alan Heston, Robert Summers and Bettina Aten, Penn World Table Version 7.0, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, May 2011.

Figure 21.2.6 shows investment rates in the United States, India, and Niger from 1960 to 2009. A number of features of this picture are striking:International Monetary Fund, World Economic Outlook Database, April 2011, accessed July 29, 2011, www.imf.org/external/pubs/ft/weo/2011/01/weodata/index.aspx.

- For most of the period, India had a higher investment rate than the other two countries. As we saw earlier, India was also the fastest growing of the three countries. These facts are connected: capital accumulation plays an important role in the growth process.
- The investment rate in the United States has been relatively flat over time, though it has been noticeably lower in recent years.
- Investment rates in Niger have been more volatile than in the other two countries. They were low in the mid 1980s but have increased substantially in recent years.

Low investment rates may be due to low savings rates. They may also reflect relatively low returns to increases in the capital stock in a country. The low investment rate that prevailed for many years in Niger not only reflected a low saving rate but also indicates that something is limiting investment from external sources. For the United States, in contrast, a significant part of the high investment rate is due not to domestic savings but to inflows from other countries.

We know that output per person is a useful indicator of living standards. Increases in output per person generally translate into increases in material standards of living. But to the extent that an economy trades with other countries, the two are not equivalent. If an economy borrows to finance its investment, output per person will exaggerate living standards in the country because it does not take into account outstanding obligations to other countries. If an economy places some of its savings elsewhere, then measures of output per person will understate living standards.The national accounts deal with this issue by distinguishing between GDP, which measures the production that takes place within a country's borders, and gross national product (GNP), which corrects for income received from or paid to other countries.

Solovenia and Solovakia

Now add another country, Solovakia (with a single inhabitant named Una), and compare it to Solovenia. We can compare Solovenia and Solovakia by investigating which is producing more output per person and why. Imagine, for example, that Solovenia is a relatively poor country, and Solovakia is richer. Using our knowledge of the aggregate production function, we can understand how this difference might arise. It might be because Una has more human capital or knowledge than Juan, or because Una has a larger stock of physical capital.

Another basis for comparison is the rate at which the two economies are growing. If Solovakia is richer, and if it is also growing faster than Solovenia, then the gap between the two countries will become wider over time. We call such a process **divergence**. Conversely, if Solovenia is growing faster than Solovakia, then the gap between Juan's and Una's living standards will become smaller over time. Such a situation, where poorer countries catch up to richer ones, is called **convergence**.

Why might we see either convergence or divergence? Part of the answer has to do with the **marginal product of capital** in the two countries. Suppose that Solovakia is richer because it has a larger stock of physical capital than Solovenia. In that case, we expect the marginal product of capital to be larger in Solovenia. Solovenia is a more competitive economy than Solovakia. Juan will want to invest at home, while Una will take some of the output that she produces in Solovakia and invest it in Solovenia. Therefore we expect capital to migrate from Solovakia to Solovenia. As a consequence, it is likely that Solovenia will grow faster than Solovakia, leading to convergence.

Key Takeaways

- The capital stock increases through investment.
- Because physical capital is an input in the aggregate production function, growth in capital stock is one source of output growth. The other sources are the accumulation of human capital and increases in technology.
- In a closed economy, investment equals national savings. In an open economy, investment equals national savings plus inflows of funds from abroad. So in an open economy, growth in the capital stock and hence output growth can be financed both by domestic savings and borrowing from other countries.

Exercises

1. Draw a version of Figure 21.2.2 with labor hours instead of physical capital on the horizontal axis. Explain how the figure illustrates the positive marginal product of labor and diminishing marginal product of labor. How would you illustrate a change in the capital stock using this figure?
2. Explain how an economy can have an investment rate of 10 percent but a savings rate of only 3 percent.

This page titled 21.2: The Single-Person Economy is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

21.3: Four Reasons Why GDP Varies across Countries 21.3: Four Reasons Why GDP Varies across CountriesEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What are the main possible explanations for real GDP differences across countries?
2. How important are differences in technology for explaining differences in real GDP across countries?

We started this chapter with the following question: "Why are some countries rich and other countries poor?" The aggregate production function and the story of Juan help us to understand what determines the amount of output that an economy can produce, taking us the first step toward explaining why some countries are richer than others.

The production function tells us that if we know four things—the size of the workforce, the amount of physical capital, the amount of human capital, and the level of technology—then we know how much output we are producing. When comparing two countries, if we find that one country has more physical capital, more labor, a better educated and trained workforce (that is, more human capital), and superior technology, then we know that country will have more output.

Differences in these inputs are often easy to observe. Large countries obviously have bigger workforces than small countries. Rich countries have more and better capital goods. In the farmlands of France, you see tractors and expensive farm machinery, while you see plows pulled by oxen in Vietnam; in Hong Kong, you see skyscrapers and fancy office buildings, while the tallest building in Burkina Faso is about 12 stories high; in the suburbs of the United States, you see large houses, while you see shacks made of cardboard and corrugated iron in the Philippines. Similarly, rich countries often have well-equipped schools, sophisticated training facilities, and fine universities, whereas poorer countries provide only basic education. We want to be able to say more, however. We would like to know *how much* these different inputs contribute to overall economic performance.

To get some sense of this, we look at some rough numbers for the United States, India, and Niger. We carried out this exercise using data from 2003, but the fundamental message does not depend on the year that we have chosen; we would get very similar conclusions with data from any recent year. To start, let us look at the different levels of output in these countries. Table 21.3.1 "Real GDP in the United States, India, and Niger" gives real gross domestic product (real GDP) in these countries. Note that we are now looking at the overall level of GDP, rather than GDP per person as we did at the beginning of this chapter. Real GDP in the United States was about $10.2 trillion. In India, real GDP was about one-third of US GDP: $3.1 trillion. In Niger, real GDP was under $10 billion. In other words, the United States produces about 1,000 times as much output as Niger.

Table 21.3.1: Real GDP in the United States, India, and Niger

Country	Real GDP in 2003 (Billions of Year 2000 US Dollars)
United States	10,205
India	3,138
Niger	9

Source: Alan Heston, Robert Summers and Bettina Aten, Penn World Table Version 7.0, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, May 2011.

In the following subsections, we look at how the different inputs contribute to bring about these large differences in output. We go through a series of thought experiments in which we imagine putting the amount of each input available in the United States into the production functions for the Indian and Niger economies.

Differences in the Workforce across Countries

The United States, India, and Niger differ in many ways. One is simply the number of people in each country. The workforce in the United States is about 150 million people. The workforce in India is more than three times greater—about 478 million in 2010—while the workforce in Niger is only about 5 million people. Thus India has much more labor to put into its production function than does Niger.

In Table 21.3.2 "Real GDP in 2003 in the United States, India, and Niger if All Three Countries Had the Same Workforce" we look at what would happen to output in India and Niger if—counterfactually—each had a workforce the size of that in the United States while their other inputs were unchanged. Output in the United States is, of course, unchanged in this experiment. India's output would decrease to about $1.4 trillion because they would have a smaller workforce. Niger's output would increase about tenfold to $88 trillion. Differences in the workforce obviously matter but do not explain all or even most of the variation across the three countries. Niger's output would still be less than 1 percent of output in the United States.

Table 21.3.2: Real GDP in 2003 in the United States, India, and Niger if All Three Countries Had the Same Workforce

Country	Real GDP in 2003 (Billions of Year 2000 US Dollars)
United States	10,205
India	1,475
Niger	88

Differences in Physical Capital across Countries

Not surprisingly, the United States also has a much larger capital stock than does Niger. The capital stock in the United States is worth about $30 trillion. India's capital stock is about $3 trillion, and Niger's capital stock is much, much smaller—about $9 billion. So what would happen if we also gave India and Niger the same amount of physical capital as the United States? Table 21.3.3 "Real GDP in the United States, India, and Niger if All Three Countries Had the Same Workforce and Physical Capital Stock" shows the answer.

India's GDP, in this thought experiment, goes back to something close to its actual value of around $3 trillion. In other words, the extra capital compensates for the smaller workforce. Real GDP in the United States is still more than three times larger than that in India. The extra capital makes a big difference in Niger, increasing its output about ten-fold. Even if Niger had the same size workforce and the same amount of capital as the United States, however, it would still have only a tenth of the amount of output. The other two inputs—human capital and technology—evidently matter as well.

Table 21.3.3: Real GDP in the United States, India, and Niger if All Three Countries Had the Same Workforce and Physical Capital Stock

Country	Real GDP in 2003 (Billions of Year 2000 US Dollars)
United States	10,205
India	3,054
Niger	1,304

Differences in Human Capital across Countries

Differences in education and skills certainly help to explain some of the differences among countries. Researchers have found evidence that measures of educational performance are correlated with GDP per person. The causality almost certainly runs in both directions: education levels are low in Niger because the country is so poor, and the country is poor because education is low.

We can include measures of education and training in an attempt to measure the skills of the workforce. In fact, economists Robert Hall and Chad Jones have constructed a measure that allows us to compare the amount of human capital in different countries.To estimate relative human capital levels in different countries, we use the figures in Robert Hall and Chad Jones, "Why Do Some Countries Produce So Much More Output per Worker Than Others?" *Quarterly Journal of Economics* 114, no. 1 (1999): 83–116. In Table 21.2.4 "Real GDP in 2003 in the United States, India, and Niger if All Three Countries Had the Same Workforce, Physical Capital Stock, and Human Capital Stock", we bring the human capital level in India and Niger up to the level in the United States and, as before, suppose that all three countries have the same amount of labor and physical capital. Real GDP in India would climb to about $5.2 trillion, or a little over half the level in United States. Niger's real GDP would equal about $2.8 trillion, meaning the increased human capital would more than double Niger's GDP. However, real GDP in the United States would still be more than three times greater than that of Niger.

Table 21.3.4: Real GDP in 2003 in the United States, India, and Niger if All Three Countries Had the Same Workforce, Physical Capital Stock, and Human Capital Stock

Country	Real GDP in 2003 (Billions of Year 2000 US Dollars)

Country	Real GDP in 2003 (Billions of Year 2000 US Dollars)
United States	10,205
India	5,170
Niger	2,758

Differences in Technology across Countries

To summarize, even after we eliminate differences in labor, physical capital, and human capital, much is still left to be explained. According to our production function, the remaining variation is accounted for by differences in technology—our catchall term for everything apart from labor, physical capital, and human capital.

Just as firms accumulate physical capital, they also accumulate knowledge in various ways. Large firms in developed countries *develop* new knowledge through the activities of their research and development (R&D) divisions.Gains in productivity of this form sometimes end up embodied in capital stock—think of a computer operating system, such as Windows or Linux. Such knowledge increases the value of capital stock and is already captured by looking at the ratio of capital stock to GDP. In poorer countries, firms may access existing knowledge by importing technology from more developed countries.

Differences in knowledge help to explain differences in output per worker. The rich countries of the world tend to have access to state-of-the-art production techniques. We say that they are on the **technology frontier**; they use the most advanced production technologies available. Factories in poor countries often do not use these production techniques and lack modern machinery. They are *inside* the technology frontier.

As economists have researched the differences in economic performance in rich and poor countries, they have found that success depends on more than physical capital, human capital, and knowledge. Appropriate institutions—the social infrastructure—also need to be in place. These are institutions that allow people to hold property and write and enforce contracts that ensure they can enjoy the fruits of their investment. Key ingredients are a basic rule of law and a relative lack of corruption. An ability to contract and trade in relatively free markets is also important.

Particularly in more advanced countries, we need the right institutions to encourage technological progress. This is complicated because there is a trade-off between policies to encourage the *creation* of knowledge and policies to encourage the *dissemination* of knowledge. Knowledge is typically a **nonexcludable good**, so individuals and firms are not guaranteed the rights to new knowledge that they create. This reduces the incentive to produce knowledge. To counter this problem, governments establish certain property rights over new knowledge, in the form of patent and copyright laws. Knowledge is also typically a **nonrival good**, so everyone can, in principle, benefit from a given piece of knowledge. Once new knowledge exists, the best thing to do is to give it away for free. Patent and copyright laws are good for encouraging the development of knowledge but bad for encouraging the dissemination of knowledge. Current debates over intellectual property rights (file sharing, open source, downloading of music, etc.) reflect this trade-off.

Differences in **natural resources** can also play a role in explaining economic performance. Some countries are lucky enough to possess large amounts of valuable resources. Obvious examples are oil-producing states such as Saudi Arabia, Kuwait, Venezuela, the United States, and the United Kingdom. Yet there are many countries with considerable natural resources that have not enjoyed great prosperity. Niger's uranium deposits, for example, have not helped that country very much. At the same time, some places with very little in the way of natural resources have been very successful economically: examples include Luxembourg and Hong Kong. Natural resources help, but they are not necessary for economic success, nor do they guarantee it.

Key Takeaways

- Differences in real GDP across countries can come from differences in population, physical capital, human capital, and technology.
- After controlling for differences in labor, physical capital, and human capital, a significant difference in real GDP across countries remains.

Exercises

1. In Table 21.3.2 "Real GDP in 2003 in the United States, India, and Niger if All Three Countries Had the Same Workforce", Table 21.3.3 "Real GDP in the United States, India, and Niger if All Three Countries Had the Same Workforce and Physical Capital Stock", and Table 21.3.4 "Real GDP in 2003 in the United States, India, and Niger if All Three Countries Had the Same Workforce, Physical Capital Stock, and Human Capital Stock", the level of real GDP for the United States is the same as it is in Table 21.3.1 "Real GDP in the United States, India, and Niger". Why is this the case?
2. What kinds of information would help you measure differences in human capital?
3. How can human capital and knowledge flow from one country to another?

This page titled 21.3: Four Reasons Why GDP Varies across Countries is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

21.4: The Accumulation of Physical Capital 21.4: The Accumulation of Physical CapitalEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What factors determine the growth rate of the capital stock?
2. Will poorer countries catch up to richer countries because of faster growth of capital stock?
3. What is the evidence on convergence?
4. Why might countries not converge?

Our first task in this chapter was to explain the vast differences in living standards that we observe in the world. We now know that this variation is due to differences in physical capital, human capital, and technology. The rough calculations in 21.2 Section "Four Reasons Why GDP Varies across Countries" tell us that variations in physical capital, human capital, and technology *all* play a role in explaining differences in economic performance.

Now we consider these inputs separately. In this section, we look at the accumulation of physical capital. In 21.4 Section "Balanced Growth", we look at the role of human capital and technology. Our main aim is to consider one of our two remaining questions:

Will poorer countries catch up to richer countries?

The Growth Rate of the Capital Stock

Capital goods are goods such as factories, machines, and trucks. They are used for the production of other goods and are not completely used up in the production process. Economies build up their capital stocks by devoting some of their gross domestic product (GDP) to new capital goods—that is, investment. As we saw in our discussion of Solovenia in 21.2 Section "Four Reasons Why GDP Varies across Countries", if a country does not interact much with other countries (that is, it is a closed economy) the amount of investment reflects savings within a country. In open economies, the amount of investment reflects the perceived benefits to investment in that country compared to other countries.

Capital goods wear out over time and have to be scrapped and replaced. A simple way to think about this depreciation is to imagine that a fraction of the capital stock wears out every year. A reasonable average **depreciation rate** for the US economy is 4 or 5 percent. To understand what this means, think about an economy where the capital stock consists of a large number of identical machines. A depreciation rate of 5 percent means that for every 100 machines in the economy, 5 machines must be replaced every year.The depreciation rate can be understood in terms of the average lifetime of a typical machine. For example, a depreciation rate of 5 percent is the same as saying that, on average, machines last for 20 years. To see this, imagine that capital stock is kept constant at 100 machines, and each machine lasts for 20 years. Imagine also that 5 machines are 1 year old, 5 machines are 2 years old, and so forth, with the oldest 5 machines being 20 years old. Each year, these 5 oldest machines would wear out (5 percent depreciation) and have to be replaced by 5 new machines. After a year has passed, the situation will be exactly the same as the previous year: there will be 5 machines that are 1 year old, 5 machines that are 2 years, and so forth. Mathematically, we are saying that the lifetime of a machine = 1/depreciation rate: $20 = 1/0.05$.

The depreciation of capital goods reduces the capital stock. The total amount of capital lost to depreciation each year is calculated by multiplying the depreciation rate and the capital stock together. If the capital stock is \$30 trillion, for example, and the depreciation rate is 5 percent, then \$1.5 trillion ($= \$30\ trillion \times 0.05$) worth of capital is lost each year.

The capital stock increases as long as there is enough new investment to replace the worn out capital and still contribute some extra. The overall change in the capital stock is equal to new investment minus depreciation:

$$change\ in\ capital\ stock\ =\ new\ investment\ -\ depreciation\ rate\ \times capital\ stock.$$

For example, suppose that the current capital stock (measured in trillions of dollars) is 40, and the depreciation rate is 10 percent per year. Then the capital stock after depreciation is $40 - (.1 \times 40) = 40 - 4 = 36$. Suppose that new investment is \$4.8 trillion. Then the new capital stock is 36 + 4.8 = 40.8. In this case, capital stock has increased by \$0.8 trillion, or 2 percent.

The equation for the change in the capital stock is one of the fundamental ingredients of economic growth. It tells us that economies build up their capital stock—and therefore their real GDP—by devoting enough output to new investment to both

replace worn out capital and then add some more. If we divide both sides of the previous equation by the capital stock, we can obtain the **growth rate** of the capital stock. (Remember that the growth rate of a variable is the change in the variable divided by its initial level.)

The growth rate of the capital stock depends on three things:

1. **The amount of investment.** The more investment the economy carries out, the more quickly the capital stock grows.
2. **The current capital stock.** The larger the capital stock, other things being equal, the lower its growth rate.
3. **The depreciation rate.** If existing capital wears out faster, the capital stock grows more slowly.

It is intuitive that a higher investment rate increases the growth rate of the capital stock, and a higher depreciation rate decreases the growth rate of the capital stock. It is less obvious why the growth rate of the capital stock is lower when the capital stock is higher. The growth rate measures the change in the capital stock *as a percentage of the existing capital stock*. A given change in the capital stock results in a smaller growth rate if the existing capital stock is larger. For example, suppose that the current capital stock is 100, and the change in the capital stock is 10. Then the growth rate is 10 percent. But if the current capital stock is 1,000, then the same change of 10 in the capital stock represents only a 1 percent growth rate.

Toolkit: Section 31.21

The toolkit contains more information on how growth rates are calculated.

Convergence through Capital Accumulation

Why are we so interested in the accumulation of capital? One reason is that poverty of the kind we observe in Niger and elsewhere is a massive problem for the world. About 40 percent of the world's population—close to 2.5 billion people—live in conditions of poverty. (The World Bank defines poverty as living on less than US$2 per day.) We are not going to solve the problem of mass poverty overnight, so we would like to know whether this gap between the rich and the poor is a permanent feature of the world. It might be that economies will diverge, meaning that the disparities in living standards will get worse and worse, or it might be that they will converge, with poorer countries catching up to richer countries.

When comparing two countries, if we find that the poorer economy is growing faster than the richer one, then the two are converging. If we find that the richer country is growing faster than the poorer one, they are diverging. Moreover, if a country has a small capital stock, we know that—other things being equal—it will tend to be a poorer country. If a country has a large capital stock, then—again, other things being equal—it is likely to be a richer country. The question of convergence then becomes: other things being equal, do we expect a country with a small capital stock to grow faster than an economy with a large capital stock?

The answer is yes, and the reason is the marginal product of capital. From the production function, the marginal product of capital is large when the capital stock is small. Think again about Juan in Solovenia. A large marginal product of capital means that he can obtain a lot of extra output if he acquires some extra capital. This gives him an incentive to save rather than consume. A large marginal product of capital also means that Juan can attract investment from other countries.

A country where the marginal product of capital is high is a competitive economy—one where both domestic savers and foreign savers want to build up the capital stock. The capital stock will grow quickly in such an economy. This is precisely what we saw in the equation for the growth rate of the capital stock: higher investment and a lower capital stock both lead to a larger capital stock growth rate. Both of these imply that a country with a large marginal product of capital will tend to grow fast.

We illustrate this idea in Figure 21.4.1. Country A has a small capital stock. The aggregate production function tells us that this translates into a large marginal product of capital—the production function is steep. In turn, a large marginal product of capital means that country A will grow quickly. Country B has an identical production function but a larger capital stock, so the marginal product of capital is lower in country B than in country A. There is less incentive to invest, implying that country B, while richer than country A, grows more slowly.

Figure 21.4.1 also shows that it is possible for a country to have such a large capital stock that it shrinks rather than grows. Country C has so much capital that its marginal product is very low. There is little incentive to build up the capital stock, so the capital stock depreciates faster than it is replaced by new investment. In such an economy, the capital stock and output would decrease over time.

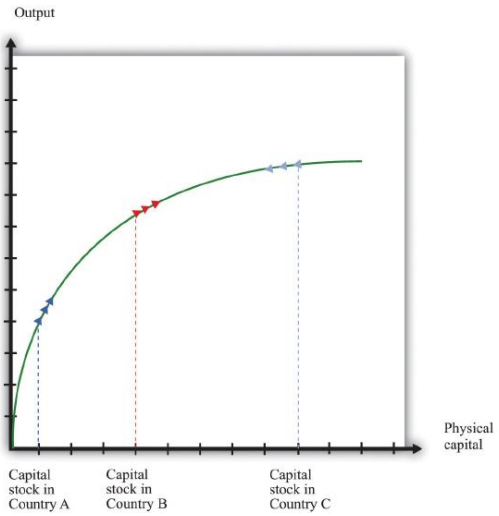


Figure 21.4.1: Convergence through the Accumulation of Capital. The growth rate of the capital stock depends on the marginal product of capital. Country A has little capital, so the marginal product of capital is large, and the capital stock will grow rapidly. Country B has more capital, so the capital stock grows more slowly. Country C has so much capital that the capital stock decreases.

Figure 21.4.1 suggests an even stronger conclusion: all three economies will ultimately end up at the same capital stock and the same level of output—complete convergence. This conclusion is half right. If the three economies were identical except for their capital stocks and if there were no growth in human capital and technology, they would indeed converge to exactly the same level of capital stock and output. In 21.4 Section "Balanced Growth", we look at this argument more carefully. First, though, we examine the evidence on convergence.

Convergence or Divergence? Two Contrasting Pictures

Convergence is a very pretty theory but is it borne out by the evidence? Figure 21.4.2 shows the growth experience of several countries in the second half of the 20th century. These countries are all members of the Organisation for Economic Co-operation and Development (OECD) and are, relatively speaking, rich.The countries are Australia, Austria, Belgium, Canada, Denmark, Finland, France, Iceland, Ireland, Italy, the Netherlands, Norway, Portugal, Spain, Switzerland, and the United States. The median real GDP per capita in 1950 for these countries was about $6,000, in year 1996 dollars. Data for Figure 21.4.2 and Figure 21.4.3 come from Alan Heston, Robert Summers, and Bettina Aten, “Penn World Table Version 6.2,” Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, September 2006, accessed June 29, 2011, pwt.econ.upenn.edu/php_site/pwt_index.php. Figure 21.4.2 shows real GDP per person in these countries *relative to the United States* (the United States itself is the horizontal line near the top of the figure.) Figure 21.4.2 does show some evidence of convergence. Countries with higher levels of real GDP person in 1950 tended to grow more slowly than countries with lower levels of real GDP per person. Poorer countries in this group tended to catch up with richer countries.

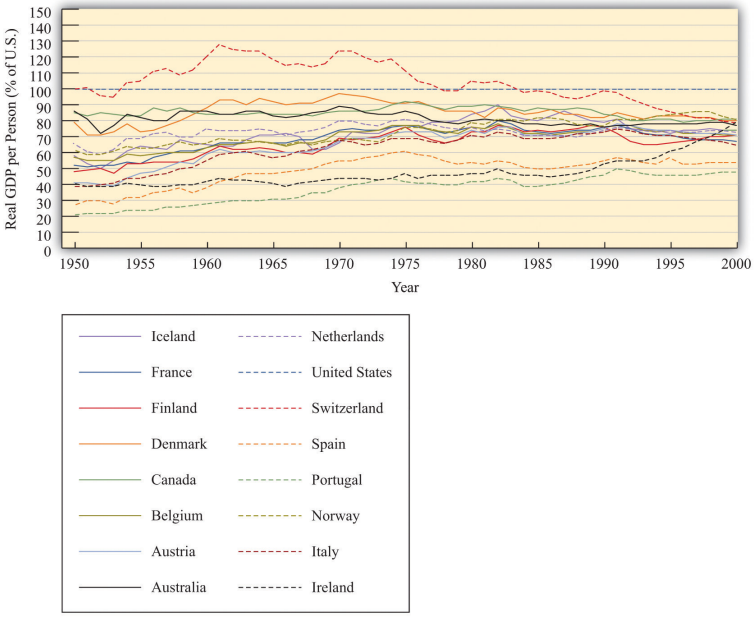


Figure 21.4.2: Some Evidence of Convergence

The growth experience of 16 relatively developed countries, measured as real GDP per person relative to the level in the United States, shows considerable evidence of convergence.

So far so good. But Figure 21.4.3 shows the growth experience over the same period for a more diverse group of countries. This group is largely composed of poorer countries. The picture here is very different: we do not see convergence. There is no evidence that the poorer countries are growing faster than the richer countries. In some cases, there even appears to be divergence: poor countries growing more slowly than rich countries so that output levels in rich and poor countries move further apart.

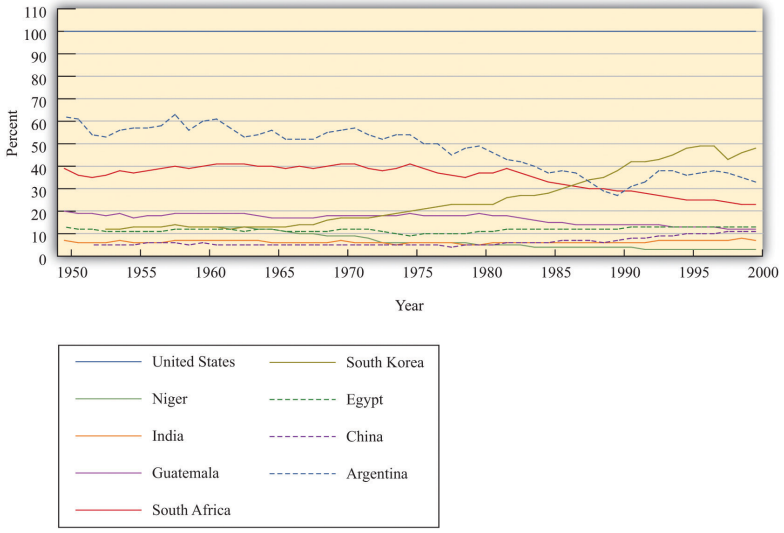


Figure 21.4.3: *Some Evidence of Divergence*

In contrast to Figure 21.4.2, the countries in this sample do not appear to be converging. Many countries that were poor in 1950 were just as poor, relatively speaking, in 2000.

Table 21.4.1 "Evidence from Select Countries" shows more data for some of these countries. It lists the level of initial GDP per person and the average growth rate in GDP per person between the early 1950s and the end of the century. For example, Argentina had real GDP per person of $6,430 in 1950 (in year 1996 dollars) and grew at an average rate of 1.25 percent over the 50-year period. Egypt and South Korea had very close levels of GDP per person in the early 1950s, but growth in South Korea was much higher than that in Egypt: by the year 2000, GDP per person was $15,876 in South Korea but only $4,184 in Egypt. These two countries very clearly diverged rather than converged. Looking at China, the level of GDP per person in the early 1950s was less than 10 percent that of Argentina. By 2000, GDP per person in China was about 33 percent of that in Argentina.

Table 21.4.1: Evidence from Select Countries

Country (Starting Year)	Real GDP per Capita (Year 1996 US Dollars)	Percentage Average Growth Rate to 2000
Argentina (1950)	6,430	1.25
Egypt (1950)	1,371	2.33
China (1952)	584	4.0
South Korea (1953)	1,328	5.5

Source: Penn World Tables

Overall, this evidence suggests that our theory can explain the behavior over time of some but not all countries. If we look at relatively rich countries, then we do see evidence of convergence. Across broader groups of countries, we do not see convergence, and we see some evidence of divergence.

Explaining Divergence

Why is it that, contrary to what Figure 21.4.1 seems to suggest, not all countries converge? The logic of that picture rests on the diminishing marginal product of capital. If rich countries have lower marginal product of capital than poor countries, then we expect poor countries to catch up. If, for some reason, richer countries sometimes also have a higher marginal product of capital than poorer countries, then the argument for convergence disappears.

Figure 21.4.4 shows an example where the aggregate production function looks a bit different. This production function has a range where increases in capital stock lead to a higher rather than a lower marginal product of capital. That is, for some amounts of capital, we see increasing marginal product of capital rather than diminishing marginal product of capital. In the figure, country A and country B converge, just as in our previous diagram. But country C is rich enough to lie on the other side of the range where there is an increasing marginal product of capital. Country C therefore has a higher marginal product of capital than country B, even though country C is richer. Countries B and C will diverge, rather than converge.

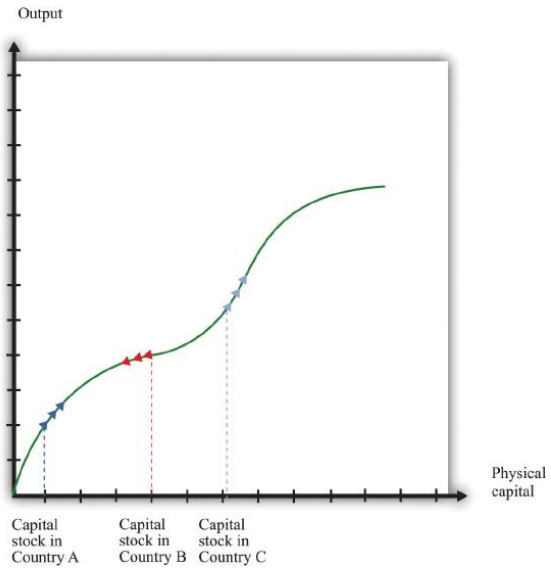


Figure 21.4.4: Divergence Arising from Increasing Marginal Product of Capital. In this diagram, three countries have an aggregate production function that does not always exhibit diminishing marginal product of capital. As a result, the economies need not converge.

Figure 21.4.5 shows another reason why a richer country might have a higher marginal product of capital than a smaller country. In Figure 21.4.1 we supposed that the three countries had the same production function and differed only in terms of their capital stock. In Figure 21.4.5, country B is richer than country A for two reasons: it has more capital *and* has a superior technology (or more labor or human capital). The higher capital stock, other things being equal, means a lower marginal product of capital in country B. But the superior technology, other things being equal, means a higher marginal product of capital in country B. In the picture we have drawn, the technology effect dominates. Country B has the higher marginal product of capital, so it is the more

attractive location for capital—it is more competitive. Because of this, the capital stock increases in country B. Indeed, if the only factor driving investment is the marginal product of capital, then we would expect capital to flow among countries until the marginal product of capital is equal everywhere.We discuss capital migration across countries in more detail in Chapter 20.

One reason why a richer economy might have better technology is because it has better social infrastructure. In particular, developed economies often have the legal and cultural institutions that preserve **property rights**. The return on investment is higher, other things being equal, when property rights are protected. In economies with less well-developed institutions, investors need a higher rate of return to compensate them for the additional risk of placing their capital in those countries.

Measuring these aspects of social infrastructure is a challenge. The World Bank has attempted to do so in its 2005 *World Development Report*.World Bank, *World Development Report 2005: A Better Investment Climate for Everyone* (New York: World Bank and Oxford University Press, 2004), 8, accessed August 22, 2011, siteresources.worldbank.org/INTWDR2005/Resources/complete_report.pdf. The study looks at various aspects of doing business in 48 countries. The top constraints on investment reported by firms were policy uncertainty, macroeconomic instability, and taxes. Many of the risks of doing business are directly associated with government action in the present and in the future. This is nicely stated in the World Bank report: "Because investment decisions are forward looking, firms' judgments about the future are critical. Many risks for firms, including uncertain responses by customers and competitors, are a normal part of investment, and firms should bear them. But governments have an important role to play in maintaining a stable and secure environment, including by protecting property rights. Policy uncertainty, macroeconomic instability, and arbitrary regulation can also cloud opportunities and chill incentives to invest. Indeed, policy-related risks are the main concern of firms in developing countries."World Bank, *World Development Report 2005: A Better Investment Climate for Everyone* (New York: World Bank and Oxford University Press, 2004), 5, accessed August 22, 2011, siteresources.worldbank.org/INTWDR2005/Resources/complete_report.pdf.

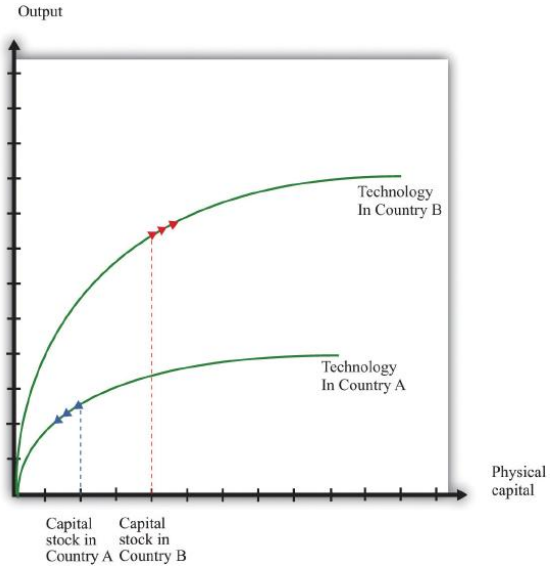


Figure 21.4.5: Divergence Arising from Differences in Technology. In this diagram, country B has a better technology or more human capital than country A. Even though country B has a larger capital stock, it also has a larger marginal product of capital.

Key Takeaways

- Capital stock increases from investment and decreases due to the depreciation of capital stock.
- All else being the same, poorer countries have a lower capital stock and therefore a higher marginal product of capital compared to rich countries. Thus capital accumulation should be faster in poor countries, which will lead to convergence with richer countries.
- The evidence suggests convergence between some but not all economies.
- Divergence of output across countries might come from the presence of an increasing marginal product of capital or from one country having a superior technology to another.

Exercises

1. Suppose we have 100 units of capital stock at the beginning of 2012 and the following table gives the investment for the next 5 years. Suppose the depreciation rate is 5 percent. Fill in the blanks in the table for the years 2012–2017.

Year	Capital Stock (Start of Year)	Investment	Depreciation
2012	100	80	
2013		20	
2014		50	
2015		120	
2016		10	

2. If one country has a higher level of real GDP than another, does that mean it must have a higher growth rate as well?
3. If citizens of a relatively poor country are educated in a richer country, does this help or hinder convergence?

This page titled 21.4: The Accumulation of Physical Capital is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

21.5: Balanced Growth 21.5: Balanced GrowthEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What is balanced growth?
2. Why does balanced growth matter?
3. When will economies converge to a balanced-growth path?

We have seen that the accumulation of capital—other things being equal—leads economies to converge over time. However, we saw that the evidence for such convergence in the data is highly mixed. To understand more about when economies will and will not converge, we need a more complete theory of the sources of economic growth. In this section, we develop such a theory and then use it to look again at the question of convergence. We initially take as given—that is, as **exogenous**—the growth rates of human capital, the workforce, and the technology.

Growth Accounting

We begin with the tool of **growth accounting**. The growth accounting equation for our aggregate production function is as follows:Growth accounting is discussed in more detail in Chapter 20.

$$\begin{aligned} &output\ growth\ rate \\ =\ &[a \times (capital\ growth\ rate)] + [(1-a) \times (workforce\ growth\ rate + \ human\ capital\ growth\ rate)] + technology\ growth\ rate. \end{aligned}$$

Toolkit: Section 31.28

You can review growth accounting in the toolkit.

In this equation, *a* is just a number. For the US economy, *a* is approximately equal to 1/3. Remember that output is just another term for real gross domestic product (real GDP).

It turns out that, in the very long run, we expect the capital stock and the level of output to grow at exactly the same rate. We see why later in this section. Such a situation is called **balanced growth**. When this is true, the growth accounting equation then becomesYou don't need to worry about the mathematical details, but if you are interested, we obtain this equation by setting the capital growth rate equal to the output growth rate:

$$\begin{aligned} &output\ growth\ rate^{BG} = [a \times (output\ growth\ rate^{BG})] \\ &+ [(1-a) \times (workforce\ growth\ rate + \ human\ capital\ growth\ rate)] + \ technology\ growth\ rate, \end{aligned}$$

which implies

$$\begin{aligned} &(1-a) \times output\ growth\ rate^{BG} \\ &= [(1-a) \times (workforce\ growth\ rate \ + \ human\ capital\ growth\ rate)] + \ technology\ growth\ rate. \end{aligned}$$

Dividing this equation by (1 − *a*) gives us the equation in the text.

For example, suppose that *a* = 1/3, the human capital growth rate = 0.01, the technology growth rate = 0.02, and the workforce growth rate = 0.03. Then

The Growth Rate of Output per Worker in a Balanced-Growth Economy

When we are comparing living standards across countries, it is better to adjust for differences in the size of the workforce to obtain output per worker. This is a measure of the overall **productivity** of an economy—that is, the effectiveness of an economy for producing output. (Of course, output per worker and output per person are very closely related. For the US economy, the workforce is roughly half the total population, so output per person is therefore approximately half as much as output per worker.) The growth rate of output per worker equals the growth rate of output minus the growth rate of the workforce:

This equation tells us that, in the end, the secret to economic growth is the development of knowledge and skills. Invention, innovation, education, training, and improvements in social infrastructure are the drivers of economic growth in the very long run.

Perhaps surprisingly, the growth rate of the capital stock is not a fundamental determinant of the growth rate. When we have balanced growth, the capital stock grows, which contributes to the overall growth of output. But if we ask what determines the overall growth rate in an economy, it is the growth of technology and human capital. The capital stock then adjusts to keep the economy on its balanced-growth path. By the definition of balanced growth, the growth rate of the capital stock is equal to the output growth rate.

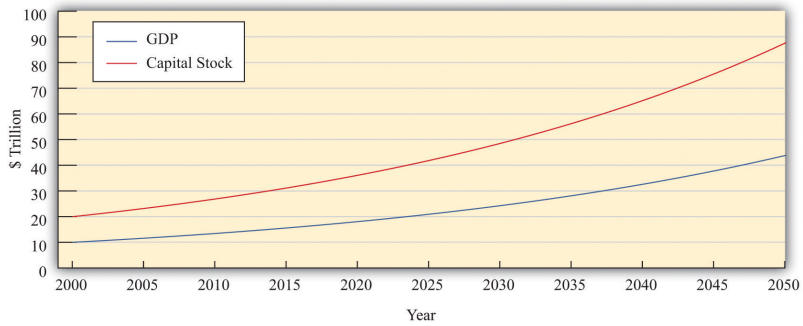


Figure 21.5.1: Output and Capital Stock in a Balanced-Growth Economy. This picture shows an example of an economy on a balanced-growth path. Both variables grow at 3 percent per year and the capital stock is always equal to exactly twice the level of GDP.

Figure 21.5.1 illustrates balanced growth. Look first at output. Notice that even though the *growth rate* of output is constant, the graph is not a straight line. Instead, it curves upward: the change in the *level* of output increases over time. This is because a growth rate is a percentage change. In our example, output in 2000 is $10 trillion, and the growth rate is 3 percent. From 2000 to 2001, output increases by $300 billion (= $10 trillion × 0.03). By 2050, output is equal to $44 trillion. Between that year and the next, output increases by $1.3 trillion (= $44 trillion × 0.03). Even though the growth rate is the same, the change in the level of output is more than four times as large.

When output and the capital stock grow at the same rate, the *ratio* of the capital stock to GDP does not change. In Figure 21.5.1, the value of the capital stock is always twice the value of output. The capital stock and real GDP both grow at the same rate (3 percent per year), so the ratio of the capital stock to GDP does not change over time.

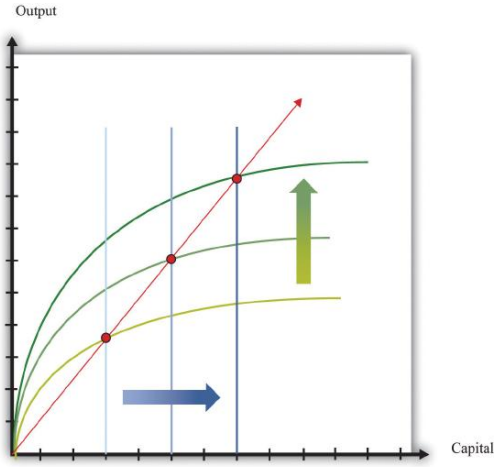


Figure 21.5.2. Balanced growth means that the ratio of the capital stock to output does not change. On a balanced-growth path, output and the capital stock grow at the same rate, so the ratio of the capital stock to output is always the same: the growth path of the economy is a straight line from the origin.

Figure 21.5.2 shows what a constant ratio of the capital stock to GDP looks like in our production function diagrams. Along any straight line from the origin, the ratio of the capital stock to output does not change. As a simple example, suppose that (as in Figure 21.5.1) the capital stock is always twice the level of output. This means that output is always half of the capital stock:

$$output \ = \ 0.5 \times capital \ stock.$$

This is just the equation of a straight line that passes through the origin. In Figure 21.5.2, increases in human capital or technology shift the production function upward. On the balanced-growth path, capital stock grows at exactly the right rate so that the economy grows along a straight line from the origin.

The Transition to Balanced Growth

If an economy is not yet on its balanced-growth path, it will tend to go toward that path. If a country has a small capital stock relative to GDP, then its capital stock will grow faster than real GDP. Countries that are still developing may well be in this position. Countries that are further along in the development process are likely to be (approximately) on their balanced-growth paths. For such countries, the ratio of capital stock to output is unchanging.

Economies that have not yet accumulated enough capital to be on their balanced-growth paths will have a growth rate that equals the balanced-growth rate *plus* an additional factor due to the growth rate of capital relative to GDP.If you are interested in the mathematical derivation of this equation, you can find it in the toolkit.

The first term is the growth rate along the balanced-growth path. The second term is the *additional* component to growth that comes about whenever the capital stock is growing faster than output.

Table 21.5.1 "Approaching the Balanced-Growth Path" gives an example of an economy that is approaching a balanced-growth path. Like the economy in Figure 21.5.1, the balanced-growth output growth rate is 3 percent. The workforce grows at 1 percent, so output per worker grows at 2 percent along the balanced-growth path. However, this economy starts off (in the year 2000) with a smaller capital stock than is needed for balanced growth. Looking at the first row of the table, you can see that the capital stock grows at 14.4 percent, while output grows at 6.8 percent. Because capital grows faster than output, there is an additional component to growth, as we have just explained. This contributes an extra 3.8 percentage points to the growth rate, so output per worker grows at 5.8 percent.

As time goes on, the capital stock grows relative to output, and the economy gets closer to the balanced-growth path. As this happens, the additional component of growth becomes smaller. For example, in 2010, the capital stock grows at 6.8 percent, and output grows at 4.3 percent. The growth rate of output per worker is 3.3 percent—2 percentage points being the balanced-growth contribution and 1.3 percent due to the faster growth rate of capital stock compared to output. By 2050, the economy is close to balanced growth: output per worker grows at 2.3 percent, with capital stock growing only a little bit faster than output.

Table 21.5.1: Approaching the Balanced-Growth Path

Year	Balanced-Growth Output Growth Rate (%)	Balanced-Growth Output per Worker Growth Rate (%)	Capital Growth Rate (%)	Output Growth Rate (%)	Output per Worker Growth Rate (%)
2000	3.0	2.0	14.4	6.8	5.8
2005	3.0	2.0	9.3	5.1	4.1
2010	3.0	2.0	6.8	4.3	3.3
2015	3.0	2.0	5.5	3.8	2.8
2020	3.0	2.0	4.7	3.6	2.6
2025	3.0	2.0	4.1	3.4	2.4
…	…	…	…	…	…
2050	3.0	2.0	3.8	3.3	2.3

Countries that are well below their growth path will see their capital stock grow rapidly relative to GDP. They will experience relatively rapid GDP growth. Countries that are close to their balanced-growth path will see their capital stock grow more slowly relative to GDP and have a GDP growth rate that is only slightly bigger than the balanced-growth rate. Although the economy will eventually reach its balanced-growth value, this adjustment may take decades. For this reason, we say that the economy will achieve balanced growth only in the very long run.To be mathematically precise, the economy gets closer and closer to its balanced-growth path but never quite gets there. Over a period of decades, it gets close enough that it makes no practical difference.

Convergence Revisited

We can now use our theory of balanced growth to make our earlier argument for convergence more precise. Then we consider whether we might also see convergence from changes in human capital and technology.

Convergence in Physical Capital

Imagine that we are comparing two countries that are identical in almost every respect. They both have the same levels of technology and human capital and the same balanced-growth ratio of capital stock to GDP. However, they have different amounts of physical capital. Suppose that one of the countries has a large capital stock (call it the rich country) and the other country has a much smaller capital stock (call it the poor country).

These two economies will initially have different levels of output and living standards. Our model predicts, however, that these differences will be temporary. Both economies will approach the balanced-growth path. The poor country will grow more rapidly because its ratio of capital stock to GDP will be increasing more quickly as it moves toward the balanced-growth path. Over time, we expect to see the poor country catch up to the rich one. We illustrate this in Figure 21.5.3.

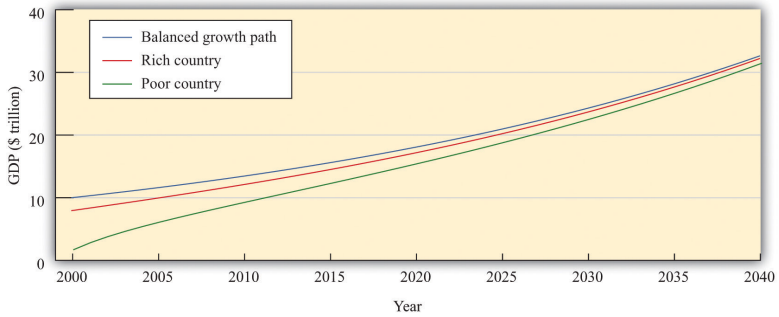


Figure 21.5.3: Convergence of a Rich Country and a Poor Country. Consider two economies, identical in all respects except that one has a smaller capital stock than the other. The poorer country accumulates capital faster than the richer country and grows faster.

This is exactly the same mechanism for convergence that we saw before. The country with a smaller capital stock will have a higher marginal product of capital and will grow faster because the country is a more attractive place for investment. Because the poor country accumulates capital more rapidly than the richer country, it will grow faster. The two countries will converge to the same balanced-growth path and to the same level of output per person.

Convergence in Human Capital

So far we have not considered why human capital might change over time. If there are reasons to think that this variable might grow more quickly in poor countries than in rich countries, we have another force that might drive convergence.

In some ways, human capital resembles physical capital. As with the physical capital stock, some accumulation is the result of decisions by governments, and some comes from decisions by private agents. From the government side, it is likely that economies with low levels of human capital might also be economies in which there is a high return to basic education. If literacy rates are low and most children do not receive much education, even straightforward investments in schooling might yield big gains in terms of the ultimate capabilities of the workforce. Governments in poor countries might see big potential gains from investment in education. Private individuals and firms may also perceive that the returns on education are larger in poorer economies. If very few people in the economy have college degrees, an individual might find that a college education yields a very large payoff. By contrast, if the population as a whole is highly educated, it might take a much larger investment to stand out from others.

This discussion is somewhat speculative. Human capital is difficult to measure, and the marginal product of human capital is even harder to quantify. Nevertheless, there are some good reasons to believe that the incentives to invest in human capital are greater in poorer economies. If so, we have another reason to expect convergence.

Convergence in Technology

What about technology? Will it grow faster in poorer economies? The answer depends on which aspect of technology we are talking about.

Differences in knowledge between rich and poor countries are likely to diminish over time. Rich economies are typically close to the technology frontier, meaning that they are using state-of-the-art production techniques. For countries on the technology frontier, growth in knowledge can only come the hard way, through investment in research and development (R&D). Countries inside the technology frontier are typically poorer developing countries. These economies can grow their stock of knowledge simply by importing knowledge from countries at the frontier. Technological advance is much cheaper and easier if you can use others' inventions and innovations rather than coming up with your own. We therefore expect such countries to have faster growth rates of

knowledge. As they become more developed, the growth of knowledge in these economies will slow down to the rate experienced by other countries near the technology frontier. But in the meantime, they will grow faster than rich countries. Technology transfer to developing economies is surely a force leading to convergence of economies.

There is less to say about social infrastructure and natural resources. The amount of natural resources available in an economy is largely due to accidents of history and geography: there is no obvious reason to expect the growth rate of natural resources to be linked to the level of development. Social infrastructure, meanwhile, is a complicated mix of institutions, customs, and other factors. Again, there is no obvious reason to expect social infrastructure to grow more quickly in poorer economies.

Divergence Revisited

Now that we have incorporated human capital and technology into our framework, we can identify some further possible explanations of divergence. Our theory says that economies will converge if they differ only in terms of their initial capital stock. But it is possible that different economies will also have different balanced-growth paths. Figure 21.5.4 shows what this looks like. The ratio of capital stock to output in the very long run depends on a number of different factors, including the growth rate of technology and the growth rate of the workforce. If these differ across countries, then their balanced-growth paths will differ as well, and we will not observe convergence.The toolkit presents a complete model of balanced growth, including a formula for the balanced-growth ratio of capital stock to output.

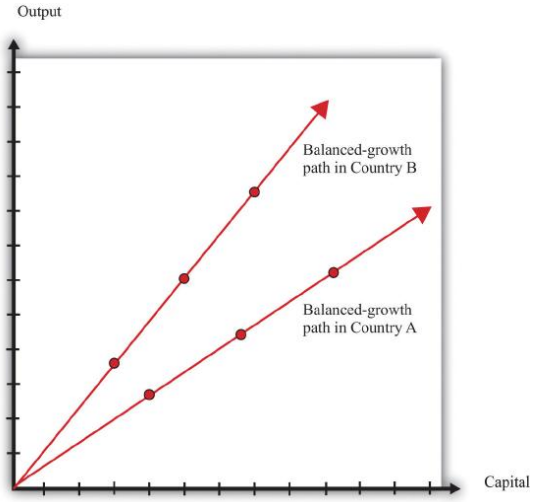


Figure 21.5.4: Balanced Growth in Two Countries with Different Ratios of Capital to Output. Economies may have different balanced-growth paths. In this example, the ratio of capital stock to output is higher in country A than in country B.

This explanation—and our previous stories of divergence—tells us why different economies will not necessarily end up at exactly the same level of output per worker. But the problem of divergence is in some ways worse than that. Some countries are not only failing to converge but also moving further and further apart. In other words, in some cases, richer economies are growing faster than poorer economies. Indeed, as we saw with Niger, some of the poorest economies in the world have been shrinking rather than growing.

Remember that the growth rate of output per worker on a balanced-growth path is as follows:

We can explain divergence in our framework if human capital or technology is growing more slowly in poor countries than in rich ones. Are there reasons to expect this to be the case? Earlier, we said that countries with low levels of human capital might also be countries where the return to human capital investment was large, which is a force for convergence. We also pointed out, however, that the marginal product of physical capital might be larger in an economy with a superior technology, even if that economy had more capital. The same is true of human capital. Countries can build up their human capital through schooling and training. They can also build up their human capital by attracting skilled workers from other countries. If richer countries are able to attract skilled workers, then we will see divergence rather than convergence. In Chapter 20, we discuss how economies actively seek to attract human capital.

Turning to technology, divergence in social infrastructure is certainly a possibility. Social infrastructure includes the rule of law, the general business climate, social attitudes toward corruption, the protection of property rights, and many other intangible factors. These influences on economic growth are difficult to define and almost impossible to measure accurately. Yet economists are

convinced that successful economies must have a good set of such social institutions. It is likely that it is easier to build and improve such institutions in countries that are relatively prosperous, which would again lead richer countries to grow more rapidly than poorer countries.

Economists have built some of these ideas into the theoretical framework of economic growth. Unfortunately, the models are too complicated for an introductory economics textbook, so we will not go into them in any detail here. We can, however, provide a simple example that conveys the flavor of these more complex ideas. The story goes as follows. We know that workers acquire human capital through education and on-the-job training. Suppose that, when there is more *physical* capital in the economy (relative to the number of workers), it is easier to acquire *human* capital. You can study in modern facilities with up-to-date computers. You work with state-of-the-art machinery and become more skilled. In this story, human capital is **endogenous**: it depends on the amount of physical capital.

To be concrete, imagine that technology is constant, and the amount of human capital is proportional to the amount of physical capital per worker. When we incorporate this assumption into the production function, we end up with a very straightforward relationship:

$$output \ = \ B \times physical \ capital,$$

where *B* is just a number.The derivation of this equation is not very difficult; it is explained in the toolkit.

In this economy, the ratio of capital stock to GDP is constant at all times (capital/output = 1/*B*). This economy is *always* on a balanced-growth path. Because of this, the growth rate of output equals the growth rate of capital stock:

$$output \ growth \ rate \ = \ physical \ capital \ growth \ rate.$$

The more important point, though, is that this technology *does not exhibit diminishing marginal product of capital*. The marginal product of capital is constant: it equals *B*. If this number were different in different economies, then we would expect to see capital stock flowing from economies where *B* is small toward economies where *B* is large. We would see divergence rather than convergence.

The model that we have described in this subsection is simplistic. Its point is simply to show that, if we make human capital endogenous, it is much easier to explain divergence. Economists have built more complicated and realistic models with endogenous human capital and technology that give similar results.

Key Takeaways

- Balanced growth occurs when capital stock grows at the same rate as output. Along a balanced-growth path, the ratio of output to capital stock does not change.
- Balanced growth is important to understand because over long periods of time, we expect economies to reach their balanced-growth path.
- There are reasons to expect at least some convergence in physical capital, human capital, and knowledge. However, there is no strong argument for why we would see convergence in social infrastructure.

Exercises

1. Suppose that an economy has a balanced-growth path where the physical capital stock is three times the level of GDP. If the current capital stock is four times the level of GDP, do you expect capital stock to grow faster or slower than GDP?
2. Suppose we have two economies that are currently identical, except in the first economy *a* is 0.3 and in the second economy *a* is 0.5. Will the balanced-growth path be the same in both countries? Which economy will converge more quickly to the balanced-growth path?

This page titled 21.5: Balanced Growth is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

21.6: The Role of International Institutions in Promoting Growth 21.6: The Role of International Institutions in Promoting GrowthEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What are the main international organizations that help to promote growth?
2. What do these institutions do to achieve their stated goals?

Governments acting alone can do a lot to promote economic growth. We have discussed the importance of protecting property rights and establishing a climate of political stability. These efforts by individual governments are complemented by international actions to promote growth and development in poorer countries. In this section, we describe three powerful and controversial international economic organizations: the World Bank, the International Monetary Fund (IMF), and the World Trade Organization (WTO). We briefly explain what these institutions do and how they go about reaching their goals.The websites of the World Bank, the IMF, and the WTO are, respectively, as follows: http://www.worldbank.org, http://www.imf.org/external/index.htm, and http://www.wto.org.

The World Bank

The World Bank is an international intermediary funded by 184 member countries. Its goal is to provide loans and grants to developing countries with the aim of eliminating poverty by promoting economic growth. Economists working at the World Bank rely on variants of the growth model used in this chapter to understand the growth experiences of different countries and determine the effects of policies in those countries.

The World Bank borrows money on international capital markets and also receives funds directly from member countries. The World Bank is similar to a bank that a household or a firm would approach for a loan to build a factory or a house, except that its borrowers are national governments. It often funds projects that would otherwise not be undertaken. In many cases, these are projects that promote infrastructure, education, health, and so forth. Projects like these may have social benefits yet not be profitable enough for private firms to undertake. Building a road in a rural part of a developing country is not the type of investment project one normally associates with a profit-seeking firm, for example, even though the road may have spurred rural development.

In 2010 the World Bank made about $45 billion in loan commitments and $29 billion in loan disbursements.“Annual Report 2010: Financial Information,” World Bank, accessed August 22, 2011, web.worldbank.org/WBSITE/EXTERNAL/EXTABOUTUS/EXTANNREP/EXTANNREP2010/0,,contentMDK:22626599~menuPK:7115719~pagePK:64168445~piPK:64168309~theSitePK:7074179,00.html. At one level, this is evidently substantial—a project worth $100 million or more can certainly have a large impact on a poor country. At another level, it is not a huge sum of money in the global economy. For comparative purposes, BP set aside over $40 billion to pay for the cleanup of its 2010 oil spill in the Gulf of Mexico.

1. World Bank projects range broadly. They include funding for infrastructure construction, promoting health care (such as HIV/AIDS programs), promoting education, and so forth. Many of these projects involve the provision of public goods, so they create benefits for society as a whole that exceed the direct return on investment. That is, many of the projects that are funded by national governments in richer countries are funded through the World Bank in developing countries. At the beginning of this chapter, we saw an example of a World Bank project in Niger, which was aimed at increasing human capital in that country. As another example, here is a description of a recent World Bank loan to Guyana to provide water access to the poor.

Guyana: Water Sector Consolidation Project

GRANT AMOUNT: $12.3 million

PROJECT DESCRIPTION: This project’s main objective is to increase access to safe water among the poor. The project seeks to support the achievement of sustainable universal access to safe and affordable water for the population of Guyana, especially the poor. The project will also help to consolidate the water sector modernization and reform process undertaken by the government with support of the International Development Association (IDA) and other donors in recent years.“Water Sector

Consolidation Project," World Bank, July 12, 2005, accessed June 30, 2011, web.worldbank.org/external/projects/main?pagePK=64283627&piPK=73230& theSitePK=40941&menuPK=228424&Projectid=P088030&cid=3001_72.

The project described here would not likely be a profitable private sector project, but it is important for the development of Guyana. Notice, too, that this loan, like many other World Bank loans, is for the development of infrastructure (roads, bridges, schools, communication systems, etc.). In more developed countries, such projects are usually performed by governments, but in developing countries, these investments are frequently undertaken through the World Bank.

Investment in infrastructure is typically complementary to the accumulation of other physical capital, such as machines and plants. Even though developing countries have relatively low capital stocks, investment in plants and equipment may not be very profitable if basic infrastructure is lacking. There is no point in building a factory if there are no roads to take your goods to market. Investment in infrastructure can increase the marginal product of capital and make other investment more attractive.

The International Monetary Fund

The IMF was established to (among other things) provide short-term support for countries facing financial difficulties. This is explicitly stated in the IMF's Articles of Agreement: "To give confidence to members by making the general resources of the Fund temporarily available to them under adequate safeguards, thus providing them with opportunity to correct maladjustments in their balance of payments without resorting to measures destructive of national or international prosperity.""Articles of Agreement of the International Monetary Fund," International Monetary Fund, February 22, 2010, accessed June 30, 2011, http://www.imf.org/external/pubs/ft/aa/aa01.htm.

A country's balance of payments has two main components. The first is the trade balance. A balance of payment maladjustment may mean that a country is running persistent trade deficits—that is, its imports are greater than its exports. This means the country is borrowing from other countries and is building up its external debt. The second component of the balance of payments is the interest that a country must pay on its *existing* external debt. This means that imbalances in the past lead to worse imbalances in the present. Imagine, for example, that Juan in Solovenia borrowed extensively in the past. It is then difficult for him to get out of debt because he has to pay so much interest. Moreover, the amount of external debt in a country cannot grow forever. When countries get into trouble by accumulating large amounts of debt, there is a temptation to default on outstanding debt. A key role of the IMF is to help countries through these difficult episodes.

IMF help has strings attached. A controversial aspect of the IMF's mode of operation is in the phrase…*under adequate safeguards*. As part of a deal to provide resources to countries in need of funds, the IMF often makes explicit demands about government fiscal and monetary policies. This is termed IMF "conditionality" and is described by the IMF as follows: "When a country borrows from the IMF, its government agrees to adjust its economic policies to overcome the problems that led it to seek financial aid from the international community. These loan conditions also serve to ensure that the country will be able to repay the Fund so that the resources can be made available to other members in need. In recent years, the IMF has streamlined conditionality in order to promote national ownership of strong and effective policies.""IMF Conditionality," International Monetary Fund, March 18, 2011, accessed August 22, 2011, http://www.imf.org/external/np/exr/facts/conditio.htm.

A quick tour of the IMF website (http://www.imf.org/external/index.htm) provides a lot of information about past and ongoing loans. One example is the ongoing relationship between the IMF and Argentina.The IMF formulates country reports on an annual basis, and these are available on the IMF website. These reports summarize the dealings between individual countries and the IMF. Argentina had reached an agreement with the IMF in September 2003 providing Argentina with access to SDR 8,981 million. SDR means "special drawing right." It is a unit of account used by the IMF whose value is an average of four key currencies. Its actual value on any given date can be found at www.imf.org/external/np/fin/data/rms_sdrv.aspx. In May 2011, 1 SDR was worth US$1.59. This agreement with Argentina came after Argentina was unable to meet demands for payment on some of its external debt and after real gross domestic product (real GDP) had fallen by nearly 11 percent in 2002. Agreement with the IMF was not immediate, partly due to the conditionality of a prospective loan. Though agreement was ultimately reached, there were lengthy negotiations regarding the conduct of fiscal and monetary policy in Argentina as a condition for IMF assistance.

The World Trade Organization

The WTO "makes the rules" for international trade. It is a relatively new organization—having been founded in 1995—and has 150 member countries. It arose from earlier trade agreements between countries, most notable the *General Agreement on Tariffs and Trade*. The WTO website describes the role of the organization as follows:

[…]

Essentially, the WTO is a place where member governments go, to try to sort out the trade problems they face with each other. The first step is to talk. The WTO was born out of negotiations, and everything the WTO does is the result of negotiations. The bulk of the WTO's current work comes from the 1986–94 negotiations called the Uruguay Round and earlier negotiations under the General Agreement on Tariffs and Trade (GATT). The WTO is currently the host to new negotiations, under the "Doha Development Agenda" launched in 2001.

Where countries have faced trade barriers and wanted them lowered, the negotiations have helped to liberalize trade. But the WTO is not just about liberalizing trade, and in some circumstances its rules support maintaining trade barriers—for example to protect consumers or prevent the spread of disease.

[…]"What Is the World Trade Organization?" World Trade Organization, accessed June 30, 2011, http://www.wto.org/english/thewto_e/whatis_e/tif_e/fact1_e.htm.

The negotiations at the WTO set the ground rules for international trade. Using the mechanisms of the WTO, countries agree on trade policies, such as the levels of tariffs. This is also a forum for designing policies on the protection of intellectual property rights. The WTO also provides a forum for dispute resolution.

Many critics of globalization have focused their attention on the WTO. For example, the nongovernmental organization Global Exchange (http://www.globalexchange.org) lists 12 "top reasons to oppose the WTO," including the claims that the WTO is increasing hunger, increasing inequality, trampling human rights, destroying the environment, and killing people through its policies. Critics such as this group argue that the WTO is fundamentally undemocratic, writing the rules so as to favor powerful corporations and rich countries. Defenders of the WTO argue that it gives poorer countries a much greater voice in international economic decision making. They point out, for example, WTO decisions are based on consensus, meaning that all 150 member countries must agree to them.

Key Takeaways

- The World Bank, the IMF, and the WTO are three leading international organizations that help countries in the development process.
- The World Bank funds projects in recipient countries, the IMF provides balance of payments support, and the WTO works to reduce trade barriers.

Exercises

1. In what way does the IMF work to promote convergence across countries?
2. The WTO helps to govern intellectual property rights. What is the impact of those rights on development?

This page titled 21.6: The Role of International Institutions in Promoting Growth is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

- **Current page** by Anonymous is licensed CC BY-NC-SA 3.0. Original source: https://2012books.lardbucket.org/books/theory-and-applications-of-economics.
- **23.3: Job and Worker Flows** by Anonymous is licensed CC BY-NC-SA 3.0. Original source: https://2012books.lardbucket.org/books/theory-and-applications-of-economics.

21.7: End-of-Chapter Material 21.7: End-of-Chapter MaterialEdit Update titlecancel

In Conclusion

We live in a world today that would be unrecognizable and unimaginable to those born two centuries ago. Things we take for granted—jet travel, antibiotics, electricity, the Internet, dentistry—are all products of the extraordinary growth of the last 200 years. Yet despite all our technological advances, billions of people in the world still live in poverty. Although some countries continue to grow rapidly, others stagnate or even go backward. If we could unlock the secrets of economic growth, we would have the means to help people to permanently better lives.

Even as economists emphasize economic growth as a way to combat poverty, noneconomists are often critical of economic growth, pointing out that it comes with costs as well as benefits. For example, as countries become richer, they use more energy and more of the world's natural resources. Oil reserves are being depleted, and rainforests are disappearing. Growth may lead to increased pollution, such as greenhouse gas emissions that in turn contribute to climate change. These are serious and legitimate concerns. In brief, economists have four main responses.

1. The framework we presented in this chapter does, in fact, capture the effect of declining natural resources. They lead to a slower rate of growth in technology. Indeed, it is possible that declining natural resources could more than offset growth in knowledge and social infrastructure so that the technology growth rate becomes negative. As yet, there is no evidence that this is a significant concern, but—at least until we have a better understanding of the drivers of knowledge and social infrastructure growth—it certainly might become relevant in the future.
2. There are indeed uncompensated side effects of economic growth, such as increased pollution. Economists agree that such effects can be very important. However, they can and should be corrected directly. Curtailing growth is an extremely indirect and inefficient response to its adverse side effects. As Nobel Prize–winner Robert Solow put it, "What no-growth would accomplish, it would do by cutting off your face to spite your nose."Robert M. Solow, "Is the End of the World at Hand?" *Challenge* 16, no. 1 (March/April 1973): 39–50. Also available at www.jstor.org/stable/40719094.
3. The evidence reveals that *some* environmental problems are solved rather than exacerbated by growth. Air pollution is a much more serious problem in the developing countries of the world than in the rich countries of the world. In part this is because a clean environment is a luxury good; people only worry about the state of the environment once their basic needs of food and shelter are addressed.
4. The most serious problems are those where we cannot rely on market mechanisms. If oil becomes scarce, then increases in the price of oil will provide incentives for people to economize on their use of fuel and look for alternative sources of energy. These incentives will at least ease the adjustment of the world economy. But there are no functioning market mechanisms to deal with climate change, for example.

Decades of research by economists have told us that there is no magic bullet, no simple and painless way to encourage economic growth. At the same time, we have learned a great deal about how and why countries grow. We have learned that growth depends on the accumulation of both physical and human capital. We have learned that growth ultimately hinges on the growth of knowledge, highlighting the importance of education, training, and research and development (R&D). And we have learned that good institutions are critical for countries that want to promote economic growth.

We have made progress, but the study of economic growth remains one of the most fascinating and challenging problems in all economics. There is no doubt that economists will continue their search for the elusive secrets of prosperity. As the Nobel Prize–winning economist Robert Lucas observed, "Once one starts to think about [economic growth], it is hard to think about anything else."

Key Links

- Penn World Tables: pwt.econ.upenn.edu
- International Monetary Fund (IMF): http://www.imf.org/external/index.htm
- World Bank: http://www.worldbank.org
  - World Bank Development Report: econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/EXTWDRS/0,,contentMDK:20227703~pagePK:478093~piPK:477627~theSitePK:477624,00.html

- World Bank Development Indicators: http://www.google.com/publicdata/overview?ds=d5bncppjof8f9_&ctype=l&strail=false&nselm=h&hl=en&dl=en
- World Trade Organization: http://www.wto.org
- CIA World Factbook: www.cia.gov/library/publications/the-world-factbook
- Angus Maddison's home page: http://www.ggdc.net/maddison
- Excel file of long-run data: http://www.ggdc.net/maddison/Historical_Statistics/horizontal-file_03-2007.xls

exercises

1. Think about your last visit to a shopping center or a large food store in the United States or other developed economy. Which of these goods and services do you think are available in a typical market in Niger? Which were available in the United States 50 years ago? 100 years ago?
2. (Advanced) In the late 1990s, the US government was running a surplus of about 1 percent of gross domestic product (GDP). Current projections show that the government is going to run deficits in excess of 5 percent of GDP in the future. Let us imagine that there are no changes in private saving or in foreign borrowing/lending.The condition that private savings do not change is important. For example, if the government cuts taxes, it is possible that people will predict that taxes will be higher in the future and will increase their savings in anticipation. We will say more about this in Chapter 29. In this case, the increased deficit translates directly into a decrease in the investment rate. To investigate the implications of such a decrease, suppose that, in the year 2000

   investment rate = 0.24,

   depreciation rate = 0.085,

   and output growth rate = 0.035.

   1. What would output per worker equal in 2040 if there had been no change in the investment rate?
   2. Suppose that by 2040, improvements in technology and human capital mean that the production function is given by output per worker = 30,000 × capital/output. Suppose also that the economy has reached its new balanced-growth path. What is output per worker in 2040?
   3. Now suppose that the increase in the government deficit means that the investment rate decreases to 0.18. What is the new balanced-growth ratio of the capital stock to GDP?
   4. Suppose the production function for this economy is output per worker = 15,000 × capital/output. What is output per worker in 2000?
   5. On a balanced-growth path, the ratio of capital stock to output is given by the following formula:

      Suppose that the economy was on a balanced-growth path in 2000. Calculate the balanced-growth ratio of the capital stock to GDP.

1. Try to estimate approximately how much you spend every day. Be sure to include an amount for rent, utilities, and food. Do you think it would be possible for you to live on $2 per day?
2. Suppose there are two economies. The first has a current level of real GDP of 100, and the second has a current level of real GDP of 200. The poorer country is forecasted to grow at 10 percent in the coming year, while the richer country is forecasted to grow at 15 percent. If these forecasts are true, what will their levels of real GDP be next year? Is this a case of divergence or convergence?
3. When capital's share of output (a) is larger, does an economy move to its balanced-growth path more quickly or more slowly? Explain.
4. Suppose that capital's share of output is 0.5, the human capital growth rate is 2 percent, the technology growth rate is 1 percent, and the workforce is not growing. What is the balanced-growth growth rate of output?
5. Look at Table 21.5.1 "Approaching the Balanced-Growth Path". Explain why the output growth rate decreases over time.
6. (Advanced) Think about Juan in Solovenia. Consider two cases. In the first case, he experiences an increase in his productivity that he knows will last for only one month. In the second, he experiences a permanent increase in his productivity. How do you think his decisions about how hard to work will be different in the two cases?
7. On a balanced-growth path, the ratio of capital stock to output is given by the following formula:

   Use the formula for the balanced-growth rate of output to determine how the ratio of capital stock to output depends on the growth rate of the workforce. Does an increase in the growth rate of the workforce lead to an increase or a decrease in the

ratio of capital stock to output?

Economics Detective

1. Find savings rates for the United States, India, and Niger and compare these to the investment rates for these countries. What can you say about capital inflows from other countries?
2. Go to the Penn World Tables (http://datacentre2.chass.utoronto.ca/pwt61). Click on “Alphabetical List of Countries.” Select the United States and two other countries of your choice. Look at the data for real GDP per capita and real GDP per worker. Briefly describe in words what has happened to these two variables over the period for which data are available.

Spreadsheet Exercises

1. Using a spreadsheet, reproduce Figure 21.7.1. Specifically, suppose that GDP starts with the value 10 in the year 2000, and capital stock in the same year has the value 20. Now set the growth rate of each series equal to 3 percent (0.03). What is the capital stock in 2050? What is GDP? Has the ratio of capital stock to GDP stayed constant?
2. Using the same spreadsheet and keeping the growth rate of GDP equal to 3 percent, examine what happens if the growth rate of capital is (a) 1 percent; (b) 5 percent.
3. Suppose that an economy has the following production function:

   Suppose that the workforce is growing at 1 percent per year, and human capital is growing at 2 percent per year. (We are assuming technology is constant in this example.) Suppose that we find that the ratio of capital stock to GDP is 4 on all dates and, initially, human capital is 15,000. What are the values for the growth rate of output per worker, the growth rate of output, and the growth rate of capital?
4. By experimenting with a spreadsheet, find out how long it will take for output per worker to double in this example.

This page titled 21.7: End-of-Chapter Material is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

22: The Great Depression

22.1: Lessons from History
22.2: What Happened during the Great Depression?
22.3: The Great Depression- A Decrease in Potential Output?
22.4: The Components of GDP during the Great Depression
22.5: The Great Depression- A Decrease in Aggregate Spending?
22.6: Policy Interventions and the Great Depression
22.7: End-of-Chapter Material

This page titled 22: The Great Depression is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

22.1: Lessons from History 22.1: Lessons from HistoryEdit Update titlecancel

Newspaper headlines around the world in 2008 asked whether the world's economies were heading for another "Great Depression." Long-past economic history suddenly captured the attention of economists, journalists, and others. But what was this event and why—even though it occurred the best part of a century ago—does it still hold such a prominent place in our economic memories?

In the early 1930s, instead of benefiting from economic growth and improved standards of living, people witnessed a huge decline in the level of economic activity. There was great economic hardship: large numbers of families struggled to obtain even basic food and shelter. Some sense of the desperation during these times can be found in oral histories. Here, for example, is one person's story of what it was like trying to find a job:

I'd get up at five in the morning and head for the waterfront. Outside the Spreckles Sugar Refinery, outside the gates, there would be a thousand men. You know dang well there's only three or four jobs. The guy would come out with two little Pinkerton cops: 'I need two guys for the bull gang. Two guys to go into the hole.' A thousand men would fight like a pack of Alaskan dogs to get through there. Only four of us would get through. I was too young a punk.See Studs Terkel, *Hard Times: An Oral History of the Great Depression* (New York: Pantheon Books, 1970), 30.

The personal suffering is less apparent in the figure below, but this picture does reveal the extraordinary nature of those times. It shows real gross domestic product (real GDP) in the United States from 1890 to 1939. Three things stand out. First, the level of economic activity grew substantially during this half century. This is normal: economies typically grow over the long haul, becoming more productive and producing more output. Second, although the level of US economic activity grew substantially over this half century, there were many ups and downs in the economy during the late 19th century and early 20th century. Third—and most important for our purposes—the period from 1929 to 1937 stands out from the rest. This was not a minor blip in economic activity; the US economy suffered a collapse that persisted for many years. At the same time, unemployment climbed to a staggering 25 percent in 1933—one out of four people was unemployed—compared to a rate of only 3.2 percent in 1929.

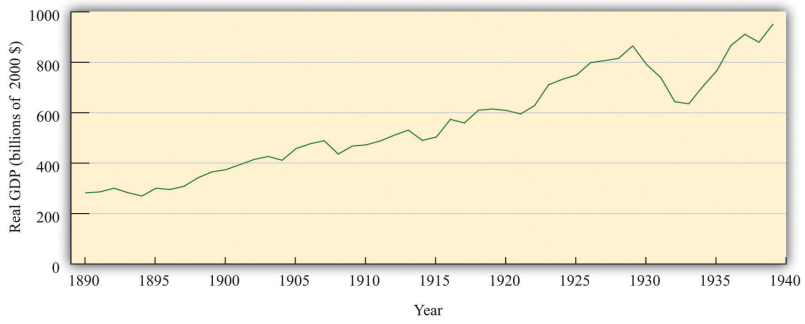


Figure 22.1.1: US Real GDP, 1890–1939. Real GDP increased considerably between 1890 and 1939, but the Great Depression of the early 1930s is a striking exception.

Source: Data from "What Was the U.S. GDP Then?," Measuring Worth, accessed August 22, 2011, www.measuringworth.org/datasets/usgdp/result.php.

The United States was not the only country to experience such hard economic times in this period. Many other countries, such as the United Kingdom, Canada, France, Germany, and Italy also saw their economic progress reversed for a period of years. The Great Depression, as this economic cataclysm came to be called, was a shock to the economists of the day. Prior to that time, most economists thought that, though economies might grow fast in some years and decline slightly in others, prolonged unemployment and underutilization of resources was impossible. The Great Depression proved this view to be erroneous and eventually led to a fundamental change in the way in which economists thought about the aggregate economy. The idea that the economy was naturally stable was replaced with a view that severe economic downturns could recur at any time.

Along with this change in thinking about the economy came a change in attitudes toward macroeconomic policy: economists began to believe that the government could play an active role to help stabilize the economy, perhaps by increasing government spending in bad times. Prior to the Great Depression, nobody even thought that the government should try to keep the economy stable. Both Democrats and Republicans in the 1932 election advocated *less* government spending because government revenues had fallen. Yet, by the end of the 1930s, the United States and other countries had adopted the view that active policy measures were useful or even essential for the proper functioning of economies.

Three-fourths of a century later, these events are part of economic history. Few people still alive experienced those terrible years directly, yet the time remains part of our collective memory. Above all, we need to know what went wrong if we hope to ensure that such punishing times do not come again. Indeed, the world economy recently suffered the most severe recession since the 1930s, and it is unclear at the time of this writing how long or how bad the current crisis will be. The insights of the economists who explained the Great Depression are still at the heart of today's discussions of economic policy. Understanding what happened to the economy in the 1930s is more than an exercise in economic history; it is essential for understanding modern macroeconomics. We want to know—

What caused the Great Depression?

Road Map

We begin by looking at some facts about the Great Depression and the boom that preceded it. Our goal is to see if we can develop a good explanation of these facts. The most fundamental defining feature of the Great Depression was the large and sustained decrease in real GDP. In the figure below, which shows the circular flow of income, reminds us that real GDP measures both production and spending.

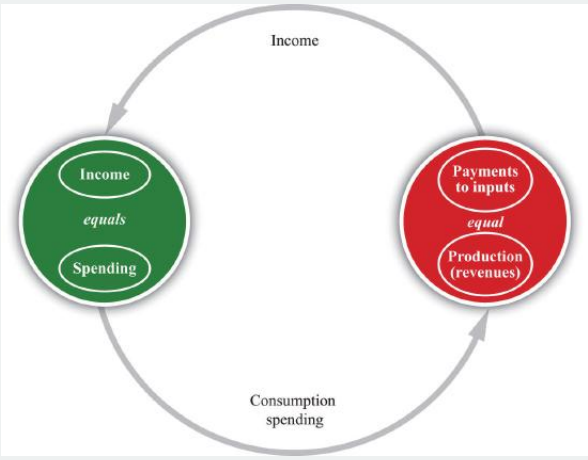


Figure 22.1.2: The Circular Flow of Income. GDP measures the production of an economy *and total income in an economy. We can use the terms production, income, spending, and GDP interchangeably.*

It follows that during the Great Depression, both output and spending decreased. Perhaps it is the case that production in the economy declined for some reason, and spending decreased as a consequence. Or perhaps spending declined for some reason, and production decreased as a consequence. We examine two approaches to the Great Depression, based on these ideas. One sees the root cause of the Great Depression as a decline in the productive capabilities of the economy, meaning that firms—for some reason—were unable to produce as much as they had before. This then led to decreased spending. The other approach sees the root cause of the Great Depression as a decline in spending, meaning that households and firms—for some reason—decided that they wanted to purchase fewer goods and services. This then led to decreased production.

We look at each explanation in turn. We investigate which inputs contributed the most to the decrease in output and also look at what happened to the different components of spending. This more careful look at the data helps us to evaluate the two competing theories of the Great Depression. We conclude by examining the implications for economic policy and considering what policies were actually conducted at the time of the Great Depression.

This page titled 22.1: Lessons from History is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

22.2: What Happened during the Great Depression? 22.2: What Happened during the Great Depression?Edit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What are the main facts about the Great Depression?
2. What is puzzling about the Great Depression?
3. What are the two leading strands of thought about the cause of the Great Depression?

We begin with some facts. Table 22.2.1 "Major Macroeconomic Variables, 1920–39*" shows **real gross domestic product (real GDP)**, the **unemployment rate**, the **price level**, and the **inflation rate** from 1920 to 1939 in the United States. Real GDP measures the overall production of the economy, the unemployment rate measures the fraction of the labor force unable to find a job, the price level measures the overall cost of GDP, and the inflation rate is the growth rate of the price level.

Table 22.2.1: Major Macroeconomic Variables, 1920–39*

Year	Real GDP	Unemployment	Price Level	Inflation Rate
1920	606.6	5.2	11.6	
1921	585.7	11.7	10.4	−10.3
1922	625.9	6.7	9.8	−5.8
1923	713.0	2.4	9.9	1.0
1924	732.8	5.0	9.9	0.0
1925	748.6	3.2	10.2	3.0
1926	793.9	1.8	10.3	1.0
1927	798.4	3.3	10.4	1.0
1928	812.6	4.2	9.9	−4.8
1929	865.2	3.2	9.9	0.0
1930	790.7	8.9	9.7	−2.0
1931	739.9	16.3	8.8	−9.3
1932	643.7	24.1	8.0	−9.1
1933	635.5	25.2	7.5	−6.3
1934	704.2	22.0	7.8	4.0

*GDP is in billions of year 2000 dollars (Bureau of Economic Analysis [BEA]). The unemployment rate is from the US Census Bureau, *The Statistical History of the United States: From Colonial Times to the Present* (New York: Basic Books, 1976; see also www.census.gov/prod/www/abs/statab.html). The base year for the price index is 2000 (that is, the index equals 100 in that year) and comes from the Bureau of Labor Statistics (BLS; http://www.bls.gov), 2004.

Year	Real GDP	Unemployment	Price Level	Inflation Rate
1935	766.9	20.3	8.0	2.6
1936	866.6	17.0	8.1	1.3
1937	911.1	14.3	8.4	3.7
1938	879.7	19.1	8.2	−2.4
1939	950.7	17.2	8.1	−1.2

*GDP is in billions of year 2000 dollars (Bureau of Economic Analysis [BEA]). The unemployment rate is from the US Census Bureau, *The Statistical History of the United States: From Colonial Times to the Present* (New York: Basic Books, 1976; see also www.census.gov/prod/www/abs/statab.html). The base year for the price index is 2000 (that is, the index equals 100 in that year) and comes from the Bureau of Labor Statistics (BLS; http://www.bls.gov), 2004.

Looking at these data, we see first that the 1920s were a period of sustained growth, sometimes known as the "roaring twenties." Real GDP increased each year between 1921 and 1929, with an average growth rate of 4.9 percent per year). Meanwhile the unemployment rate decreased from 6.7 percent in 1922 to 1.8 percent in 1926. Real GDP reached a peak of $865 billion in 1929. This number is expressed in year 2000 dollars, so we can compare that number easily with current economic data. In particular, if we divide by the population at that time, we find that GDP per person was the equivalent of about $7,000, in year 2000 terms. Real GDP per person has increased about fivefold since that time.

Toolkit: Section 31.21

You can review growth rates in the toolkit.

The Great Depression began in late 1929 as a recession not unlike those experienced previously—a decrease in GDP from one year to the next was common—but it rapidly blossomed into a four-year reduction in economic activity. By 1933, real GDP had fallen by over 25 percent and was only $636 billion. At the same time, unemployment increased from around 3 percent to 25 percent. In 1929, jobs were easy to come by. By 1933, they were almost impossible to find. More than a quarter of the people wishing to work were unable to find a job. Countless others, no doubt, had given up even looking for a job and were out of the labor force.

The experience of the 1920s and 1930s tells us that when real GDP increases, unemployment tends to decline and vice versa. We say that unemployment is **countercyclical**, meaning that it typically moves in the direction opposite to the movement of real GDP. An economic variable is **procyclical** if it typically moves in the same direction as real GDP, increasing when GDP increases and decreasing when GDP decreases. The countercyclical behavior of unemployment is not something that is peculiar to the Great Depression; it is a relatively robust fact about most economies. It is also quite intuitive: if fewer people are employed, less labor goes into the production function, so we expect output to be lower.

An event occurred in September 1929 that, at least with hindsight, marks a turning point. The stock market, as measured by the Dow Jones Industrial Average, had been increasing until that time but then decreased by 48 percent in less than 2.5 months. The value of the stock market is a measure of the value, in the minds of investors, of all the firms in the economy. Investors suddenly decided that the US economy was worth only half what they had believed three months earlier. It is unlikely that two such dramatic economic events occurred at almost the same time and yet are unconnected. We should not make the claim that the stock market crash *caused* the Great Depression. But the stock market decrease was *correlated* with declining output in the early days of the Great Depression. **Correlation** is distinct from causation. It is possible, for example, that the stock market crash and the Great Depression were both caused by some other event.

Toolkit: Section 31.23

Correlation is a statistical measure of how closely two variables are related. If the two variables tend to increase together, we say that they are "positively correlated"; if one increases when the other decreases, then they are "negatively correlated." If the relationship between the two variables is an exact straight line, we say that they are "perfectly correlated." The fact that two variables are correlated does not necessarily mean that changes in one variable *cause* changes in the other. The toolkit contains more information.

Table 22.2.1 "Major Macroeconomic Variables, 1920–39*" also contains information on the price level and the inflation rate. The most striking fact from this table is that the price level declined over this period—on average, goods were considerably cheaper in dollar terms in 1940 than they were in 1920. We see this both from the decrease in the price level and from the fact that the inflation rate was *negative* in several years (remember that the inflation rate is the growth rate of the price level). If we look at the more recent history of the United States and at most other countries, we rarely observe negative inflation. Decreasing prices are an unusual phenomenon.

Other countries had similar experiences during this time period. Figure 22.2.1 shows that France, Germany, and Britain all experienced very poor economic performance in the early 1930s. Output was lower in each country in 1933 compared to four years earlier, and each country also saw a decline in the price level. Many other countries around the world had similar experiences. The Great Depression was a worldwide event.

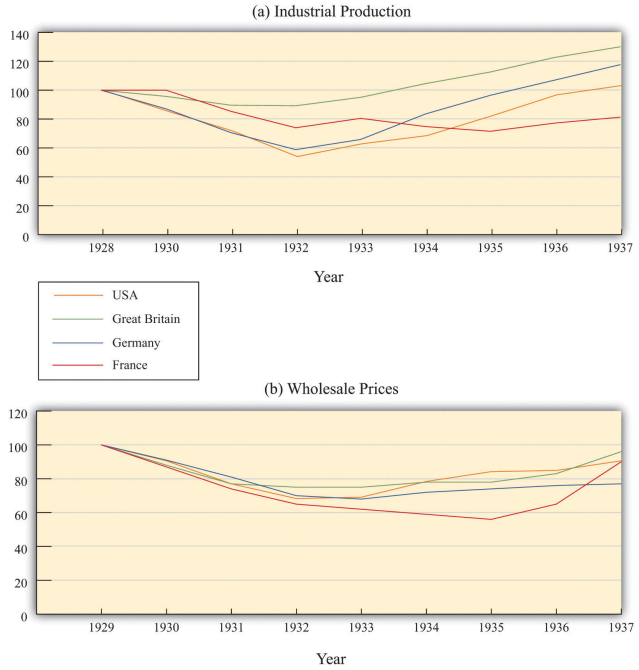


Figure 22.2.1: The Great Depression in Other Countries. France, Germany, and Britain also experienced declines in output (a) and prices (b) during the Great Depression. The output data are data for industrial production (manufacturing in the case of the United States), and the price data are wholesale prices.

Source: International Monetary Fund, "World Economic Outlook: Crisis and Recovery," April 2009, Box 3.1.1, http://www.imf.org/external/pubs/ft/weo/2009/01/c3/Box3_1_1.pdf.

Why this was the case remains one of the puzzles of the period. There were events at the time that had international dimensions, such as concerns about the future of the "gold standard" (which determined the **exchange rates** between countries) and various policies that disrupted international trade. Still, economists are unconvinced that such factors can explain why the Great Depression occurred in so many countries. Three-fourths of a century later, we still do not have a complete understanding of the Great Depression and are still unsure exactly why it happened. From one perspective this is frustrating, but from another it is exciting: the Great Depression maintains an air of mystery.

Toolkit: Section 31.20

You can review the meaning and definition of the exchange rate in the toolkit.

The Puzzle of the Great Depression

Try to imagine yourself in the United States or Europe in the early 1930s. You are witnessing immense human misery amid a near meltdown of the economy. Friends and family are losing their jobs and have bleak prospects for new employment. Stores that you had shopped in all your life suddenly go out of business. The bank holding your money has disappeared, taking your savings with it. The government provides no insurance for unemployment, and there is no system of social security to provide support for your elderly relatives.

Economists and government officials at that time were bewildered. The experience in the United States and other countries was difficult to understand. According to the economic theories of the day, it simply was not possible. Policymakers had no idea how to bring about economic recovery. Yet, as you might imagine, there was considerable pressure for the government to do something about the problem. The questions that vexed the policymakers of the day—questions such as “What is happening?” and “What can the government do to help?”—are at the heart of this chapter.

Economists make sense of events like the Great Depression by first accumulating facts and then using frameworks to interpret those facts. We have a considerable advantage relative to economists and politicians at the time. We have the benefit of hindsight: the data we looked at in the previous subsection were not known to the economists of that era. And economic theory has evolved over the last seven decades, giving us better frameworks for analyzing these data.

Earlier, we said there are two possible reasons why output decreased.

1. There was a decrease in production due to a decrease in the available inputs into the **aggregate production function**. Since there was no massive decrease in the amount of **physical capital** or the size of the workforce, and people presumably did not suddenly lose all their **human capital**, this means that the culprit must have been a decrease in **technology**.
2. There was a decrease in aggregate spending. Households chose to reduce their consumption, firms chose to reduce their investment, and governments chose to reduce their spending. As a consequence, firms scaled back their production.

We look at each of these candidate explanations in turn.

Toolkit: Section 31.26

You can review the aggregate production function and the inputs that go into it in the toolkit.

Key Takeaways

- During the Great Depression in the United States from 1929 to 1933, real GDP decreased by over 25 percent, the unemployment rate reached 25 percent, and prices decreased by over 9 percent in both 1931 and 1932 and by nearly 25 percent over the entire period.
- The Great Depression remains a puzzle today. Both the source of this large economic downturn and why it lasted for so long remain active areas of research and debate within economics.
- One explanation of the Great Depression rests on a reduction in the ability of the economy to produce goods and services. The second leading explanation focuses on a reduction in the overall demand for goods and services in the economy.

Exercises

1. The notes in Table 22.2.1 "Major Macroeconomic Variables, 1920–39*" state that the base year for the price level is 2000, so the price index has a value of 100 in that year. Approximately how much would you expect to have paid in the year 2000 for something that cost $2 in the late 1920s?
2. Using Table 22.2.1 "Major Macroeconomic Variables, 1920–39*", how can you see that the unemployment rate is countercyclical?

This page titled 22.2: What Happened during the Great Depression? is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

22.3: The Great Depression- A Decrease in Potential Output? 22.3: The Great Depression- A Decrease in Potential Output?Edit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What is potential output?
2. How could a decrease in potential output create the Great Depression?
3. How does the theory that the Great Depression was caused by a decrease in potential output match the facts?

Our first approach to interpreting the Great Depression focuses on **potential output**, which is the amount of real gross domestic product (real GDP) an economy produces when the labor market is in equilibrium and capital goods are not lying idle. We start here because this approach corresponds reasonably closely to the economic wisdom of the time.

A Decrease in Technology: The Multiple-Markets Perspective

Comparative statics is a technique that allows us to understand the effects of a decrease in technology in a particular market, such as the market for new homes. In a comparative statics exercise, we look at what happens to **endogenous variables** (in this case, production and prices of new homes) when we change an **exogenous variable** (in this case, technology). A decline in technology shifts the market supply curve leftward: at any given price, the decrease in technology means that the firm can produce less output with its available inputs. The result is shown in part (a) of Figure 22.3.1 for the housing market: output of new homes decreases and the price of new homes increases.

Toolkit: Section 31.16

You can review the technique of comparative statics and the definition of endogenous and exogenous variables in the toolkit.

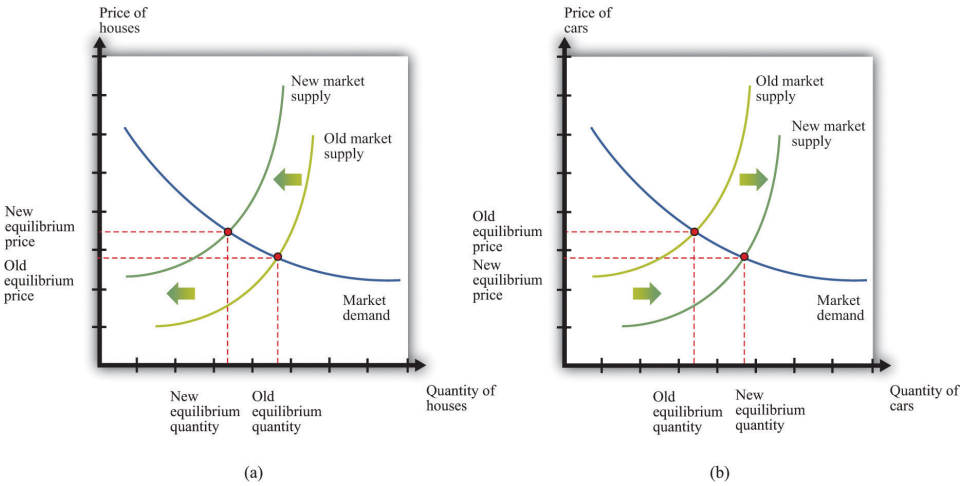


Figure 22.3.1: An Inward Shift in the Market Supply of Houses. (a) A decrease in technology leads to an inward shift of the market supply curve for houses. (b) The labor and other resources that are not being used to produce houses can now be used to produce other goods, such as cars.

If this decline in technology in the housing market were the only change in the economy, what would happen? Construction firms would fire workers because these firms were building fewer new homes. Over time, however, the fired construction workers would find new jobs in other sectors of the economy. The same logic applies to other inputs: capital and other inputs that were being used in the construction industry would be redeployed to other markets. For example, there would be additional labor and other inputs available for automobile production. Part (b) of Figure 22.3.1 shows the resulting outward shift in the supply curve for cars. It is difficult to explain the big decrease in output and the high rate of unemployment in the Great Depression through a change in technology in a single market.

Suppose, however, that this change in technology does not happen in just one market but occurs across the entire economy. Then a version of part (a) in Figure 22.3.1 would hold for each market in the economy. We would see declines in economic activity across a wide range of markets. Moreover, with declines in so many industries, we would expect to see lower real wages and less employment. The idea that workers could easily move from one industry to another is not as persuasive if the entire economy is hit by an adverse technology shock.

Using Growth Accounting to Understand the Great Depression

We use **growth accounting** to show how changes in output are driven by changes in the underlying inputs—capital, labor, and technology. Equivalently, we use the technique to give us a measure of the growth rate of technology, given data on the growth rates of output, capital, and labor:

$$technology\ growth\ rate\ =\ output\ growth\ rate\ -[a \times capital\ stock\ growth\ rate]-[(1-a) \times labor\ growth\ rate$$

.

We have omitted human capital from this growth accounting equation. We do so because, unfortunately, we do not have very good human capital measures for the period of the Great Depression. Human capital typically changes very slowly, so this is not too much of a problem: over a period of a decade, we do not expect big changes in human capital. Any changes in human capital that do occur are included in the catchall “technology” term.

Toolkit: Section 31.28

You can review the technique of growth accounting in the toolkit.

The key ingredient needed for the growth accounting equation is the number *a*. It turns out that a good measure of *a* is the fraction of real GDP that is paid to owners of capital. Roughly speaking, it is the amount of GDP that goes to the profits of firms. Equivalently, (1 − *a*) is the fraction of GDP that is paid to labor. The circular flow of income reminds us that all income ultimately finds its way back to households in the economy, which is why these two numbers sum to one.

Can Technology Changes Explain the Roaring Twenties?

The economist John Kendrick applied such growth accounting to data from the Great Depression.See John W. Kendrick, *Productivity Trends in the United States* (Princeton, NJ: Princeton University Press, 1961), particularly Table A-XXII, p. 335, and the discussion of these calculations. Table 22.3.1 "Growth Rates of Real GDP, Labor, Capital, and Technology, 1920–39*" summarizes his findings. Each row in Table 22.3.1 "Growth Rates of Real GDP, Labor, Capital, and Technology, 1920–39*" decomposes output growth into three components. In 1923, for example, output grew at a very high rate of 14.2 percent. This growth in output came from labor growth of 9.9 percent and capital stock growth of 2.0 percent. The remainder, which we interpret as growth in technology, grew at 9.5 percent. By all accounts, 1923 was a good year. The other entries in the table can be read in the same way.

Table 22.3.1: Growth Rates of Real GDP, Labor, Capital, and Technology, 1920–39*

Year	Real GDP	Labor	Capital	Technology
1920	0.4	1.4	2.1	−1.2
1921	−3.6	−11.5	1.5	4.0
1922	6.4	8.7	0.7	0.1
1923	14.2	9.9	2.0	9.5
1924	2.0	−3.2	2.6	4.9
1925	3.6	4.0	2.4	0.1
1926	6.2	4.2	3.2	3.4

*All entries are annual growth rates calculated using data from John W. Kendrick, *Productivity Trends in the United States* (Princeton, NJ: Princeton University Press, 1961), Table A-XXII, 335. Following the discussion in Kendrick, the capital share (*a*) was 0.30 until 1928 and 0.25 thereafter.

Year	Real GDP	Labor	Capital	Technology
1927	1.1	−0.2	2.9	0.5
1928	1.0	0.6	2.4	−0.3
1929	6.5	2.2	2.4	5.7
1930	−9.2	−8.1	2.0	−4.8
1931	−7.5	−10.5	0.1	0.4
1932	−14.5	−13.5	−2.2	−5.2
1933	−2.5	−1.0	−3.4	−1.2
1934	9.9	0.4	−2.8	13.7
1935	9.0	5.8	−1.4	6.6
1936	12.8	10.3	0.0	6.8
1937	6.9	5.8	1.4	2.9
1938	−5.5	−9.3	0.9	1.2
1939	9.1	6.2	−0.3	4.6

*All entries are annual growth rates calculated using data from John W. Kendrick, *Productivity Trends in the United States* (Princeton, NJ: Princeton University Press, 1961), Table A-XXII, 335. Following the discussion in Kendrick, the capital share (a) was 0.30 until 1928 and 0.25 thereafter.

Real GDP and technology were both growing in most years in the 1920s. In the early 1930s both variables decreased, and both grew again as the economy recovered from the Great Depression. In other words, technology growth and output growth are positively correlated over this period. This suggests the possibility that changes in technology *caused* the changes in output—always remembering that, as we observed earlier, correlation need not imply a causal relationship. An improvement in technology causes firms to want to produce more. They demand more workers, so employment and real wages increase. The increased output, through the circular flow, means that there is increased income. Households increase both consumption and savings. Higher savings means higher investment, so, over time, the economy accumulates more capital. Exactly the opposite holds if there is a decrease in technology: in this case, employment, consumption, and investment all decrease.

Does this theory fit the facts? For the roaring twenties, we see growth in output, labor, and capital. In addition, there was a positive technology growth rate in almost all the years of the decade. These movements are indeed consistent with the behavior of an economy driven by improvements in technology. Jumping back for a moment to individual markets, improvements in technology shift supply curves rightward. Increased output is therefore accompanied by decreased prices. The aggregate price level is nothing more than a weighted average of individual prices, so price decreases in individual markets translate into a decrease in the overall price level. From Table 22.2.1 "Major Macroeconomic Variables, 1920–39*", the price level actually moved very little between 1922 and 1929, so this fits less well.

Overall, the view that technological progress fueled the growth from 1922 to 1929 seems broadly consistent with the facts. Given the simplicity of the framework that we are using, “broadly consistent” is probably the best it is reasonable to hope for.

Can Technology Changes Explain the Great Depression?

Now let us apply the same logic to the period of the Great Depression. Negative growth in output from 1930 to 1933 was matched by negative growth in labor and technology (except for 1931). The capital stock decreased from 1932 to 1935, reflecting meager investment during this period. When the economy turned around in 1934, technology growth turned up as well.

Imagine that the economy experienced negative technology growth from 1929 to 1933. The reduced productivity of firms leads to a decrease in demand for labor, so real wages and employment decrease. Lower productivity also means that firms did not think it was worthwhile to invest in building new factories and buying new machinery. Both labor and capital inputs into the production function declined. Once technology growth resumed in 1934, the story was reversed: labor and capital inputs increased, and the

economy began to grow again. In this view, there was a substantial decline in the production capabilities of the economy, leading to negative growth in output, consumption, and investment. The Great Depression, in this account, was driven by technological regress.

Many economists are skeptical of such an explanation of the Great Depression. They have three criticisms. First, large-scale technological regress is difficult to believe on its face. Did people know an efficient way to manufacture something in 1929 but then forget it in 1930? Even remembering that technology includes social infrastructure, it is hard to imagine any event that would cause a decrease of 3 percent or more in technology—and if such an event did occur, surely we would be able to point to it and identify it.

Second, this explanation claims that labor input decreased because households saw lower real wages and voluntarily chose to consume leisure rather than work. By most measures, though, real wages *increased*. Moreover, it is difficult to equate a 25 percent unemployment rate, not to mention all the stories of how people could not find work, with a labor market in which households are simply moving along a labor supply curve.

Third, a prominent feature of the Great Depression is the decrease in the price level that occurred from 1929 to 1933. Table 22.2.1 "Major Macroeconomic Variables, 1920–39*" tells us that prices decreased by over 9 percent in both 1931 and 1932. However, a reduction in the level of potential GDP would cause an inward shift of market supply curves and thus an *increase*, rather than a decrease, in prices.

For most economists, the view of the Great Depression as a shift in technology is not convincing. Something else must have been going on. In particular, the very high unemployment rate strongly suggests that labor markets were malfunctioning. Thus, rather than viewing the large decreases in output in economies around the world as part of the normal functioning of supply and demand in an economy, we should perhaps consider it as evidence that sometimes things can go badly wrong with the economy's self-correction mechanisms. If we want to explain the Great Depression, we are then obliged—as were the economists at the time—to find a new way of thinking about the economy. It was an economist named John Maynard Keynes who provided such a new approach; in so doing, he gave his name to an entire branch of macroeconomic theory.

Key Takeaways

- Potential output is the amount of real GDP an economy could produce if the labor market is in equilibrium and capital goods are fully utilized.
- A large enough decrease in potential output, say through technological regress, could cause the large decrease in real GDP that occurred during the Great Depression.
- A reduction in potential output would lead to a decrease in real wages and an increase in the price level. Those implications are inconsistent with the facts of the Great Depression years. Further, it is hard to understand how potential output could decrease by the extent needed to match the decrease in real GDP during the Great Depression. Finally, a 25 percent unemployment rate is not consistent with labor market equilibrium.

Exercises

1. Draw the comparative statics picture for the labor market for the situation in which the Great Depression is a consequence of technological regress—that is, negative technology growth. Which curve shifts? Does it shift leftward or rightward?
2. Suppose the supply curve in a market shifts rightward. What must happen to the demand curve if the price in the market does not change?

This page titled 22.3: The Great Depression- A Decrease in Potential Output? is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

22.4: The Components of GDP during the Great Depression 22.4: The Components of GDP during the Great DepressionEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What are the main components of aggregate spending?
2. What is the national income identity?
3. What happened to consumption and investment spending during the Great Depression?
4. What is consumption smoothing?

In his analysis of the Great Depression, John Maynard Keynes contrasted his new approach with the prevailing "classical" theory:John Maynard Keynes, *The General Theory of Employment, Interest and Money* (Orlando: First Harvest/Harcourt, 1964[1936]), 3. "I shall argue that the postulates of the classical theory are applicable to a special case only and not to the general case.…Moreover, the characteristics of the special case assumed by the classical theory happen not to be those of the economic society in which we actually live, with the result that its teaching is misleading and disastrous if we attempt to apply it to the facts of experience." Keynes claimed that there was a fundamental failure in the economic system that prevented markets from fully coordinating activities in the economy. He argued that, as a consequence, the actual output of the economy was not determined by the productive capacity of the economy, and that it was "misleading and disastrous" to think otherwise. In more modern terms, he said that actual output need not always equal potential output but was instead determined by the overall level of spending or demand in the economy.

Keynes provided a competing story of the Great Depression that did not rely on technological regress and in which unemployment truly reflected an inability of households to find work. Keynes gave life to **aggregate spending**—the total spending by households, firms, and governments—as a determinant of aggregate gross domestic product (GDP). With this new perspective, Keynes also uncovered a way in which government intervention might help the functioning of the economy.

To understand how Keynes approached the puzzle of the Great Depression, we must first look more closely at the components of GDP. Figure 22.4.1 shows the circular flow, emphasizing the flows in and out of the firm sector of the economy. Accounting rules tell us that in every sector of the circular flow, the flow of dollars in must equal the flow of dollars out. We know that the total flow of dollars from the firm sector measures the total value of production in the economy. The total flow of dollars into the firm sector equals total expenditures on GDP. The figure therefore illustrates a fundamental relationship in the national accounts.

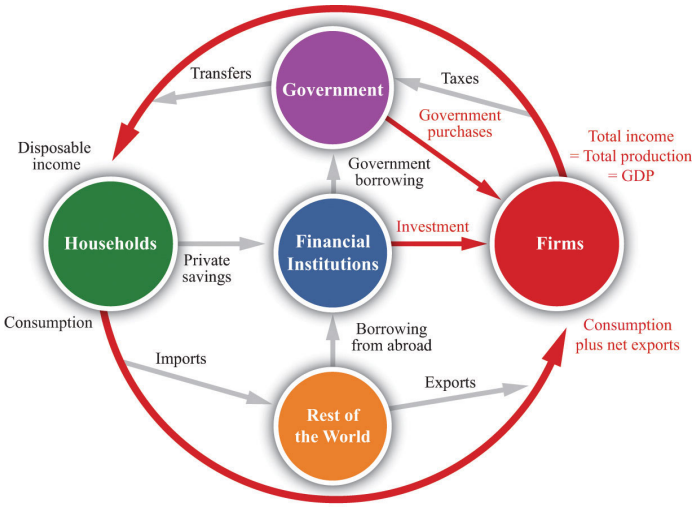


Figure 22.4.1: The Firm Sector in the Circular Flow. The flow of dollars into the firm sector equals consumption plus net exports plus investment plus government purchases. The flow of dollars from the firm sector equals total GDP in the economy.

The National Income Identity

The **national income identity** states that

$$production = consumption + investment + government\ purchases + net\ exports.$$

$production = consumption + investment + government\ purchases + net\ exports.$

Toolkit: Section 31.27

The toolkit describes the circular flow of income in more detail.

Consumption refers to total consumption spending by households on final goods and services. Consumption is divided into three categories.

1. **Services**. These are items such as haircuts, restaurant meals, hotel nights, legal services, and movies. There is often no tangible product; the consumer purchases the time and skills of individuals (such as barbers, chefs, and lawyers). Production and consumption of services usually occur together.
2. **Nondurable goods**. Examples include groceries, clothing, and DVDs—tangible products that (usually) have a fairly limited lifespan (typically less than three years).
3. **Durable goods**. These are items such as automobiles, "white goods" (washing machines, refrigerators, and other appliances), and computers. They are tangible products that usually have a lifespan of several years.

The distinctions among these categories are not always as clear-cut as the definitions suggest. A good pair of blue jeans might outlast a shoddy dishwasher, even though the jeans are classified as a nondurable good and the dishwasher as a durable good.

Investment is the purchase of new goods that increase the capital stock, allowing us to produce more output in the future. Investment is divided into three categories.

1. **Business fixed investment**. Purchases of physical capital (plants, machines) for the production of goods and services
2. **New residential construction**. The building of new homes
3. **Inventory investment**. Change in inventories of final goods

The economist's definition of investment is precise and differs from the way we often use the word in everyday speech. Specifically, economists do *not* use the term to mean the purchase of financial assets, such as stocks and bonds. Most of the time when we talk about investment in this book, we are referring to business fixed investment—the production of new physical capital goods. Inventory investment is a special category of investment that we explain in 22.3 Section "Inventory Investment".

As a rough rule of thumb, consumption spending is carried out by households, and investment spending is carried out by firms. But there is one important exception: new residential construction is included in investment. A new house purchased by a household is treated as investment, not consumption.

Government purchases include all purchases of goods and services by the government. We include in our definition of "government" local as well as national government activity. In the United States, this means that we collapse together federal, state, and local governments for the purpose of our analysis.

This component of spending refers only to *purchases* of goods and services, not to **transfers**. So, if the federal government buys aircraft from Boeing or the local police department buys a fleet of Volvos, these are included in government purchases. However, a transfer you receive from the government—say, because you are unemployed and are being paid unemployment insurance—is not counted in GDP. (Of course, if you then use this income to purchase goods and services, that consumption is part of GDP.)

Net exports simply equal exports minus imports. They are included because we must correct for the expenditure flows associated with the rest of the world. Some spending in the economy goes to imported goods, which is not associated with domestic production. We must subtract these imports from total expenditures. Against that, some demand for domestically produced goods comes from other countries. We add these exports to total expenditure.

Inventory Investment

Inventory investment is a relatively minor component of GDP, but we need to understand it in some detail because it plays a key role in the Keynesian approach. When a firm produces output, it does one of two things with it: it either sells it or adds it to inventory. Thus an accounting relationship within a firm is that

$$production = sales + changes\ in\ inventory.$$

If a firm produces more than it sells, its stocks of inventories increase. If a firm sells more than it produces, its stocks of inventories decrease. The inventories that a firm holds are counted as part of its capital stock, so any change in firms' inventories is counted as a component of investment.

Suppose General Motors (GM) produces 10 million cars, anticipating that it will sell them all. Then imagine that demand is lower than expected, so it only sells 9.9 million. The result is that 100,000 cars pile up on GM's lots, and the GM accountants record this as an addition to inventory. We want GDP to measure both production and spending, but we have 100,000 cars that have been produced but not purchased. The national income accounts get around this problem by effectively pretending that GM bought the cars from itself.

If the cars are then sold in the following year, they will not contribute to GDP in that year—quite properly, since they were not produced that year. The national accounts in the next year will show that 100,000 cars were sold to households, but they will also show that inventories decreased by 100,000 cars. Thus the accounts record expenditures on these cars as part of durable goods consumption, but the accounts also contain an offsetting reduction in inventory investment.

In some cases, firms change their stocks of inventory as a part of their business strategy. More often, changes in inventories occur because a firm did not correctly forecast its sales. **Unplanned inventory investment** is an increase in inventories that comes about because a firm sells less than it anticipated. Because GM *expected* to sell all 10 million cars but sold only 9.9 million, GM had 100,000 cars of unplanned inventory investment.

Moreover, GM is likely to react swiftly to this imbalance between its production plans and its sales. When it sees its sales decrease and its inventory increase, it will respond by cutting its production back until it is in line with sales again. Thus, when an individual firm sees inventories increase and sales decrease, it typically scales down production to match the decrease in demand.

Now let us think about how this works at the level of an economy as a whole. Suppose we divide total spending in the economy into unplanned inventory investment and everything else, which we call **planned spending**.

Toolkit: Section 31.30

Planned spending is all expenditure in the economy except for unplanned inventory investment:

$$GDP = planned\ spending + unplanned\ inventory\ investment.$$

This equation must always hold true because of the rules of national income accounting.

Begin with the situation where there is no unplanned inventory investment—so GDP equals planned spending—and then suppose that planned spending decreases. Firms find that their production is in excess of their sales, so their inventory builds up. As we just argued, they respond by decreasing production so that GDP is again equal to planned spending, and unplanned inventory investment is once again zero. Thus, even though unplanned inventory investment *can* be nonzero for very short periods of time, we do not expect such a situation to persist. We expect instead that actual output will, in fact, almost always equal planned spending.

What Happened to the Components of GDP during the Great Depression?

Now let us look at how these components of GDP behaved during the 1930s. Table 22.4.1 "Growth Rates of Key Macroeconomic Variables, 1930–39*" presents these data in the form of growth rates. Remember that a positive growth rate means the variable in question increased from one year to the next, while a negative growth rate means it decreased.

Table 22.4.1: Growth Rates of Key Macroeconomic Variables, 1930–39*

Growth Rates	1930	1931	1932	1933	1934	1935	1936	1937	1938	1939
Real GDP	−8.6	−6.4	−13.0	−1.3	10.8	8.9	13.0	5.1	−3.4	8.1
Consumption	−5.3	−3.1	−8.9	−2.2	7.1	6.1	10.1	3.7	−1.6	5.6
Investment	−33.3	−37.2	−69.8	47.5	80.5	85.1	28.2	24.9	−33.9	28.6

*This table shows growth rates in real GDP, consumption, investment, and government purchases. All data are from the National Income and Product Accounts web page, Bureau of Economic Analysis, Department of Commerce (www.bea.gov/national/nipaweb/index.asp).

Growth Rates	1930	1931	1932	1933	1934	1935	1936	1937	1938	1939
Government Purchases	10.2	4.2	−3.3	−3.5	12.8	2.7	16.7	−4.2	7.7	8.8

*This table shows growth rates in real GDP, consumption, investment, and government purchases. All data are from the National Income and Product Accounts web page, Bureau of Economic Analysis, Department of Commerce (www.bea.gov/national/nipaweb/index.asp).

We see again that real GDP decreased for four years in succession (the growth rates are negative from 1930 to 1933). The decrease in real GDP was accompanied by a decline in consumption and investment: consumption likewise decreased for four successive years, and investment decreased for three successive years. The decline in consumption was not as steep as the decline in real GDP, while the decline in investment was much larger. Were we to drill deeper and look at the components of consumption, we would discover that expenditures on durable goods decreased by 17.6 percent in 1930 and 25.1 percent in 1932, while expenditures on services decreased by only 2.5 percent in 1930 and 6.3 percent in 1932.

Whatever was happening during this period evidently had a much larger influence on firms' purchases of investment goods, and on households' spending on cars and other durable goods, than it did on purchases of nondurable goods (such as food) and services (such as haircuts). A similar pattern can be observed in modern economies: consumption is smoother than output, and spending on services is smoother than spending on durables. The reason for this is a phenomenon that economists call **consumption smoothing**.

Toolkit: Section 31.34

Consumption smoothing is the idea that households like to keep their flow of consumption relatively steady over time. When income is unusually high, the household saves (or pays off existing loans); when income is unusually low, the household borrows (or draws down existing savings). Consumption smoothing is a key ingredient of the life-cycle model of consumption, which is discussed in more detail in the toolkit.

If your company has a good year and you get a big bonus, you will increase consumption spending not only this year but also in future years. To do so, you must *save* a portion of your bonus to pay for this higher consumption in the future. By the same logic, if your income decreases, your consumption will not decrease as much. People who became unemployed during the Great Depression did not reduce their consumption of services and nondurable goods to zero. Instead, as far as was possible, they drew on their existing savings, borrowed, and postponed purchases of durable goods.

Consumption of durable goods, in other words, resembles investment rather than consumption of nondurable goods and services. This makes sense because durable goods resemble investment goods that are purchased by households. Like investment goods, they yield benefits over some prolonged period of time. As an example, consider automobile purchases during the Great Depression. Although 5.4 million cars were produced in 1929, only 3.4 million were produced in 1930—a reduction of more than 37 percent in a single year. Instead of buying new cars, households simply held onto their existing cars longer. As a consequence of the boom of the 1920s, there were a lot of relatively new cars on the road in 1929: the number of cars less than 3 years old was about 9.5 million. Two years later, this number had fallen to 7.9 million.These figures are from Michael Bernstein, *The Great Depression: Delayed Recovery and Economic Change in America, 1929–39* (Cambridge, MA: Cambridge University Press, 1987).

This reduction in activity in the automobile industry was matched by a reduction of inputs into the production process. By early 1933, there were only 4 workers for every 10 who had been employed 4 years previously. Equipment purchases for the transportation sector were so low that capital stock for this sector decreased between 1931 and 1935. In the turmoil of the Great Depression, many small car producers went out of business, leaving a few relatively large companies—such as Ford Motor Company and GM—still in business.

Similar patterns arose as the economy recovered. Investment, in particular, was astonishingly volatile. It decreased by about one-third in 1930 and again in 1931, and by over two-thirds in 1932, but rebounded at an astoundingly high rate after 1933. Consumption, meanwhile, grew at a slower rate than GDP as the economy recovered.

Key Takeaways

- The components of aggregate spending are consumption, investment, government purchases of goods and services, and net exports.
- The national income identity states that real GDP is equal to the sum of the components of aggregate spending.
- During the Great Depression, both consumption spending and investment spending experienced negative growth.
- Households use savings to retain relatively smooth consumption despite fluctuations in their income.

Exercises

1. Explain the difference between investment spending in the national income and product accounts and a decision to buy shares of a company.
2. If someone is unemployed and receives unemployment benefits from a state government, are those funds counted in aggregate expenditure?

This page titled 22.4: The Components of GDP during the Great Depression is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

22.5: The Great Depression- A Decrease in Aggregate Spending? 22.5: The Great Depression- A Decrease in Aggregate Spending?Edit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. How did the perspective of Keynes differ from the "classical theory" of the macroeconomy?
2. How does a decrease in aggregate spending lead to a reduction in real gross domestic product (real GDP)?
3. Can a decrease in consumption explain the Great Depression?
4. Can a decrease in investment explain the Great Depression?

Now that we understand the components of aggregate spending, we can consider whether a decrease in one or more of these components can explain the Great Depression.

A Decrease in Aggregate Spending: The Multiple-Markets Perspective

Consider, as before, the market for new houses and suppose there is a reduction in spending on houses. Market demand shifts inward, causing a decrease in the price of houses, as shown in Figure 22.5.1. The lower price means that construction firms choose to build fewer houses; there is a movement along the supply curve.

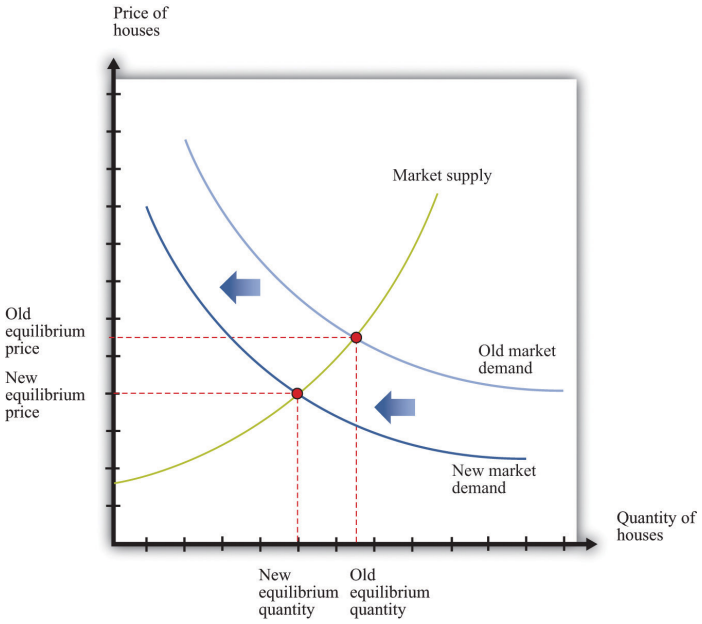


Figure 22.5.1: An Inward Shift in Market Demand for Houses. A decrease in demand for houses leads to a decrease in the price of houses and a lower quantity of houses being produced and sold.

As before, the effects are not confined to the housing market. Construction firms demand less labor, so the wages of these workers decrease. Employment in the construction industry declines, but these workers now seek jobs in other sectors of the economy. The increased supply of labor in these sectors reduces wages and thus makes it more attractive for firms to increase their hiring. Supply curves in other sectors shift rightward. Moreover, the income that was being spent on housing will instead be spent somewhere else in the economy, so we expect to see rightward shifts in demand curves in other sectors as well. In summary, if we are looking at the whole economy, a decrease in spending in one market is not that different from a decrease in technology in one market: we expect a reduction in one sector to lead to expansions in other sectors. The economy still appears to be self-stabilizing.

In this story, as is usual when we use supply and demand, we presumed that prices and wages adjust quickly to bring supply and demand into line. This is critical for the effective functioning of markets: for markets to do a good job of matching up demand and supply, wages and prices must respond rapidly to differences between supply and demand. **Flexible prices** adjust immediately to shifts in supply and demand curves so that price is always at the point where supply equals demand. If, for example, the quantity of

labor supplied exceeds the quantity of labor demanded, flexible wages decrease quickly to bring the labor market back into equilibrium.

Suppose we instead entertain the possibility that wages and prices do not immediately adjust. **Sticky prices** do not react immediately to shifts in supply and demand curves, and the adjustment to equilibrium can take some time. We defer for the moment the discussion of *why* prices might be sticky and concentrate instead on the implications of this new idea about how markets work. The easiest way to see the effects of price stickiness is to suppose that prices do not change at all. Figure 22.5.2 shows the impact of a decrease in demand for houses when the price of houses is completely sticky. If you compare Figure 22.5.2 to Figure 22.5.1, you see that a given shift in demand leads to a larger change in the quantity produced.

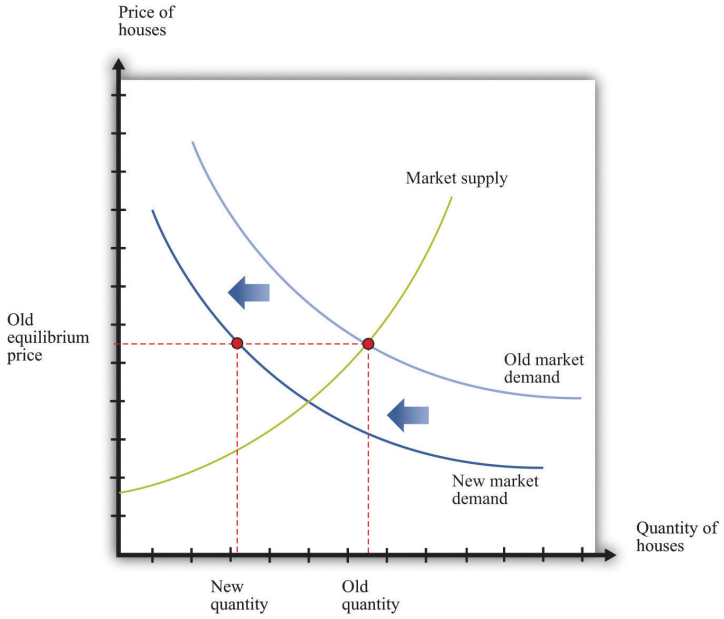


Figure 22.5.2: A Shift in Demand for Houses When Prices Are Sticky. If the price in the market is "sticky," it may not adjust immediately to the change in demand, resulting in a large decrease in the quantity of houses that are produced and sold.

What about the effects on other markets? As before, a decrease in demand for housing will cause construction workers to lose their jobs. If wages are sticky, these workers may become unemployed for a significant period of time. Their income decreases, and they consume fewer goods and services. So, for example, the demand for beef in the economy might decrease because unemployed construction workers buy cheaper meat. This means that the demand for beef shifts inward. The reduction in activity in the construction sector leads to a reduction in activity in the beef sector. And the process does not stop there—the reduced income of cattle farmers and slaughterhouse workers will, in turn, spill over to other sectors.

What has happened to the self-stabilizing economy described earlier? First, sticky wages and prices impede the incentives for workers to flow from one sector to another. If wages are sticky, then the reduction in labor demand in the construction sector does not translate into lower wages. Thus there is no incentive for other sectors to expand. Instead, these other sectors, such as food, see a decrease in demand for their product, which leads them to contract as well. Second, the decrease in income means that it is possible to see decreases in demand across the entire economy. It no longer need be the case that reductions in spending in one area lead to increased spending in other sectors.

The Circular Flow of Income during the Great Depression

So far, we have told this story in terms of individual markets. The circular flow helps us see how these markets come together in the aggregate economy. When we looked at the markets for housing and beef, we saw that a decrease in demand for housing led to a decrease in demand for labor and, hence, to lower labor income. We also saw that as income earned in the housing market decreased, spending decreased in the beef market. Such linkages are at the heart of the circular flow of income. Household spending on goods and services is made possible by a flow of income from firms. Firms' hiring of labor is made possible by a flow of revenue from households. Keynes argued that this was a delicate process that might be prone to malfunction in a variety of ways.

Households are willing to buy goods and services if they have a reasonable expectation that they can earn income by selling labor. During the Great Depression, however, household expectations were surely quite pessimistic. Individuals without jobs believed that their chances of finding new employment were low. Those lucky enough to be employed knew that they might soon be out of work.

Thus households believed it was possible, even likely, that they would receive low levels of income in the future. In response, they cut back their spending.

Meanwhile, the willingness of firms to hire labor depends on their expectation that they can sell the goods they manufacture. When firms anticipate a low level of demand for their products, they do not want to produce much, so they do not need many workers. Current employees are laid off, and there are few new hires.

Through the circular flow, the pessimism of households and the pessimism of firms interact. Firms do not hire workers, so household income is low, and households are right not to spend much. Households do not spend, so demand for goods and services is low, and firms are right not to hire many workers. The pessimistic beliefs of firms and workers become self-fulfilling prophecies.

The Aggregate Expenditure Model

In the remainder of this section, we build a framework around the ideas that we have just put forward. The framework focuses on the determinants of aggregate spending because, in this approach, the output of the economy is determined not by the level of potential output but by the level of total spending. This model is based around the idea of sticky prices—or, more precisely, it tells us what the output of the economy will be, *at a given value of the overall price level*. Once we understand this, we can add in the effects of changing prices.

Earlier, we introduced the national income identity:

$$production = consumption + investment + government\ purchases + netexports.$$

This equation must be true by the way the national income accounts are constructed. That is, it is an accounting *identity*. We also explained that

$$GDP = planned\ spending + unplanned\ inventory\ investment.$$

It is possible for firms to accumulate or decumulate inventories unintentionally, but such a situation will not persist for long. Firms quickly respond to such imbalances by adjusting their production. The aggregate expenditure model takes the national income identity and adds to it the condition that unplanned inventory investment equals zero—equivalently, gross domestic product (GDP) equals planned spending:

$$planned\ spending = consumption + investment + government\ purchases + net\ exports.$$

Another way of saying this is that as long as we interpret investment to include only planned investment, the national income equation is no longer an identity but instead a condition for equilibrium.

The Relationship between Planned Spending and Output

We could now examine all four components of planned spending separately. Different chapters of this book delve deeper into these types of spending. For the moment, however, we group them all together. We focus on the fact that total planned spending depends positively on the level of income and output in an economy, for two main reasons:

1. If households have higher income, they are likely to increase their spending on many goods and services. The relationship between income and consumption is one of the cornerstones of macroeconomics.
2. Firms are likely to decide that higher levels of output—particularly if expected to persist—mean that they should build up their capital stock and thus increase their investment.

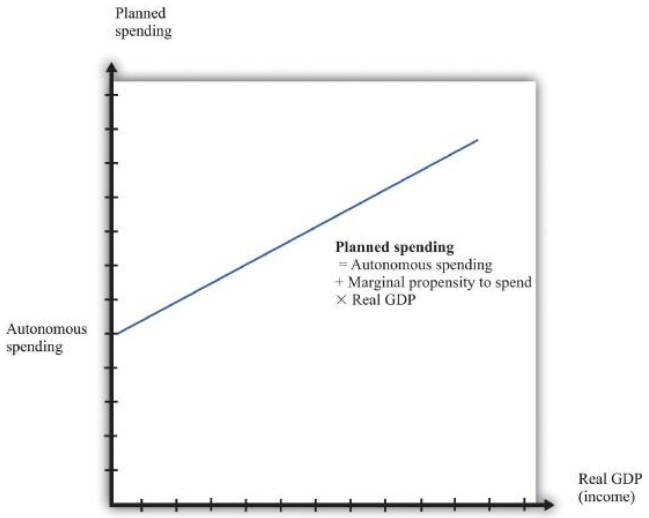


Figure 22.5.3: The Planned Spending Line. Planned spending is composed of autonomous spending (the amount of spending when real GDP equals zero) and induced spending (spending resulting from real GDP).

In summary, we conclude that when income increases, planned expenditure also increases. We illustrate this in Figure 22.5.3, where we suppose for simplicity that the relationship between planned spending and GDP is a straight line:

$$planned\ spending\ =\ autonomous\ spending\ +\ marginal\ propensity\ to\ spend \times GDP.$$

Autonomous spending is the *intercept* of the planned spending line. It is the amount of spending that there would be in an economy if income were zero. It is positive, for two reasons: (1) A household with no income still wants to consume something, so it will either draw on its existing savings or borrow against future income. (2) The government purchases goods and services even if income is zero.

The **marginal propensity to spend** is the *slope* of the planned spending line. It tells us how much planned spending increases if there is a $1 increase in income. The marginal propensity to spend is positive: Increases in income lead to increased spending by households and firms. The marginal propensity to spend is less than one, largely because of consumption smoothing by households. If household income increases by $1, households typically consume only a fraction of the increase, saving the remainder to finance future consumption. This equation, together with the condition that GDP equals planned spending, gives us the **aggregate expenditure model**.

Toolkit: Section 31.30

The aggregate expenditure model takes as its starting point the fact that GDP measures both total spending and total production. The model focuses on the relationships between output and spending, which we write as follows:

$$planned\ spending\ =\ GDP$$

and

$$planned\ spending\ =\ autonomous\ spending\ +\ marginal\ propensity\ to\ spend \times GDP.$$

The model finds the value of output for *a given value of the price level*. It is then combined with a model of price adjustment to give a complete picture of the economy.

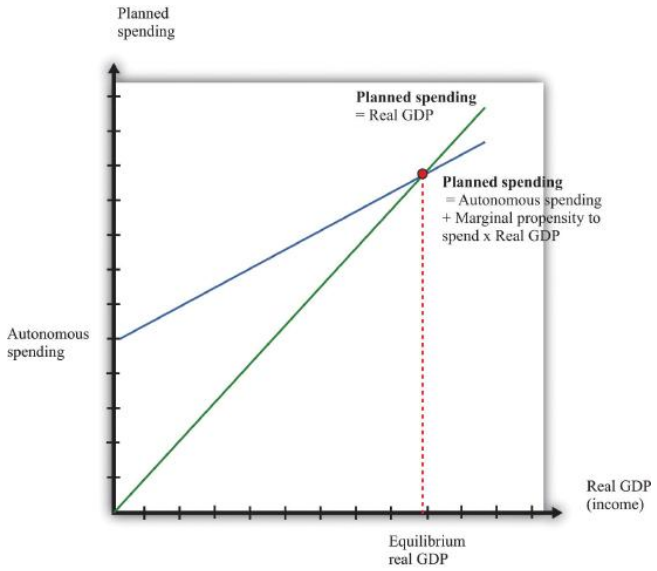


Figure 22.5.4: Equilibrium in the Aggregate Expenditure Model. The aggregate expenditure framework tells us that the economy is in equilibrium when planned spending equals real GDP.

We can solve the two equations to find the values of GDP and planned spending that are consistent with both equations:

We can also take a graphical approach, as shown in Figure 22.5.4. On the horizontal axis is the level of real GDP, while on the vertical axis is the overall level of (planned) spending in the economy. We graph the two relationships of the aggregate expenditure model. The first line is a 45° line—that is, it is a line with a slope equal to one and passing through the origin. The second is the planned spending line. The point that solves the two equations is the point where the two lines intersect. This diagram is the essence of the aggregate expenditure model of the macroeconomy.

The aggregate expenditure model makes no reference to potential output or the supply side of the economy. The model assumes that the total amount of output produced will always equal the quantity demanded at the given price. You might think that this neglect of the supply side is a weakness of the model, and you would be right. In 22.4 Section "Price Adjustment", when we introduce the adjustment of prices, the significance of potential output becomes clear.

Can a Decrease in Consumption Spending Explain the Great Depression?

We now apply this framework to the Great Depression. The aggregate expenditure approach suggests that output decreased in the Great Depression because aggregate spending decreased. Part (a) of Figure 22.5.5 shows how this process begins: a decrease in autonomous spending shifts the spending line down. The interpretation of such a shift is that, *at every level of income*, spending is lower. Such a decrease in spending is due to a decrease in (the autonomous component of) consumption, investment, government spending, or net exports (or some combination of these). Part (b) of Figure 22.5.5 shows what happens when the planned spending line shifts downward. The equilibrium level of real GDP decreases. So far, therefore, the aggregate expenditure model seems to work: a decrease in autonomous spending leads to a decrease in real GDP at the given price level. But we need to know *why* planned spending decreased.

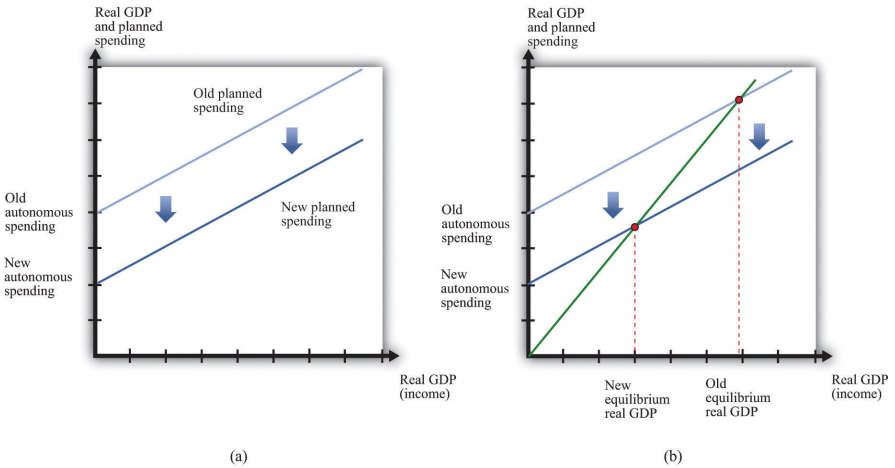


Figure 22.5.5: A Decrease in Aggregate Expenditures. The Keynesian explanation of the Great Depression is that a decrease in autonomous spending caused the planned spending line to shift downward (a) leading to a decrease in the equilibrium level of real GDP (b).

Let us first consider the possibility that a reduction in consumption triggered the Great Depression. Recall that, between September and November 1929, the stock market in the United States crashed. This collapse meant that many households were suddenly less wealthy than they had been previously. A natural response to a decrease in wealth is to decrease consumption; this is known as a **wealth effect**.

Wealth is distinct from income. Income is a flow: a household's income is the amount that it receives over a period of time, such as a year. Wealth is a stock: it is the cumulated amount of the household's savings. Is it plausible that wealth effects could explain a collapse of the magnitude of the Great Depression? To answer this, we need to determine how much real GDP decreases for a given change in autonomous spending.

The Multiplier

Toolkit: Section 31.30

The solution for output in the aggregate expenditure model can be written in terms of changes as follows:

$$change\ in\ GDP\ =\ multiplier \times change\ in\ autonomous\ spending,$$

where the **multiplier** is given by

Suppose that the marginal propensity to spend is 0.8. Then

A given change in autonomous spending will lead to a fivefold change in real GDP. Economists refer to this as a multiplier process. Because (1 − marginal propensity to spend) is less than one, the multiplier is a number greater than one. This means that any change in autonomous spending is multiplied up to result in a larger change in GDP. Even relatively small decreases in spending can end up being damaging to an economy.

The economics behind the multiplier comes from the circular flow of income. Begin with a decrease in autonomous spending. The reduction in spending means less demand for firms' goods and services. Firms respond by cutting output. (As a reminder, the signal to firms that they should cut their output comes from the fact that they see a buildup of their inventory.) When firms cut their output, they require less labor and pay out less in wages, so household income decreases. This causes households to again cut back on consumption, so spending decreases further. Thus we go round and round the circular flow diagram: decreased spending leads to decreased output, which leads to decreased income, which leads to decreased spending, which leads to decreased output, and so on and so on. The process continues until the reductions in income, output, and consumption in each round are tiny enough to be ignored.

We use the multiplier to carry out comparative static exercises in the aggregate expenditure model. In this case, the endogenous variable is real GDP, and the exogenous variable is autonomous spending. Given a change in autonomous spending, we simply multiply by the multiplier to get the change in real GDP when the price level is fixed. Let us do some back-of-the-envelope comparative static calculations, based on the assumption that the marginal propensity to spend is 0.8, so the multiplier is 5.

Table 22.2.1 "Major Macroeconomic Variables, 1920–39*" tells us that real GDP decreased by approximately $75 billion between 1929 and 1930. With a multiplier of 5, we would need a drop in autonomous spending of $75 billion divided by 5, or $15 billion, to get this large a decrease in GDP. The population of the United States in 1930 was approximately 123 million, so a $15 billion decrease in spending corresponds to about $122 per person. Remember that the figures in Table 22.2.1 "Major Macroeconomic Variables, 1920–39*" are in terms of year 2000 dollars. It certainly seems plausible that households, who had been made significantly poorer by the collapse in the stock market, would have responded by cutting back spending by the equivalent today of a few hundred dollars per year.

Our goal, you will remember, is to explain the events of the Great Depression. How are we doing so far? The good news is that we do have a story that explains how output could decrease as precipitously as it did in the Great Depression years: there was a major stock market crash, which made people feel less wealthy, so they decided to consume less and save more.

If we look more closely, though, this story still falls short. When we examined the data for the Great Depression, we saw that—while output and consumption both decreased—consumption decreased much less than did output. For example, from 1929 to 1933, real GDP decreased by 26.5 percent, while consumption decreased by 18.2 percent. By contrast, investment (that is, purchases of capital by firms, new home construction, and changes in business inventories) decreased much more than output. In 1932, purchases of new capital were $11 billion (year 2000 dollars), compared to a level of $91 billion in 1929. This is a reduction in real investment of about 82 percent. We must look more closely at investment to see if our theory can also explain the different behavior of consumption and investment.

Can a Decrease in Investment Spending Explain the Great Depression?

When GDP decreases, there can be an *induced* decrease in investment: declines in income lead firms to anticipate lower production in the future, meaning they see less of a need to build up their capital stock. But the changes in investment during the Great Depression were very large. Because it is implausible that such large variation was the result of changes in output alone, economists look for additional explanations of why investment decreased so much during the Great Depression.

During the Great Depression, the link between savings and investment was disrupted by bank failures. Between 1929 and 1933, a number of US banks went out of business, often taking the savings of households with them. People began to trust banks less, and many households stopped putting their savings into the financial sector. The financial sector is an intermediary between households and firms, matching up the supply of savings from households with the demand for savings by firms. Figure 22.5.6 shows the flows in and out of the financial sector. (Our focus here is on the role of this sector in matching savers and investors. As Figure 22.5.6 shows, however, funds also flow into (or from) the financial sector from the rest of the world and the government sector.)

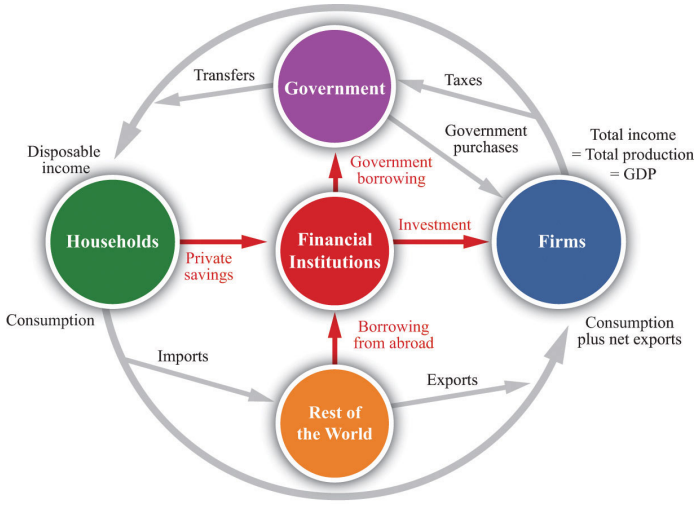


Figure 22.5.6: The Financial Sector in the Circular Flow of Income. Financial institutions such as banks act as intermediaries in the circular flow of income. During the Great Depression, many banks failed, disrupting the matching of savings and investment.

To understand bank failures in the Great Depression, we need to take a moment to review what banks do. A bank is an institution that accepts money ("bank deposits") from individuals. It then takes some of that money and puts it into longer-term projects—the construction of an apartment building, for example. The bank in this case issues a long-term loan to the company that plans to construct the new building.

At any time, a bank has a portfolio of assets. Some are **liquid**; they are easily and quickly exchanged for cash. Some are **illiquid**; they cannot easily be converted into cash. Banks keep some assets in a highly liquid form, such as cash or very short-term loans, and also hold assets that are relatively illiquid, such as a two-year loan to a construction company.

At any time, depositors at a bank can choose to withdraw their money. Under normal circumstances, people are happy to leave most of their money in the bank, so only a small fraction of depositors want to withdraw money on any given day. The bank keeps some cash in its vaults to accommodate this demand. But suppose that times are not normal. Suppose that, as was the case during the Great Depression, depositors start to see that other banks are going out of business. Then they may worry that their own bank is also at risk of failing, in which case they will lose their savings. The natural response is to rush to the bank to withdraw money before the bank fails.

If a large number of depositors all try to withdraw money at once, the bank will run out of cash and other liquid assets. It will not be able to meet the needs of its depositors. The consequence is a **bank run**. And if the bank is unable to meet its depositors' demands, it may be forced out of business altogether. This is known as a **bank failure**.

A striking feature of a bank failure caused by a bank run is that it is a self-fulfilling prophecy:

- If everybody believes that the bank is safe, then no one will withdraw money, and the bank will indeed be safe.
- If everybody believes that the bank is going to fail, then everyone will try to withdraw money, and the bank will indeed fail.

Notice that *every individual's decision about what to do is based on what that individual expects everyone else will do*.

Figure 22.5.7 presents the decisions underlying a bank run in a stylized way. Imagine that you deposit $100 in the bank. The table in the figure shows how much you obtain, depending on your own actions and those of other depositors. You and the other depositors must decide whether to leave your money in the bank ("don't run") or try to take your money out of the bank ("run"). If everyone else leaves money in the bank, then you can withdraw your money and get $100 or leave it in the bank and get the $100 plus $10 interest. If others do not run, then it is also best for you not to run. But if everyone else runs on the bank, then you get nothing if you leave your money in the bank, and you can (in this example) recover $20 if you run to the bank along with everyone else. Thus, if you expect others to run on the bank, you should do the same.

	Others leave their money in the bank	Others take their money out of the bank
You leave your money in the bank (don't run)	110	0
You leave your money in the bank (don't run)	100	20

Figure 22.5.7: Payoffs in a Bank-Run Game. This table shows the payoffs in a bank-run game. That is, it shows you what you get back depending on your choice and everybody else's choice about whether to run on the bank. If everyone else leaves money in the bank, then you should do the same, but if everyone else runs on the bank, you are better running as well.

Economists call this situation a **coordination game**. In a coordination game, there are *multiple equilibria*. In this example, there is one equilibrium where there is no run on the bank, and there is another equilibrium where everyone runs to the bank to withdraw funds.

Toolkit: Section 31.18

You can find more details on coordination games in the toolkit.

During the Great Depression, a story such as this played out not only at one bank but at many. Figure 22.5.8 shows what happened in terms of the aggregate expenditure framework. Prior to the Great Depression, the economy was in a "high confidence" equilibrium, in which the banking system was healthy and confidence was high. Then—for some reason—people became nervous about leaving money in banks, and it became much harder for firms to obtain loans. The cost of borrowing—the **real interest rate**—increased, and investment decreased substantially. The planned spending line shifted downward, and the economy moved to the bad "low confidence" equilibrium. The downward shift in planned spending leads to a decrease in real GDP, given the existing level of prices.

Figure 22.5.8 should look familiar; it is the same as part (b) of Figure 22.5.5. This is because a decrease in autonomous consumption and a decrease in autonomous investment both look the same in the aggregate expenditure model, even though the underlying story is different. Of course, it is also possible that both autonomous consumption and autonomous investment decreased.

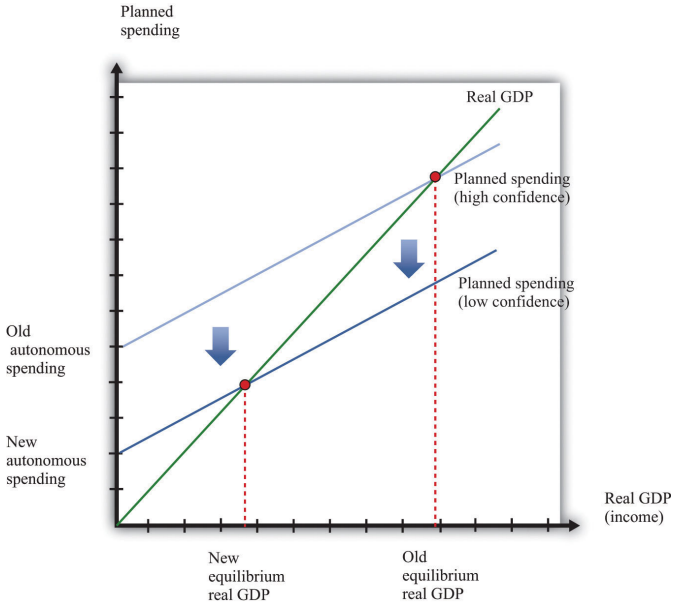


Figure 22.5.8. Failures in the financial sector lead to a drop in investment spending. During the Great Depression, a decrease in confidence in the banking system meant that many banks failed, and it became more difficult and expensive for firms to borrow. The planned spending line shifted downward, and real GDP decreased.

To summarize, the banking crisis made households reluctant to put money in the banks, and banks were reluctant to make loans. Two banking measures help us see what was happening. The **currency-deposit ratio** is the total amount of currency (that is, either banknotes or coins) divided by the total amount of deposits in banks. The **loan-deposit ratio** is the total amount of loans made by banks divided by the total amount of deposits in banks.

If the currency-deposit ratio is low, households are not holding very much cash but are instead keeping wealth in the form of bank deposits and other assets. The currency-deposit ratio increased from 0.09 in October 1929 to 0.23 in March 1933.See Milton Friedman and Anna Schwartz, *A Monetary History of the United States, 1867–1960* (Princeton, NJ: Princeton University Press, 1963), Table B3. This means that households in the economy started holding onto cash rather than depositing it in banks. You can think of the loan-deposit ratio as a measure of the productivity of banks: banks take deposits and convert them into loans for investment. During the Great Depression, the loan-deposit ratio decreased from 0.86 to 0.73.See Ben Bernanke, "Nonmonetary Effects of the Financial Crisis in the Propagation of the Great Depression," *American Economic Review* 73 (1983): 257–76, Table 1.

Price Adjustment

The story we have told explains why the economy departs from potential output but says nothing about how (if at all) the economy gets back to potential output. The answer is that prices have a tendency to adjust back toward their equilibrium levels, even if they do not always get there immediately. This is most easily understood by remembering that prices in the economy are, in the end,

usually set by firms. When a firm sees a decrease in demand for its product, it does not necessarily decrease its prices immediately. Its decision about what price to choose depends on the prices of its inputs and the prices being set by its competitors. In addition, it depends on not only what those prices are right now but also what the firm expects to happen in the future. Deciding exactly what to do about prices can be a difficult decision for the managers of a firm.

Without analyzing this decision in detail, we can certainly observe that firms often keep prices fixed when demand decreases—at least to begin with. The result looks like that in Figure 22.5.1. In the face of a prolonged decrease in demand, however, firms *will* lower prices. Some firms do this relatively quickly; others keep prices unchanged for longer periods. We conclude that prices are sticky; they do not decrease instantly, but they decrease eventually. Chapter 25 provides more detail about the price-adjustment decisions of firms. For the economy as a whole, this adjustment of prices is represented by a **price-adjustment equation**.

Toolkit: Section 31.31

The difference between potential output and actual output is called the **output gap**:

$$output\ gap\ =\ potential\ real\ GDP\ -\ actual\ real\ GDP.$$

If an economy is in recession, the output gap is positive. If an economy is in a boom, then the output gap is negative. The inflation rate when an economy is at potential output (that is, when the output gap is zero) is called **autonomous inflation**. The overall inflation rate depends on both autonomous inflation and the output gap, as shown in the price-adjustment equation:

$$inflation\ rate\ =\ autonomous\ inflation\ -\ inflation\ sensitivity \times output\ gap.$$

This equation tells us that there are two reasons for increasing prices.

1. Prices increase because autonomous inflation is positive. Even when the economy is at potential output, firms may anticipate that their suppliers or their competitors are likely to increase prices in the future. A natural response is to increase prices, so autonomous inflation is positive.
2. Prices increase because the output gap is negative. The output gap matters because, as GDP increases relative to potential, labor and other inputs become scarcer. Firms see increasing costs and choose to set higher prices as a consequence. The "inflation sensitivity" tells us how responsive the inflation rate is to the output gap.

When real GDP is above potential output, there is upward pressure on prices in the economy. The inflation rate exceeds autonomous inflation. By contrast, when real GDP is below potential, there is downward pressure on prices. The inflation rate is below the autonomous inflation rate. The price-adjustment equation is shown in Figure 22.5.9.

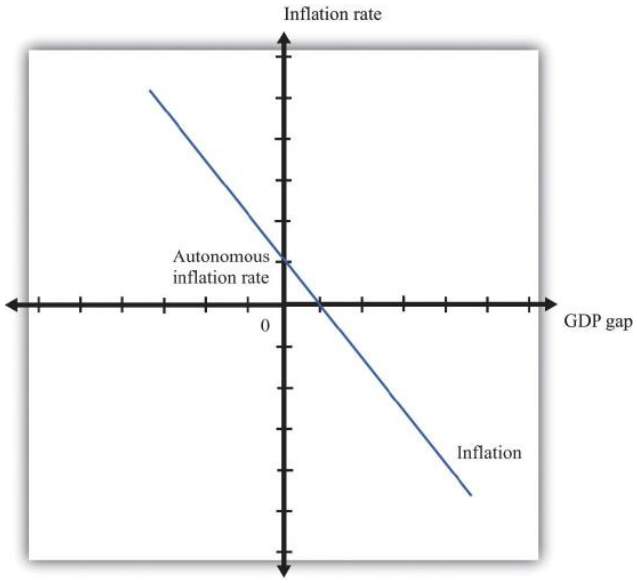


Figure 22.5.9: Price Adjustment. When an economy is in a recession, actual inflation is lower than autonomous inflation. In a boom, inflation is higher than its autonomous level.

We can apply this pricing equation to the Great Depression. Imagine first that autonomous inflation is zero. In this case, prices decrease when output is below potential. From 1929 to 1933, output was surely below potential and, as the equation suggests, this

was a period of decreasing prices. After 1933, as the economy rebounded, the increase in the level of economic activity was matched with positive inflation—that is, increasing prices. This turnaround in inflation occurred even though the economy was still operating at a level below potential output. To match this movement in prices, we need to assume that—for some reason that we have not explained—autonomous inflation became positive in this period.

Key Takeaways

- Keynes argued that, at least in the short run, markets were not able to fully coordinate economic activity. His theory gave a prominent role to aggregate spending as a determinant of real GDP.
- Given prices, a reduction in spending will lead to a reduction in the income of workers and owners of capital, which will lead to further reductions in spending. This link between income and spending is highlighted by the circular flow of income and underlies the aggregate expenditure model.
- The stock market crash in 1929 reduced the wealth of many households, and this could have led them to cut consumption. This reduction in aggregate spending, through the multiplier process, could have led to a large reduction in real GDP.
- The reductions in investment in the early 1930s, perhaps coming from instability in the financial system, could lead to a reduction in aggregate spending and, through the multiplier process, a large reduction in real GDP.

Exercises

1. Some researchers have suggested that a reduction in US net exports is another possible cause of the Great Depression. Use the aggregate expenditure model to consider the effects of a reduction in net exports. What happens to real GDP?
2. Suppose autonomous inflation is constant, but real GDP moves around. Would you expect inflation to be procyclical or countercyclical?

This page titled 22.5: The Great Depression- A Decrease in Aggregate Spending? is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

22.6: Policy Interventions and the Great Depression 22.6: Policy Interventions and the Great DepressionEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What is stabilization policy?
2. What is monetary policy, and how was it used during the Great Depression?
3. What is fiscal policy, and how was it used during the Great Depression?

Understanding why the Great Depression occurred is certainly progress. But policymakers also wanted to know if there was anything that could be done in the face of this economic catastrophe. One of Keynes' most lasting contributions to economics is that he showed how different kinds of economic policy could be used to assist economies that were stuck in recessions.

When markets are doing a good job of allocating resources, standard economic reasoning suggests that it is better for the government to stay out of the way. But when markets fail to allocate resources well, the government might be able to improve the overall functioning of the economy. The idea that markets left alone would coordinate aggregate economic activity is difficult to defend in the face of 25 percent unemployment of the labor force and a decline in economic activity of nearly 30 percent over a 4-year period. Thus the rationale for government intervention in the aggregate economy is that markets are failing to allocate resources properly, perhaps because prices and wages are sticky.

Policy Remedies

In the wake of the Great Depression, economists started advocating the use of government policy to improve the functioning of the macroeconomy. There are two kinds of government policy. **Monetary policy** refers to changes in interest rates and other tools that are under the control of the monetary authority of a country (the central bank). **Fiscal policy** refers to changes in taxation and the level of government purchases; such policies are typically under the control of a country's lawmakers. **Stabilization policy** is the general term for the use of monetary and fiscal policies to prevent large fluctuations in real gross domestic product (real GDP).

In the United States, the Federal Reserve Bank controls monetary policy, and fiscal policy is controlled by the president, the Congress, and state governments. In the countries of the European Union, monetary policy is controlled by the European Central Bank, and fiscal policies are controlled by the individual governments of the member countries.

Keynes suggested that the cause of the Great Depression was an unusually low level of aggregate spending. This diagnosis suggests an immediate remedy: use government policies to increase aggregate spending. Because

$$change\ in\ GDP\ =\ multiplier \times change\ in\ autonomous\ spending,$$

any government policy that increases autonomous spending will, through this equation, also increase GDP. There are many different policies at the disposal of the government, but they are similar at heart. The idea is to stimulate one of the components of aggregate spending—consumption, investment, government purchases, or net exports.

One fiscal policy measure is an increase in government purchases. Suppose the government increases its expenditure—perhaps by hiring more teachers, buying more tanks, or building more roads. This increases autonomous spending and works its way through the economy, just as in our earlier discussion of a decrease in autonomous consumption—except now we are talking about an increase rather than a decrease. If the government spends an extra dollar, this immediately expands income by that dollar. Extra income leads to extra spending, which leads to further increases in output and income. The process continues around and around the circular flow.

Imagine that, as before, the marginal propensity to spend is 0.8, so that the multiplier is 5. If the government increases expenditure on goods and services by $1 billion, overall GDP in the economy will increase by $5 billion. Thus to offset the decrease in real GDP of about $90 billion between 1929 and 1933, assuming a marginal propensity to spend of 0.8, the federal government should have increased government spending by $18 billion. The multiplier is a double-edged sword. It has the bad effect that it can turn small decreases in spending into big decreases in output. But it also means that relatively small changes in government spending can have a big effect on output.

Tax cuts are another way to stimulate the economy. If households have to pay fewer taxes to the government, they are likely to spend more on consumption goods. This form of policy intervention has been used over and over again by governments in the United States and elsewhere. Tax cuts, like government spending, must be paid for. If the government spends more and taxes less, then the government deficit increases. The government must borrow to finance such fiscal policy measures. Chapter 27 and Chapter 29 have more to say about fiscal policy.

The central bank can use monetary policy to affect aggregate spending. Monetary policy operates through changes in interest rates, which are—in the short run at least—under the influence of the central bank. Lower interest rates make it cheaper for firms to borrow, which encourages them to increase investment spending. Lower interest rates likewise mean lower mortgage rates, so households are more likely to buy new homes. Lower interest rates may encourage households to borrow and spend more on other goods. And lower interest rates can even encourage net exports.The link from interest rates to net exports is complicated because it involves changes in exchange rates. You do not need to worry here about how it works. We explain it, together with other details of monetary policy, in Chapter 25.

Monetary and Fiscal Policies during the Great Depression

We have argued that monetary and fiscal policies *could* have been used to help the economy out of the Great Depression. But what did policymakers actually do at the time? The answer comes in two parts: at the start of the Great Depression, they did not do much; after 1932, they did rather more.

Both presidential candidates campaigned in favor of conservative fiscal policy in 1932. Here are some excerpts from the party platforms.See John Woolley and Gerhard Peters, *The American Presidency Project*, accessed June 30, 2011, http://www.presidency.ucsb.edu.

From the Democratic Party platform:

We advocate an immediate and drastic reduction of governmental expenditures by abolishing useless commissions and offices, consolidating departments and bureaus, and eliminating extravagance to accomplish a saving of not less than twenty-five per cent in the cost of the Federal Government. And we call upon the Democratic Party in the states to make a zealous effort to achieve a proportionate result.

We favor maintenance of the national credit by a federal budget annually balanced on the basis of accurate executive estimates within revenues, raised by a system of taxation levied on the principle of ability to pay.“Democratic Party Platform of 1932,” *The American Presidency Project*, accessed June 30, 2011, http://www.presidency.ucsb.edu/ws/index.php?pid=29595#axzz1N9yDnpSR.

From the Republican Party platform:

The President’s program contemplates an attack on a broad front, with far-reaching objectives, but entailing no danger to the budget. […]

Constructive plans for financial stabilization cannot be completely organized until our national, State and municipal governments not only balance their budgets but curtail their current expenses as well to a level which can be steadily and economically maintained for some years to come. “Republican Party Platform of 1932,” *The American Presidency Project*, accessed June 30, 2011, http://www.presidency.ucsb.edu/ws/index.php?pid=29638#axz z1N9yDnpSR.

Both parties were arguing for *cuts* in government expenditures, not the increases that (with the benefit of hindsight and better theory) we have suggested were needed. Monetary policy was likewise not used to stimulate the economy at this time. It seems unlikely that the fiscal and monetary authorities knew what to do but did nothing. Instead, the tools of economic thought needed to guide policy were simply not sufficiently well developed at the time. In keeping with the prevailing view that the economy was self-correcting, the incumbent Republican president, Herbert Hoover, had insisted that “prosperity is just around the corner.”

The election of Franklin Roosevelt in 1932 was a turning point. After his election, President Roosevelt and his advisors created a series of measures—called the New Deal—that were intended to stabilize the economy. In terms of fiscal policy, the US government moved away from budget balance and adopted a much more aggressive spending policy. Government spending increased from 3.2 percent of real GDP in 1932 to 9.3 percent of GDP by 1936. These spending increases were financed by budget deficits.

Roosevelt also took action to stabilize the banking system, most notably by creating a system of deposit insurance. This policy remains with us today: if you have deposits in a US bank, the federal government insures them. According to the Federal Deposit

Insurance Corporation (http://www.fdic.gov), not a single depositor has lost a cent since the introduction of deposit insurance.The FDIC site (http://www.fdic.gov) provides a discussion the history of this fund and current activities. The discussion of "Who is the FDIC?" (http://www.fdic.gov/about/learn/symbol/index.html) is a good place to start. Finally, the 1930s was also the time of the introduction of Social Security and other measures to protect workers. The Social Security Administration (http://www.ssa.gov) originated in 1935.General information on social security is available at http://www.ssa.gov. The history of the legislation, including various House and Senate Bills, is also available at http://www.ssa.gov/history/history.html. The original act included old-age benefits and the provision of unemployment insurance. The disability part of the program was created in 1956.

The New Deal brought about changes not only in policy but also in attitudes toward policymaking. Gardiner Means, who was an economic adviser to the Roosevelt administration in 1933, said of policymaking at the time:

It was this which produced the yeastiness of experimentation that made the New Deal what it was. A hundred years from now, when historians look back on this, they will say a big corner was turned. People agreed old things didn't work. What ran through the whole New Deal was finding a way to make things work.

Before that, Hoover would loan money to farmers to keep their mules alive, but wouldn't loan money to keep their children alive. This was perfectly right within the framework of classical thinking. If an individual couldn't get enough to eat, it was because he wasn't on the ball. It was his responsibility. The New Deal said: "Anybody who is unemployed isn't necessarily unemployed because he is shiftless." See Studs Terkel, *Hard Times: An Oral History of the Great Depression* (New York: Pantheon Books, 1970), 247.

Key Takeaways

- Stabilization policy entails the use the monetary and fiscal policy to keep the level of output at potential output.
- Monetary policy is the use of interest rates and other tools, under the control of a country's central bank, to stabilize the economy. During the Great Depression, monetary policy was not actively used to stabilize the economy. A major component of stabilization after 1932 was restoring confidence in the banking system.
- Fiscal policy is the use of taxes and government spending to stabilize the economy. During the first part of the 1930s, contractionary fiscal policy may have deepened the Great Depression. After 1932, fiscal policy became more expansionary and may have helped to end the Great Depression.

Exercises

1. Suppose the government wants to increase real GDP by $1,000. Explain why a smaller multiplier implies that the government must increase its spending by more to increase real GDP by this amount.
2. Did the government miss a chance to carry out stabilization policy before 1932?

This page titled 22.6: Policy Interventions and the Great Depression is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

22.7: End-of-Chapter Material 22.7: End-of-Chapter MaterialEdit Update titlecancel

In Conclusion

We started this chapter by describing the experience of the economy of the United States and other countries in the 1930s. The catastrophic economic performance of that period was difficult to reconcile with the view of classical economists that markets always worked to coordinate aggregate economic activity. Although technological progress provides a plausible explanation of the roaring twenties, technological regress is much less convincing as a story of the Great Depression. Technological regress also cannot explain the behavior of the price level and real wages during the Great Depression.

The Keynesian view explains the Great Depression as being driven by a decrease in aggregate spending, caused primarily by two factors: household consumption decreased because the stock market crash reduced household wealth, and investment decreased because of disruption of the financial intermediation process and pessimism over the future of the economy. These reductions in spending, through the multiplier, led to large reductions in real output. This story is consistent with the observed reductions in consumption, investment, and real GDP. With sticky prices, these reductions in spending translate into lower real GDP. The simple Keynesian story also has two problems: it can explain increasing prices only by assuming an exogenous increase in autonomous inflation and it provides no explanation of why observed technology decreased in the Great Depression period.

Along with the Keynesian explanation of the Great Depression comes a solution: use government policies to manage aggregate spending. If the aggregate expenditure model were literally true, policymaking would become an exact science: the policymaker would start with a target level of output and then determine the level of, say, government purchases needed to reach that target. As you might imagine, life as an economic policymaker is more complicated. The economists and politicians designing fiscal and monetary policy do not have a perfect picture of the current state of the economy. Moreover, control over policy tools is often inexact, and policy decisions take time.

The Great Depression remains something of a puzzle to macroeconomists. This became very apparent again recently during the so-called Great Recession—the major economic downturn that began in 2008. There are some resemblances between the two episodes, and the experience of the Great Depression certainly influenced some of the monetary policy decisions that were made in recent years. In this chapter, we did not yet consider monetary policy in detail. Chapter 25, which discusses the conduct of monetary policy, also addresses monetary policy during the Great Depression.

The aggregate expenditure framework is not a very sophisticated theory of the economy. Much work in macroeconomics in the decades since the Great Depression has involved refining the various pieces of the aggregate expenditure model. Economists have developed more rigorous theories of consumption, investment, and price adjustment, for example, in which they emphasize how households and firms base their decisions on expectations about the future. But Keynes' fundamental insight—that the level of output may sometimes be determined not by the productive capacity of the economy but by the overall level of spending—remains at the heart of macroeconomic research and policymaking today.

Key Links

- The history of deposit insurance: http://www.fdic.gov/about/history
- Social Security, including details on its history: http://www.ssa.gov and http://www.ssa.gov/history/history.html
- Bureau of Labor Statistics: http://www.bls.gov
- US Census Bureau, Statistical Abstracts: www.census.gov/prod/www/abs/statab.html
- Photo exhibits about the Great Depression:
  - Photo essay: http://www.english.uiuc.edu/maps/depression/photoessay.htm
  - Library of Congress: http://memory.loc.gov/ammem/fsowhome.html
- President Hoover's library: http://hoover.archives.gov

exercises

1. Consider the bank-run game presented in 22.4 Section "Can a Decrease in Investment Spending Explain the Great Depression?". Discuss in words how you think the introduction of deposit insurance would change the incentives of an individual to run on a bank.

2. If the marginal propensity to spend is 0.6 and autonomous spending decreases by $500, what is the change in output predicted by the aggregate expenditures model?
3. During the early 1930s, the government was intent on balancing its budget. If this required a reduction in government spending, what do you predict would happen to real GDP?
4. Do you think that labor force participation (that is, the percentage of the population that is actively in the labor force, either working or looking for a job) is procyclical or countercyclical? Why?
5. What is the effect of consumption smoothing on the value of the multiplier?
6. Explain why an increase in the value of the stock market might lead to higher real GDP. [Hint: think about what happens to consumption.]
7. Suppose you plan to meet a friend at a restaurant at 7 p.m. You are worried that she might be late and not show up until 8 p.m. You would prefer to eat at 7 p.m. rather than 8 p.m., but you also would prefer not to have to stand around waiting for your friend for an hour. She has the same tastes as you do. Explain carefully how you and your friend are in a coordination game. Is it an equilibrium for you both to show up at 7 p.m.? Is it an equilibrium for you both to show up at 8 p.m.?
8. Suppose that the inflation rate is very sensitive to the output gap in the economy. What does this imply about how quickly the economy will get back to equilibrium following a shock?

Spreadsheet Exercises

1. Using the data presented in Table 22.2.1 "Major Macroeconomic Variables, 1920–39*", create a spreadsheet to look graphically at the relationship between real GDP, unemployment, and the price level from 1929 to 1933.
2. Redo Table 22.3.1 "Growth Rates of Real GDP, Labor, Capital, and Technology, 1920–39*" assuming that $a = 0.3$ throughout the period of study. How do the results change?

Economics Detective

1. Consider the town in which you were born. Try to find out what happened there during the Great Depression. Did local businesses close? Were jobs available?
2. Can you find a recent example of a bank run in some country? What happened?
3. Following the financial crisis of 2008, the United States adopted a large fiscal stimulus. Try to find some details of this stimulus. How big was it? What form did it take? How big did policymakers think the multiplier was?

This page titled 22.7: End-of-Chapter Material is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

23: Jobs in the Macroeconomy

23.1: Taking to the Streets
23.2: Unemployment
23.3: Job and Worker Flows
23.4: Hours Worked
23.5: The Government and the Labor Market
23.6: End-of-Chapter Material

This page titled 23: Jobs in the Macroeconomy is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

23.1: Taking to the Streets 23.1: Taking to the StreetsEdit Update titlecancel

In March 2006, students demonstrated on the streets of France.

Violent French Protests: 300 Held

Police detained some 300 people around France after nationwide student marches against a new labor law turned violent, as street cleaners cleared away torched cars Friday and the government braced for more protests.

A quarter of a million people took to the streets in some 200 demonstrations around the country Thursday, in a test of strength between youth and the conservative government of 73-year-old President Jacques Chirac.

Most of the violence—and the arrests—were around the Sorbonne University in Paris, where police fired rubber pellets and tear gas at youths who pelted them with stones and set cars on fire.

[…]

Many trade unionists and students oppose the new youth employment law because it allows new workers under the age of 26 to be dismissed within a two-year trial period.“Violent French Protests: 300 Held,” *VOV News*, March 18, 2006, accessed August 22, 2011, english.vov.vn/Home/Violent-French-protests-300-held/20063/36835.vov.

If, like most readers of this book, you are a student in the United States, it is unlikely that you have taken part in violent demonstrations about labor policy. It is not that such demonstrations are unheard of. In Madison, Wisconsin, in 2011, there were extended protests concerning proposed changes in public sector contracts. Still, in the United States, it is accepted that the government has a limited influence on contracts between workers and firms. It is part of economic life in the United States that employment is not protected by the government. In Europe, however, many countries have extensive laws on their books that are designed to protect workers. For example, in much of Europe, unemployment insurance is more generous than in the United States. Unemployed people obtain larger benefits and are eligible for these benefits for longer periods of time.

In many European countries, it is also much more difficult to fire workers than it is in the United States. The proposed new job contract that led to the demonstrations in France was intended to reduce the nearly 25 percent unemployment rate of the French youth. Perhaps paradoxically, the contract was designed to make it easier to make young people unemployed. The logic was that firms would be willing to hire more workers if the costs of firing them were lower.

The different systems in the United States and Europe each have their defenders. Supporters of European labor laws point to the greater job security enjoyed by workers in Europe. Supporters of the US system argue that the United States enjoys greater flexibility in the labor market, leading to a more efficient economy with less unemployment. Some feel that the United States should adopt European-style labor protection measures; others feel that Europe would benefit from becoming more like the United States.

In this chapter, we look at the different experiences of Europe and the United States in order to evaluate these different approaches to the labor market. In the end, we want to be able to answer—or at least form intelligent opinions about—the following question:

What are the results of the different labor market policies in the United States and Europe?

This is not just an academic question for discussion in a textbook. In both the United States and Europe, labor market policy is frequently debated. The US Congress has considered various labor policies, such as restrictions on plant closing to protect jobs, requirements that firms offer workers health insurance, requirements that firms include paid sick days in employment contracts, and so on. At the same time, there is considerable discussion in Germany, France, and elsewhere in Europe about the possible benefits of increased labor market flexibility.

Road Map

Employment and unemployment are ideas that most of us are familiar with. You may well have already been employed, at least in a part-time capacity, at some point in your life. It is also possible that you have been unemployed, meaning that you were without a job, but were actively seeking work. Our personal experiences, and those of our parents and friends, help us understand the basics of employment and unemployment.

Even if you have not yet been employed, you will begin searching for a job once you graduate with a college degree. As you surely know, finding a good job is not always easy. You want to find a job that you enjoy, fits your skills, and pays well. It is also not easy for prospective employers: they want to find someone who is suitably skilled, will work well within the firm, and is not too expensive. The challenge is to match workers and jobs: the worker needs to be suited to the job, and the job needs to be suited to the worker.

The process of matching does not happen just once. As time passes, your skills, ambitions, and choice of occupation may change. As time passes, your employer's needs change. You may wish to move to another city. Your employer may want to move your job to another city. Most people do not spend their entire lives in one job.

A schematic representation of this process is shown in Figure 23.1.1. Here you leave college and look for a job. Finding that job is likely to be time-consuming. You will have to contact lots of prospective employers, read newspaper ads, use search engines on the Internet, and, of course, show up for interviews. In the end, you will find your first job and begin your career.

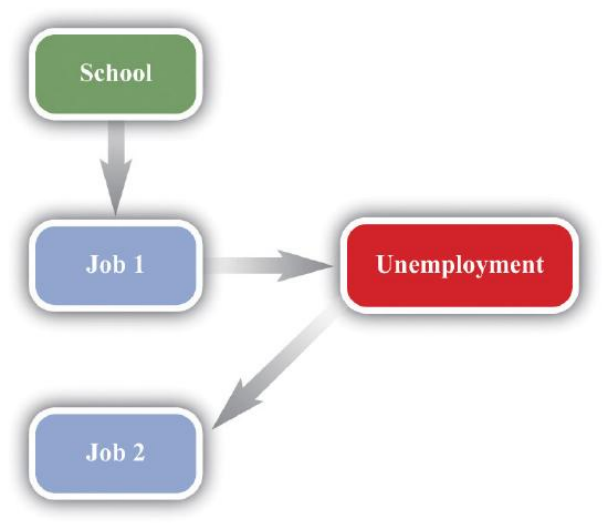


Figure 23.1.1: Employment Transitions over Your Lifetime

You might stick with this job for a while, but in all likelihood the match between you and your employer will come to an end sooner or later. You may leave the job through your own choice because you are no longer happy with it. Alternatively, you may be forced to leave because your employer no longer has need of you. You then search for another job. When you succeed in finding a new position where your needs and desires align with those of another employer, a new match is formed.

Fifty years or so ago, people often joined companies and stayed with them for life, but this is very unusual today. You are likely to move between jobs several times during your lifetime before your eventual retirement. Sometimes you may be able to move from one job to another without interruption. At other times you will be unemployed between jobs. Throughout your life, you are likely to face periods of anxiety and stress because of the employment uncertainties that you confront:

- How likely is it that you will be able to retain your current job?
- If you lose your job, will you be able to find another job that you like?
- How long will it take to find another job?
- What should you do if you do not like your job?
- How will you support yourself while you are unemployed?

This discussion makes it clear that we cannot analyze labor market policies without understanding the movements in and out of employment and unemployment. But before we can do so, we need to make sure we understand exactly what unemployment is, and what causes it. Thus we begin by carefully defining unemployment.In part, this is a review of material in Chapter 18. There, we explained that the unemployment rate is one possible indicator of the overall health of the economy. We look at the data for Europe and the United States and make sense of this data using economic reasoning. We then turn to an analysis of the matching between workers and jobs and the decisions of individual workers in this process. All this analysis gives us a better understanding of unemployment and, more generally, the operation of labor markets. We conclude by evaluating labor market policies in the United States and Europe.

This page titled 23.1: Taking to the Streets is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

23.2: Unemployment 23.2: UnemploymentEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What is the unemployment rate, and how is it measured?
2. What are the differences and the similarities in unemployment rates in the United States and Europe?
3. If the labor market functions perfectly, what is the rate of unemployment?
4. How does unemployment arise?

We begin by discussing the most commonly watched indicator of the state of the labor market: the unemployment rate. In the United States, the unemployment rate is measured by the Bureau of Labor Statistics (BLS; http://www.bls.gov/cps/home.htm). The BLS looks at the population of individuals of working age who are not in the military. It sorts such people into three separate categories:

1. **Employed.** Individuals with a job, either full time or part time
2. **Unemployed.** Individuals who do not currently have a job but are searching for employment
3. **Out of the labor force.** Individuals who are not employed and not looking for work

Thus

$$civilian\ working\ age\ population\ =\ number\ employed\ +number\ unemployed\ +number\ out\ of\ the\ labor\ force.$$

Those out of the labor force include students, stay-at-home parents, those who are prevented from working by disability, and people who have taken early retirement. The category also includes **discouraged workers**, those who are deemed to have dropped out of the labor force because they have stopped looking for a job.

The **civilian labor force** comprises the employed and the unemployed. The **unemployment rate** is calculated as follows:

and the **employment rate** is calculated as follows:

In the United States, the definition of “employed” is fairly liberal. To be classified as employed, it is sufficient to have done *any* work for pay or profit in the previous week. People may even be counted as employed if they did not work during the week—for example, if they were on vacation, out sick, on maternity/paternity leave, or unable to work because of bad weather.

In this chapter, we explore differences in unemployment in the United States and Europe. To do this properly, we need to take care that unemployment is measured in a similar way within the sample of countries. The European Commission defines as unemployed those aged 15 to 74

- who were without work during the reference week, but currently available for work,
- who were either actively seeking work in the past four weeks or who had already found a job to start within the next three months.“Employment and Unemployment (LFS): Definitions,” European Commission, accessed July 31, 2011, epp.eurostat.ec.europa.eu/portal/page/portal/employment_unemployment_lfs/methodology/definitions.

As in the United States, the unemployment rate is the number of people unemployed as a percentage of the labor force, and the labor force is the total number of people employed and unemployed.

The European Commission defines as employed those aged 15 to 74

- who during the reference week performed work, even for just one hour a week, for pay, profit or family gain,
- were not at work but had a job or business from which they were temporarily absent because of, e.g., illness, holidays, industrial dispute or education and training.“Employment and Unemployment (LFS): Definitions,” European Commission, accessed July 31, 2011, epp.eurostat.ec.europa.eu/portal/page/portal/employment_unemployment_lfs/methodology/definitions.

These descriptions reveal that the definitions used in Europe are broadly similar to those in the United States, meaning that we can legitimately compare employment and unemployment rates in the two regions.

National and local governments help people cope with the risk that they might lose their jobs. In the United States and many other countries, unemployed people are typically eligible to receive payments from the government, called **unemployment insurance**,

for some period of time after losing their jobs. Some governments help the unemployed find jobs and may even provide financial support to help people retrain and obtain marketable skills.

Unemployment in the United States and Europe

Figure 23.2.1 shows quarterly unemployment rates for the United States, France, and Europe as a whole. In the late 1980s, unemployment fell in both the United States and France, although the US unemployment rate was about two percentage points lower than the French rate. The 1990s were a different story. Unemployment rates increased in both countries at the beginning of the decade. Thereafter, the unemployment rate decreased in the United States, but it continued to increase in France for about half of the decade and decreased only near the end of century. From the early 1990s up to about 2008, the unemployment rate in Europe was substantially higher than that in the United States. The pattern for Europe as a whole closely matches the pattern for France, although unemployment in France is typically a little higher than the European average.

The crisis of 2008, however, led to a dramatic rise in the unemployment rate in the United States. At the end of 2007, the US unemployment rate was just under 5 percent. Two years later, at the start of 2010, the rate was over 10 percent. Unemployment also rose in Europe, but to nothing like the same degree. In early 2011, US and European unemployment rates were almost identical.

One other feature of the data is noticeable: there is a regular *seasonal pattern* in the data. For example, in the United States, unemployment is almost always higher in the first quarter of the year than it is in the preceding or following quarter. This is because some sectors of the economy are heavily affected by seasonal patterns. For example, stores may hire extra people during the Christmas holiday period, while construction firms may employ fewer people during the winter months. Sometimes, data such as these are "seasonally adjusted" to remove these effects.

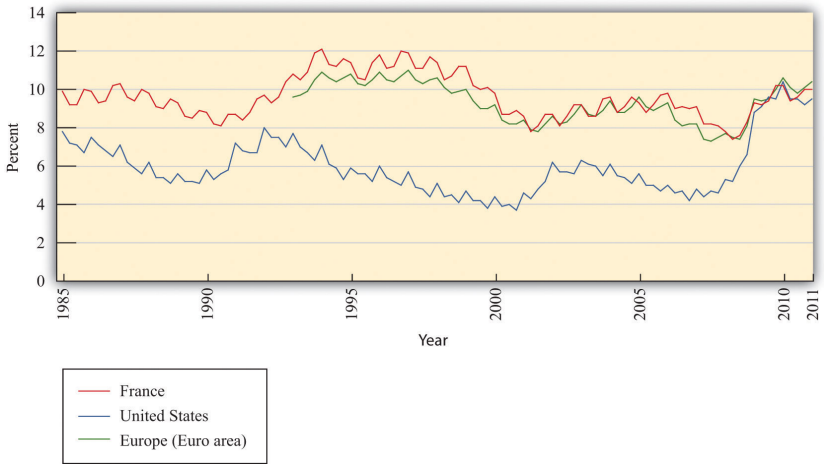


Figure 23.2.1: *Unemployment Rates in France, the United States, and the Euro Area, 1985-2011*. Source: OECD, "Statistics Portal: Labour,"

www.oecd.org/topicstatsportal/0,2647,en_2825_495670_1_1_1_1_1,00.html#499797.

The French labor law reforms with which we began the chapter were aimed at young workers, so let us also look specifically at the unemployment experience of this group. Between 2000 and 2010, the unemployment rate in France for the age group 20–24 ranged between 17 and 21 percent, with an average of 18.6 percent. In the United States, in contrast, for the same period and the same group of workers, the unemployment rate averaged 10 percent. The figures on youth unemployment come from "Statistics Portal: Labour," OECD, www.oecd.org/topicstatsportal/0,2647,en_2825_495670_1_1_1_1_1,00.html#499797. In both countries, the unemployment rate is higher for younger workers than the overall unemployment rate.

Although there are some similarities between France and the United States, there is also a clear puzzle: unemployment, for both the overall population and young workers, was, until very recently, much higher in France. We need to understand the source of this difference before we can evaluate different policy remedies.

The Labor Market

Unemployment suggests a mismatch between supply and demand. People who are unemployed want to have a job but are unable to find one. In economic language, they are willing to supply labor but cannot find a firm that demands their labor. The most natural starting point for an economic analysis of unemployment is therefore the **labor market**.

Toolkit: Section 31.3

The labor market brings together the supply of labor by households and the demand for labor by firms. You can review the labor market in the toolkit.

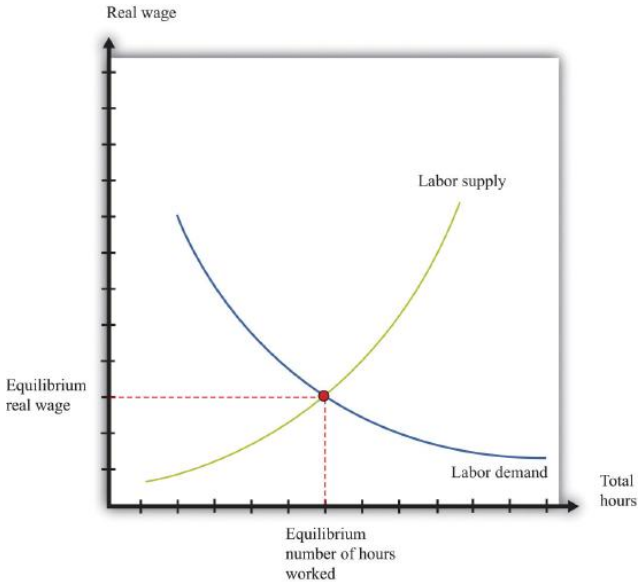


Figure 23.2.2: Labor Market

The labor market is depicted in Figure 23.2.2. "Price" on the vertical axis is the **real wage**, which is the nominal wage divided by the price level. It tells us how much you can obtain in terms of real goods and services if you sell an hour of your time. Recalling that the price level can be thought of as the price of a unit of the real gross domestic product (real GDP), you can equivalently think of the real wage as the value of your time measured in units of real GDP.

At a higher real wage, households supply more labor. There are two reasons for this. First, a higher real wage means that, for the sacrifice of an hour of time, households can obtain more goods and services than before. Households are therefore induced to substitute away from leisure to work and ultimately consume more. Second, as the wage increases, more individuals join the labor force and find a job. Embedded in the upward-sloping labor supply curve is both an increase in hours worked by each employed worker and an increase in the number of employed workers.

At a higher real wage, firms demand fewer labor hours. A higher real wage means that labor time is more expensive than before, so each individual firm demands less labor and produces less output. The point where the labor supply and demand curves meet is the equilibrium in the labor market. At the equilibrium real wage, the number of hours that workers choose to work exactly matches the number of hours that firms choose to hire.

Supply and demand in the labor market determine the real wage and the level of employment. Variations in either labor supply or labor demand show up as shifts in the curves. If we want to talk about unemployment, however, the labor market diagram presents us with a problem. The idea of a market is that the price adjusts to reach **equilibrium**—the point where supply equals demand. In the labor market, this means the real wage should adjust to its equilibrium value so that there is no mismatch of supply and demand. Everyone who wants to supply labor at the equilibrium wage finds that their labor is demanded—in other words, everyone who is looking for a job is able to find one.

Remember the definition of unemployment: it is people who are not working *but who are looking for a job*. The supply-and-demand framework has the implication that there should be no unemployment at all. Everyone who wants to work is employed; the only people without jobs are those who do not want to work.

Theories of Unemployment

So where do we go from here? One natural approach is to start from Figure 23.2.2 but look for circumstances in which we would see unemployment. Figure 23.2.3 shows us that there will be unemployment if the real wage in the market is too high—that is,

above the equilibrium real wage. In this case, the amount of labor that workers want to sell is greater than the amount that firms want to buy. Some workers will want a job at this wage but be unable to find one. They will be unemployed.

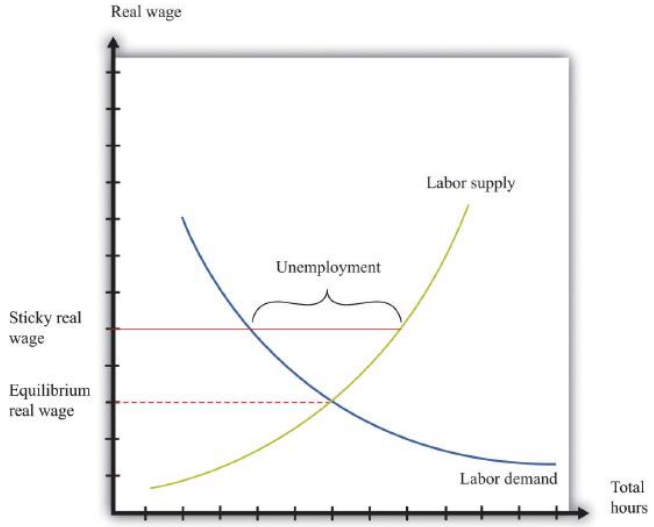


Figure 23.2.3: Unemployment in the Labor Market. If the real wage is sticky, it may be higher than the equilibrium real wage, meaning that some workers who want to work are unable to find a job.

Figure 23.2.3 shows us what the labor market must look like for there to be unemployment, but it is hardly an explanation of unemployment. Economists typically expect markets to look like Figure 23.2.2, not Figure 23.2.3. That is, they think that the price in a market—in this case, the real wage—adjusts quickly to ensure that supply equals demand. If we want to explain unemployment with a picture like Figure 23.2.3, we also need some story of why real wages might be sticky, so they remain above the equilibrium wage.

Inflexible Real Wages

Over the years, economists have offered several stories about why wages might be inflexible.

- One story is that the wage is not allowed to decrease by law. Many economies have minimum wage laws on their books. This could explain some unemployment. A difficulty with this explanation is that the minimum wage affects only low-income workers. Most workers in the economy actually earn a wage above the legal minimum and are unaffected by minimum-wage legislation.
- Another possibility is that firms find it difficult to adjust wages downward. The market for people's time is not like the market for bread. Pay cuts are very visible to workers and are likely to meet a great deal of resistance. If a firm tries to cut wages, it is likely to find that its workers become demotivated and that its best workers start looking for jobs at other firms.

Both of these stories are really explanations of why **nominal wages** may be unable to adjust. Figure 23.2.3 has the real wage on the axis. Remember that the real wage is calculated as follows:

Minimum wage laws specify a fixed minimum nominal wage. Even if the nominal wage is fixed, the real wage decreases when the **price level** increases. It follows that rigidities in the nominal wage translate into rigidities in the real wage only if the price level is also sticky.

Prices in an economy may indeed be sticky in the short run, so sticky wages and prices do provide one explanation for short periods of unemployment. Such unemployment is sometimes called **cyclical unemployment**. In the long run, however, we would expect the labor market to return to an equilibrium with zero unemployment. Cyclical unemployment is the component of unemployment that depends on the business cycle. During a recession, cyclical unemployment is relatively high. In periods of economic expansion, cyclical unemployment is low or nonexistent. But we *always* observe some unemployment, which tells us that sticky nominal wages and prices cannot be the whole story.

Figure 23.2.3 tells us that the only way to get persistent unemployment in this framework is for the *real* wage to be permanently above the equilibrium wage. We need to find some reason why market forces will not cause the real wage to adjust to the point where demand equals supply.

One possible story introduces labor unions into the picture. Unions give some market power to workers. Just as we sometimes think about firms having market power, meaning that they have some control over the prices that they set, so we can think about a union having some control over the wage that workers are paid. If there were just a single union representing all workers, then it could choose the real wage, much as monopoly firms choose their price. Firms would then hire as many hours as they wanted at that wage. Generally, unionized workers are paid more than the wage at which supply equals demand, just as in Figure 23.2.3. The union accepts some unemployment but believes that the higher wage more than compensates. A problem with this story is that, like the minimum wage, it is relevant only for a relatively small number of workers. In the United States in particular, only a small fraction of the workforce is unionized.

Another story goes by the name of **efficiency wages**. The idea here is that firms have an incentive to pay a wage above the equilibrium. Workers who are paid higher wages may feel better about their jobs and be more motivated to work hard. Firms may also find it easier to recruit good workers when they pay well and find it easier to keep the workers that they already have. The extra productivity and lower hiring and firing costs may more than compensate the firm for the higher wage that it is paying.

Inside the Labor Market

So far, we have come up with four possible stories about unemployment. Can these theories help to explain differences between Europe and the United States?

First, it is generally the case that minimum wages are more generous in Europe than in the United States, so it is certainly possible that higher minimum wages in Europe contribute to higher levels of unemployment there. Second, there is some evidence that nominal wages are in some sense "stickier" in Europe than in the United States. Third, we can observe that unions are generally more prevalent and more powerful in Europe than in the United States. Thus some of the stories that we have told are potentially helpful in explaining differences between the United States and Europe.

However, all these theories are silent about the underlying movement of workers from employment to unemployment and back again. Figure 23.2.3 paints a static picture of a world that is in fact dynamic and fluid. There is no means in the framework to explore the role of unemployment insurance and other policies that differ across Europe and the United States. In addition, market forces may work differently in the labor market. In Figure 23.2.3, there are more workers wanting to work than there are jobs offered by firms. The standard story of market adjustment is that workers willing to work for a lower wage would approach a firm, offer to undercut the wage of an existing worker, and be immediately hired as a replacement. This is not how hiring and firing usually works in the labor market. Firms have a relationship with their existing workers; they know if their workers are competent, hardworking, and reliable. Firms will not readily replace them with unknown quantities, even for a lower wage.

For these reasons, researchers in labor economics think that Figure 23.2.3 is too simple a framework to explain the realities of modern labor markets. Instead, they frequently turn to a different framework more suited to thinking about labor market flows.

Key Takeaways

- The unemployment rate is the fraction of the civilian labor force looking for a job but currently not employed. The BLS in the United States produces this number on a monthly basis.
- During the early part of the 1980s, the unemployment experiences in the United States and Europe were similar. Up until 2008, the unemployment rate in Europe had been significantly higher than the unemployment rate in the United States. Very recently, however, the US unemployment rate climbed to European levels.
- In a perfectly functioning labor market, the unemployment rate would be zero.
- Possible explanations of unemployment include rigidities in wages, the market power of unions, and incentive effects.

check your understanding

1. Explain in your own words why the standard supply-and-demand framework predicts zero unemployment when it is applied to the labor market.
2. What wage is determined in labor market equilibrium—the real wage or the nominal wage?

This page titled 23.2: Unemployment is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

23.3: Job and Worker Flows 23.3: Job and Worker FlowsEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What key features of labor markets does the static model of labor supply and labor demand fail to capture?
2. What are some of the key facts about worker labor market flows?
3. What is search theory, and how is it useful for understanding labor market outcomes?
4. What are the efficiency gains from flexible labor markets?

The labor market is a highly dynamic place. Workers are constantly moving from job to job, in and out of the workforce, or from employment to unemployment and vice versa. Large firms devote substantial resources to human resource management in general and hiring and firing in particular. By contrast, Figure 23.3.3 is *static* because it shows the labor market at a moment in time. Our understanding of the labor market—and, by extension, employment and unemployment—is badly incomplete unless we look more carefully at the movement of workers. Further, when workers and firms meet, they do not take as given a market wage but instead typically engage in some form of bargaining over the terms of employment.

This vision of a dynamic labor market with bargaining is much closer to the reality of labor relations than is the model of labor supply and demand. To better understand the determinants of employment and unemployment, we therefore turn to labor market flows. We begin with some more facts, again contrasting the experience of Europe with that of the United States, and then develop a framework that allows us to think explicitly about the dynamic labor market.

Facts

Our starting point is the classification of individuals in the civilian working age population. Recall that economic statistics place them as one of the following: employed, unemployed, or not in the labor force. Imagine taking a snapshot of the US economy each month. For a given month, you would be able to count the number of people employed, unemployed, and out of the labor force. We could call these the *stocks* of each kind of individual.

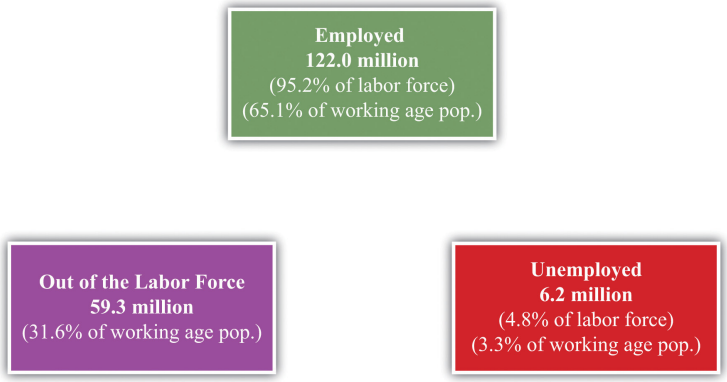


Figure 23.3.1: Worker Stocks in the United States

Figure 23.3.1 shows the number of people between 16 and 64 years old in the United States in three different "states"—employment, unemployment, and out of the labor force—over the period 1996–2003.These data come from a study using a monthly survey conducted by the Bureau of Labor Statistics (BLS) called the Current Population Survey and were compiled by Stephen J. Davis, R. Jason Faberman, and John Haltiwanger. The numbers here come from S. Davis, R. J. Faberman, and J. Haltiwanger, "The Flow Approach to Labor Market: New Data Sources and Micro-Macro Links" NBER Working Paper #12167, April 2006, accessed June 30, 2011, http://www.nber.org/papers/w12167. On average, there were 122 million people employed, 6.2 million unemployed, and 59.3 million considered out of the labor force. Adding these numbers together, there were 187.5 million working-age individuals, of whom 128.2 million were in the labor force. The average unemployment rate was 4.8 percent over this period, and the employment rate was 95.2 percent. Notice, though, that many individuals are out of the labor force: only 65 percent of the population is employed.

Figure 23.3.1 shows an average over many months, but you could also look at how these numbers change from month to month. Even more informatively, you could count the number of people who were employed in two consecutive months. This would tell you the likelihood of being employed two months in a row. These calculations for the US economy are summarized in Figure 23.3.2.

Look, for example, at the arrows associated with the box labeled *unemployed*. There are two arrows coming in: one from the employed box and one from the out-of-the-labor-force box. There are two arrows going out: one to the employed box and one to the out-of-the-labor-force box. Each of these four arrows has a percentage attached, indicating the fraction of people going from one box to another. Thus, on average, 28.3 percent of the unemployed people in one month are employed in the next and 23.3 percent leave the labor force. The remaining 48.4 percent stay in the group of unemployed.

The numbers in the figure are averages over a long period. Such flows change over the course of the year due to seasonal effects. Around Christmas, for example, it may be easier for an unemployed worker to find a job selling merchandise in a retail shop. These flows also change depending on the ups and downs of the aggregate economy.

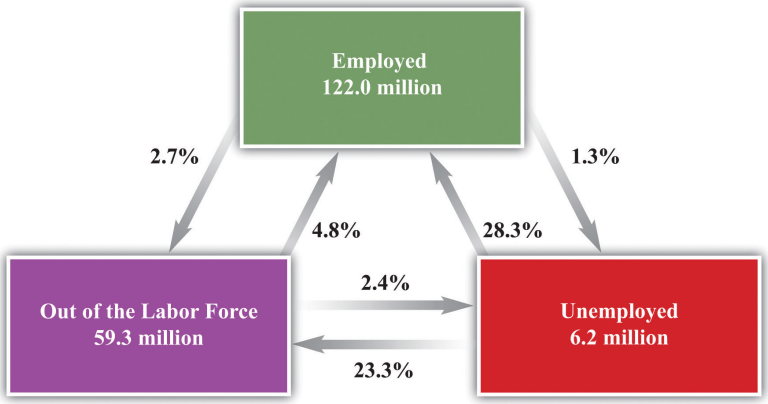


Figure 23.3.2: Worker Flows in the United States

Do European countries exhibit similar patterns? Portugal makes for a good comparison with the United States because the unemployment rates in the two countries were broadly similar over most of the last two decades. Yet Portugal has very strong employment protection laws, to the point where they are enshrined in the Portuguese Constitution: "Article 53," *Portugal-Constitution*, adopted April 2, 1976, accessed June 30, 2011, http://www.servat.unibe.ch/icl/po00000_.html#A053_.

Article 53 Job Security

The right of workers to job security is safeguarded. Dismissals without just cause or for political or ideological reasons are forbidden.

A study that compared the labor markets in Portugal and the United States uncovered the following facts: See Olivier Blanchard and Pedro Portugal, "What Hides Behind an Unemployment Rate: Comparing Portuguese and U.S. Labor Markets," *American Economic Review* 91, no. 1, (2001), 187–207.

- The flows into unemployment from employment and the flows from employment to unemployment are much lower in Portugal compared to the United States.
- Average unemployment duration in Portugal is about three times that of the United States.
- Job protection is very high in Portugal relative to the United States.

Even though Portugal and the United States have similar overall unemployment rates, the underlying flows are quite different in the two countries. Flows between employment and unemployment—and vice versa—are much smaller in Portugal. This means that if you lose your job, it is likely to take a long time to find a new one. If you have a job, you are likely to keep it for a long time. As we would expect from this, people typically spend much longer periods of time in unemployment in Portugal than they do in the United States.

If we compare the United States with Europe more generally, we see similar patterns. In 2010, the average **unemployment duration** for workers ages 15–24 was about 10.6 months in Europe but only 5.9 months for the United States. For workers in the 25–54 age group, the duration was higher in both Europe (13.7 months) and the United States (8.2 months) than for younger

workers. See "Unemployment Duration," Online OECD Employment database, accessed June 30, 2011, http://www.oecd.org/document/34/0,3746,en_2649_33927_40917154_1_1_1_1,00.html#uduration. Recall that in 2010, Europe and the United States had similar rates of unemployment. Employment duration, however, is still much higher in Europe than the United States. In both places, older workers tend to be unemployed for longer periods than younger workers. But European workers are typically unemployed for much longer periods of time than US workers. These figures come from "Average Duration of Unemployment," OECD, accessed June 30, 2011, stats.oecd.org/Index.aspx?DataSetCode=AVD_DUR.

The Organisation for Economic Co-operation and Development (OECD) conducted a large study on the employment protection legislation in a variety of developed countries. The main study (OECD Employment Outlook for 2004, http://www.oecd.org/document/62/0,3746,en_2649_33927_31935102 _1_1_1_1,00.html) created a measure of employment protection and then attempted to relate it to labor market outcomes in different countries. The reasoning we have just presented suggests that in countries with relatively high levels of employment protection, labor markets would be much more sluggish.

Formulating a comprehensive measure of employment protection is not easy. In principle, the idea is to measure the costs of firing workers and various regulations of employment. Examples would include requirements on advance notice of layoffs and the size of severance payments that firms are obliged to pay. In some countries, a firm must go to court to lay off workers. For temporary workers, there are specific restrictions placed on this form of contract, as in the discussion of France that opened this chapter. In reality, these costs are difficult to detect and convert to a single measure. The OECD findings should be interpreted with these challenges in mind.

Another OECD publication (http://www.oecd.org/dataoecd/40/56/36014946.pdf) examines employment protection legislation across OECD countries in 1998 and 2003.This discussion is based on Figure A.6 of OECD, "Annex A: Structural Policy Indicators," *Economic Policy Reforms: Going for Growth*, accessed June 30, 2011, http://www.oecd.org/dataoecd/40/56/36014946.pdf. Portugal was the country with the highest level of employment protection legislation, while the United States was the lowest. France was above average, while the United Kingdom and Canada were below average. The OECD analysis highlighted two effects of such legislation on labor market flows:

1. It limits flows from employment into unemployment because it is costly to fire workers.
2. It limits flows from unemployment to employment because firms, when deciding to hire a worker, will realize that they may wish to fire that worker sometime in the future.

The first effect is the more obvious one; indeed, it provides the rationale for employment protection. If it is hard to fire workers, then firms are less likely to do so. The second effect is less obvious and more pernicious. If it is hard to *fire* workers, then firms become more reluctant to *hire* workers. Put yourself in the place of a manager wondering whether to make a hire. One concern is that the person you are considering will turn out to be unsuitable, or a bad worker. Another is that conditions in your industry will worsen, so you may not need as many employees. In those circumstances, you want to be able to let the worker go. If you will not be able to do so, you may decide it is safer simply to make do with the workers you already have.

The OECD analysis particularly stressed the effects on the labor market experience of relatively young workers. The report emphasized that stronger legislation is linked to lower employment of young workers. If it is costly to sever a relationship, then a firm will not give a young worker a chance in a new job. The OECD also noted an important *benefit* of employment protection legislation: it enhances the willingness of young workers to invest in skills that are productive at their firms. Without a strong attachment to the firm, workers have little incentive to build up skills that are not transferable to other jobs.

Job Creation and Job Destruction

In place of the supply-and-demand diagram, we can think about the decisions that workers and firms make when they are trying to form or break an employment relationship. Individual workers search for available jobs, which are called vacancies. On the other side, vacancies are searching for workers. When a vacancy and a worker are successfully matched, a job is created. When we say that a vacancy is searching for a worker, we, of course, really mean that a firm with a vacancy is seeking to hire a worker. You can think of a firm as being a collection of jobs and vacancies.

Whereas the standard supply-and-demand picture downplays differences among workers and jobs, this "search-and-matching" approach places these differences at the center of the analysis. Workers differ in terms of their abilities and preferences. Jobs differ in terms of their characteristics and requirements. For an economy to function well, we need to somehow do a good job of matching vacancies with workers. When a successful match occurs, we call this "job creation."

Search theory is a framework for understanding this matching process. Let us think about how this process looks, first from the perspective of the worker and then from the perspective of the firm. Workers care about the various characteristics of their jobs. These characteristics might include how much the job pays, whether it is in a good location, whether it offers good opportunities for advancement, whether it is interesting, whether it is dangerous, and other attributes.

Vacancies are likewise "looking" for certain characteristics of workers, such as how much they cost, what skills they possess, whether they have relevant experience, whether they are hardworking and motivated, whether they are trustworthy, and so on. The firm cares about these characteristics because it cares about profitability: its goal is to make as much profit as possible.

Over time, the quality of the match between a worker and a vacancy may change. A job may become less profitable to the firm and/or less attractive to the worker. To put it another way, the amount of *value* created by the job may change. The worker may come to dislike particular aspects of the job or may wish to change location for family reasons. The worker may feel that he or she would be better matched with some other firm, perhaps because of changes in his or her skills and experience. From the firm's side, demand for the firm's product may decrease, or the firm might shift to a new production technique that requires different skills. If the value created by a job decreases too much, then the firm or the worker may choose to end the relationship, either by the worker's choice (quitting the job) or the firm's (firing the worker). This is "job destruction."

Jobs are created and destroyed all the time in the economy. The flows of workers among jobs and employment states are a key characteristic of the labor market. As these flows occur, workers often spend time unemployed. After a job is destroyed, the worker may spend some time unemployed until he or she finds a job with a different firm.

Labor Flows and Productivity

In a rapidly changing economy, the value of different jobs (worker-firm matches) changes over time. To function efficiently, the labor market needs to be able to accommodate such changes. For this discussion, we will think about efficiency as simply being measured by the productivity of the match between workers and firms. In an efficient match, the worker is productive at the chosen job. For the overall economy, if all matches are efficient, then it is not possible to change the assignment of workers to jobs and produce more output.

Comparative and Absolute Advantage

Let us see how this works in a simple example. Table 23.3.1 "Output Level per Day in Different Jobs" gives an example of an economy with two workers and two jobs. Each entry in the table is the amount of output that a particular worker can produce in each job in one day. For example, worker B can produce 4 units of output in job 2 and 8 units of output in job 1.

Table 23.3.1: Output Level per Day in Different Jobs

Worker	Job 1	Job 2
A	9	6
B	8	4

Before we begin, let us pause for a moment to think about this kind of example. This chapter is motivated by the desire to explain the employment and unemployment experiences of hundreds of millions of workers in the United States and Europe. It may seem ridiculous to think that a story like this—with two workers, two jobs, and some made-up numbers—can tell us anything about employment and unemployment across two continents. Economists often refer to such stories as "toy" models, in explicit recognition of their simplicity. This kind of model is not designed to tell us anything *specific* about US or European unemployment. The point of this kind of model is to keep our thinking clear. If we cannot understand the workings of a story like this, then we cannot hope to understand the infinitely more complicated real world. At the same time, if we do understand this story, then we begin to get a feel for the forces that operate in the real world.

If we were in charge of this economy, how would we allocate the workers across the jobs? In this case, the answer is easy to determine. If we assign worker A to job 1 and worker B to job 2, then the economy will produce 13 units of output per day. If we assign worker A to job 2 and worker B to job 1, then the economy will produce 14 units of output per day. This is the better option because—in the interest of efficiency—we would like the workers to be assigned to the jobs they do best.

Notice, by the way, that worker A is better than worker B at both jobs. However, worker A is a lot better at job 2 (50 percent more productive) and only a little better at job 1 (12.5 percent more productive). The best assignment of workers is an application of the

idea called comparative advantage: each worker does the job at which he or she does best *when compared to the other person*.

Comparative advantage and **absolute advantage** are used to compare the productivity of people (countries) in the production of a good or a service. We introduce this tool here assuming there are two people and two goods that they can each produce.

Toolkit: Section 31.13

A person has an absolute advantage in the production of a good if that person can produce more of that good in a unit of time than another person can. A person has a comparative advantage in the production of one good if the opportunity cost, measured by the lost output of the other good, is lower for that person than for another.

In our example, worker A has a comparative advantage in job 2, and worker B has a comparative advantage in job 1. We have defined comparative advantage in terms of **opportunity cost**, so let us go through this carefully and make sure it is clear. The opportunity cost of assigning a worker to one job is the amount of output the worker could have produced in the other job.

We can measure opportunity cost in terms of the output lost from assigning a worker to job 2 instead of job 1. The opportunity cost of assigning worker A to job 2 rather than job 1 is 3 units (9 − 6). The opportunity cost of assigning worker B to job 2 rather than job 1 is 4 units of output (8 − 4). The opportunity cost is higher for worker B, which is another way of saying that worker B has a comparative advantage in job 1. Worker B should be assigned to job 1, and worker A should take on job 2.

We could equally have measured opportunity cost the other way around: as the output lost from assigning a worker to job 1 rather than job 2. The opportunity cost of assigning worker A to job 1 rather than job 2 is −3 units (6 − 9). The opportunity cost of assigning worker A to job 1 rather than job 2 is less, it is −4 units of output (4 − 8). Worker A has the higher opportunity cost (−3 is greater than −4), so we again conclude that worker A should be assigned to job 2.

Changes in Productivity

Suppose that this simple economy is indeed operating efficiently, with worker A in job 2 and worker B in job 1. Then imagine that the productivity of one of these matches changes. For example, suppose that at some point worker B goes on a training course for job 2, so Table 23.3.1 "Output Level per Day in Different Jobs" becomes Table 23.3.2 "Revised Output Level per Hour from Assigning Jobs".

Table 23.3.2: Revised Output Level per Hour from Assigning Jobs

Worker	Job 1	Job 2
A	9	6
B	8	7

If you compare these two tables, you can see that worker B is now more productive than worker A in job 2. Worker A is still better at job 1, as before.

If we want to produce the maximum amount of output in this economy, we now want to switch the workers around: if worker A does job 1 and worker B does job 2, then the economy can produce 16 units of output per day instead of 14.

How might this change happen in practice? Here are three scenarios.

1. **Instantaneous reallocation.** In this case, the labor market is very fluid. Workers A and B trade places as soon as B becomes more productive. No one is unemployed, and real gross domestic product (real GDP) increases immediately.
2. **Stagnant labor market.** This scenario is the opposite of the first. Here, there is no reallocation at all. People are stuck in their jobs forever. In this case, worker B remains assigned to job 1, and worker A remains assigned to job 2. Although this was the best assignment of jobs when Table 23.3.1 "Output Level per Day in Different Jobs" described the economy, it is not the best assignment for Table 23.3.2 "Revised Output Level per Hour from Assigning Jobs". Relative to the better assignment, the economy loses 2 units of GDP every day.
3. **Frictional unemployment.** This scenario lies between these two extremes: workers and firms adjust but not instantaneously. How might workers A and B exchange jobs? One possibility is that worker A is fired from job 2 because the firm wants to attract worker B to the job instead. At the same time, worker B might quit in the hope of getting job 1 when it is vacant. Both workers move from employment into unemployment, as in the arrow from employment to unemployment in Figure 23.3.2.

During the time when workers A and B are unemployed, their production is reduced to zero. So, during the period of adjustment, the economy in the third scenario undergoes a recession. But once adjustments are made, the economy is much more productive than before. Economists refer to the unemployment that occurs when workers are moving between jobs as **frictional unemployment**.

How do these three scenarios compare? It is evident that fluid labor markets are the ideal scenario. In this situation, there is no lost output due to unemployment, and the economy is always operating in the most efficient manner. The choice between the second and third scenarios is not so clear-cut. In the second scenario, there is no loss of output from unemployment, but the assignment of workers to jobs is not efficient. In the third scenario, the economy eventually gets back to the most efficient assignment of jobs, but at the cost of some lost output and unemployment (and, in the real world, various other costs of transition incurred by workers and firms).

You can think of the time spent in unemployment in the second scenario as a type of investment. The economy forgoes some output in the short run to enjoy a more efficient match of workers and firms in the long run. As with any investment decision, we decide if it is worthwhile by comparing the immediate cost (the first four weeks of lost output) with the discounted present value of the future flow of benefits. Discounted present value is a technique that allows us to add together the value of dollars received at different times.

Toolkit: Section 31.5

Discounted present value is a technique for adding together flows at different times. If you are interested in more detail, review the toolkit.

Suppose, for example, that it takes four weeks for the economy to reallocate the jobs in the third scenario. Assuming the workweek has 5 working days, the economy produces 0 output instead of 14 units of output for a total of 20 days. The total amount of lost output is $20 \times 14 = 280$. Once the workers have found their new jobs, the economy produces 10 more units per week than previously. After 28 weeks, this extra output equals the 280 lost units. If we could just add together output this month and output next month, we could conclude that this investment pays off for the economy after 28 weeks. Because output produced in the future is worth less than output today, it will actually take a bit longer than 28 weeks for the investment to be worthwhile.

Provided that changes to the relative productivity of workers do not occur too frequently, the costs of adjusting the assignment of workers to jobs (the spells of unemployment) will be more than offset by the extra output obtained by putting workers into the right jobs. This is the gain from a fluid labor market, even though the process entails spells of unemployment.

Youth Unemployment

We observed earlier that the unemployment rate for young workers is higher than for older workers, in both France and the United States. We can understand why by thinking about the search and matching process.

When lawyers, doctors, professors, and other professionals change jobs, they typically do so with little or no intervening unemployment. Search and matching is easy because they have visible records, meaning their productivity at a particular job is relatively easy to figure out. In general, the longer someone has been in the workforce, the more information is available to potential new employers. Also, experienced workers have a good understanding of the kinds of job that they like.

Just the opposite is more likely in the labor market for young workers. Firms know relatively little about the young workers they hire. Likewise, young workers, with little employment experience, are likely to be very uncertain about whether or not they will like a new job. The result, at least in the United States, is a lot of turnover for young workers. Young workers sample different jobs in the labor market until they find one suited to their tastes and talents. They take advantage of the fluid nature of the US labor market to search for a good match. The gain is a better fit once they find a job they like. The cost is occasional spells of unemployment.

In Europe, search and matching is much harder. Some young workers are even effectively guaranteed jobs for life by the government from the moment they finish college. By contrast, young workers without jobs find it difficult to obtain employment. Given the lack of fluidity in European labor markets, it is surprising neither that more young workers are unemployed, nor that they stay unemployed for longer periods of time.

The Natural Rate of Unemployment

We expect there to be some frictional unemployment, even in a well-functioning economy. We also know that there is cyclical employment associated with the ups and downs of the business cycle. When cyclical unemployment is zero, we say that the economy is operating at *full employment*. The **natural rate of unemployment** is defined as the amount of unemployment we expect in an economy that is operating at full employment—that is, it is the level of unemployment that we expect once we have removed cyclical considerations.

The natural rate of unemployment can seem like an odd concept because it says that it is normal to have unemployment even when the economy is booming. But it makes sense because all economies experience some frictional unemployment as a result of the ongoing process of matching workers with jobs. Government policies that affect the flows in and out of employment lead to changes in the natural rate of unemployment.

key takwaways

- The static model of labor supply and labor demand fails to capture the dynamic nature of the labor market and does not account for job creation and destruction.
- In the United States, labor markets are very fluid. Each month, a significant fraction of workers lose their jobs, and each month a significant fraction of unemployed workers find jobs.
- Search theory provides a framework for understanding the matching of workers and jobs and wage determination through a bargaining process.
- The economy is operating efficiently when workers are assigned to jobs based on comparative advantage. Inflexible labor markets lead to inefficient allocations of workers to jobs.

Exercises

1. Is it best to assign workers to jobs based on absolute advantage or comparative advantage?
2. Why is frictional unemployment not always zero?

This page titled 23.3: Job and Worker Flows is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

23.4: Hours Worked 23.4: Hours WorkedEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What are the facts about hours worked across countries?
2. What are the explanations for these differences in hours worked?

The total number of hours worked in an economy depends on both the number of people who are employed and the number of hours worked by each employed person. So far, we have said little or nothing about this second issue. But another significant difference between Europe and the United States is that people work less in Europe than in the United States. If you hear such a statement, perhaps on the radio, you might have some questions about this comparison.

- Does this difference stem from differences in productivity? That is, is it the case that workers in Europe are less productive than workers in the United States, so it is less worthwhile for them to work as much?
- How is the difference measured? For example, suppose we simply divided the number of hours worked in an economy by the total population and found that this number was higher in the United States than in Europe. There are many possible reasons why this might be true. It could be because labor force participation is higher in the United States. Or it could be because the unemployment rate is lower in the United States. Or it could be because the average employed person in Europe works fewer hours than the average employed person in the United States.

Such questions simply mean that we had better be sure that we get our facts straight. We do this in the next part of this chapter. After that, we again turn to some theory to understand what is going on.Discussions of this topic by academics have been prompted by the work of Nobel Prize–winning economist Edward C. Prescott. The following article provides an overview and analysis of the key issues: Edward S. Prescott, "Why Do Americans Work So Much More Than Europeans?" *Federal Reserve Bank of Minneapolis Quarterly Review* 28, no. 1 (July): 2–13, accessed August 22, 2010, http://www.minneapolisfed.org/publications_papers/pub_display.cfm?id=905.

Hours Worked in Europe and the United States

Figure 23.4.1 and Figure 23.4.2 show some basic facts about hours worked in the United States and Europe.Richard Rogerson, "Understanding Differences in Hours Worked," *Review of Economic Dynamics* 9 (2006): 365–409. Figure 23.4.1 shows how hours worked in a number of different European countries compare to hours worked in the United States. More precisely, it shows the total hours worked by individuals between 15 and 64 years old divided by the number of people in that age group. The table does not distinguish by employment status: all working age people are counted, not just employed people.

Three of the largest European countries—France, Germany, and Italy—average less than 75 percent of the hours worked in the United States. Part of this difference is due to longer holidays in Europe, and part is due to the fact that the workweek in Europe is typically shorter. Because the table counts all working age people, the higher unemployment rate in Europe also contributes to the difference.

H < .75	.75 < H < .85	.85 < H < .95	H > .95
Belgium	Austria	Denmark	Australia
France	Finland	Greece	Canada
Germany	Ireland	Portugal	Japan
Italy	Netherlands	Sweden	New Zealand
	Norway	UK	Switzerland
	Spain		

H = Hours in Europe/Hours in US

Figure 23.4.1: Hours in Europe Relative to the United States

Figure 23.4.2 looks at the hours worked in various countries over the 40 years from 1970 to 2009. The measure of hours is calculated in the same manner as Figure 23.4.1. Average hours worked have declined significantly in most of these countries. Meanwhile average hours worked in the United States have been more or less flat over these four decades. As a result, hours worked are now significantly higher in the United States than in any of these countries. The data come from OECD (2010), "Hours Worked: Average annual hours actually worked", *OECD Employment and Labour Market Statistics* (database). doi: 10.1787/data-00303-en (Accessed on 18 October 2011) http://scholar.harvard.edu/alesina/files/work_and_leisure_in_the_u.s._and_europe.pdf. Figure 1 shows a similar pattern of divergence in hours worked for employed people, though the hours worked per employed person has declined in all countries over this period.

Research by the Nobel Prize–winning economist Edward Prescott paints a similar picture. He reports that from 1993 to 1996, the hours worked per person in France were about 68 percent of the level in the United States. In addition, US output per person was much higher than in Europe. Prescott explains this difference based on the number of hours worked, not by differences in output per hour worked. In other words, the United States is richer, not because it is more productive but simply because people work more.

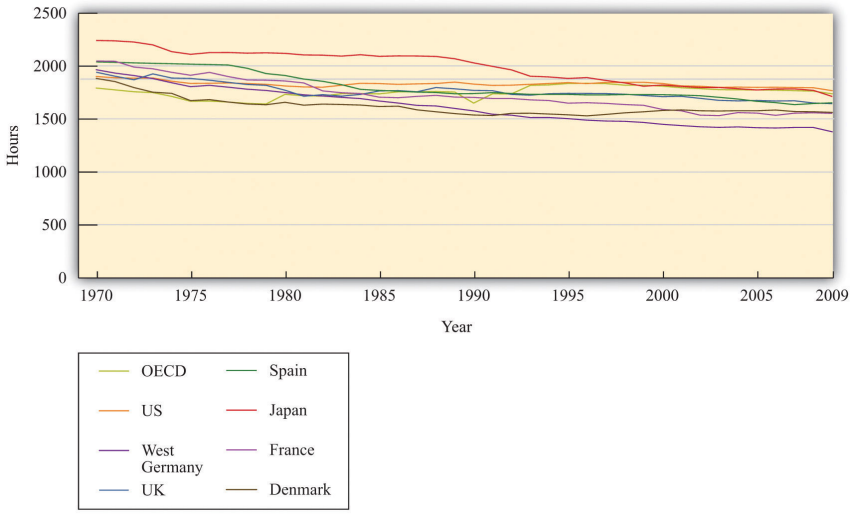


Figure 23.4.2: Annual Hours in Various Countries

Where Do Differences in Hours Worked Come From?

The immediate question is, why do people work more in the United States? A natural place to look for explanations is the labor supply decisions of households. One possibility is simply that the tastes of US and European households are different. Perhaps Europeans prefer having fewer goods and more leisure. Although this is possible, economists prefer to start from the presumption that people have broadly similar tastes and look first to see if there are other plausible explanations.

The differences in hours worked are not explained by Europeans having poorer technology. Both the United States and European countries are highly developed, so technologies used in one country are used in the others as well. Supporting this is the fact that, as we already noted, productivity does not appear to be lower in Europe.

Another candidate explanation is that there are differences in the tax system. Figure 23.4.3 shows an **individual labor supply curve**—in either Europe or the United States. Notice in Figure 23.4.3 the wage on the vertical axis is the real wage *after taxes*. This is defined as follows:

$$real\ wage\ after\ taxes\ = real\ wage \times (1 - tax\ rate).$$

In this equation, the tax rate is a *marginal* tax rate. This means that it is the tax paid on the *extra amount you earn* if you work a little bit more. Suppose the tax rate is 0.40 and your real wage per hour is $10. Then, if you work an extra hour, you pay $4 to the government, and you retain $6.

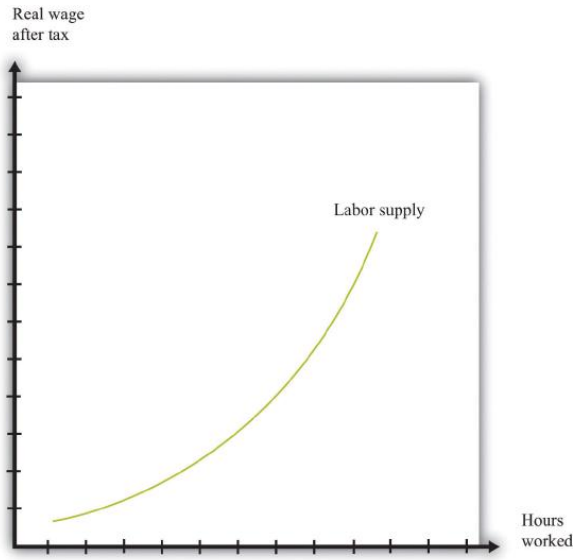


Figure 23.4.3: *Labor Supply*

Toolkit: Section 31.3

If you want to see the underpinnings of the labor supply curve, you can look in the toolkit.

Figure 23.4.3 shows that an increase in the after-tax real wage will cause an individual to supply more time to the market and thus consume less time as leisure. The increase in the wage creates an incentive for the individual to substitute away from leisure because it has become more costly.

Suppose that we compare two identical individuals in Europe and the United States. If the marginal tax rate in Europe is higher than it is in the United States, then the after-tax wage in Europe will be smaller. Since labor supply is upward sloping, individuals in Europe will work less than individuals in the United States. For this to be a convincing explanation, two things must be true:

1. Marginal tax rates must be higher in Europe.
2. Labor supply must slope upward enough to match the differences in hours.

Marginal tax rates are indeed lower in the United States than in Europe. Recent research finds that the marginal tax rate on labor income is about 34.5 percent in the United States compared to 57.7 percent in Europe (Germany, France, Italy, and the United Kingdom).Alberto F. Alesina, Edward L. Glaeser, and Bruce Sacerdote, “Work and Leisure in the U.S. and Europe: Why So Different?” (Harvard Institute for Economic Research, Working Paper #2068, April 2005), accessed June 30, 2011, www.colorado.edu/Economics/morey/4999Ethics/AlesinaGlaeserSacerdote2005.pdf. So, if you work an extra hour and earn a pretax wage of $10, then you would keep $6.55 in the United States and $4.23 in Europe.

The evidence is also consistent with the view that labor supply increases as the after-tax real wage increases. Figure 23.4.4 shows the implication of this. On the vertical axis are two different levels of the after-tax real wage: a low one for Europe and a higher one for the United States. These differences in the after-tax real wage translate into differences in hours, using the labor supply curve of an individual. Thus, as in Figure 23.4.4, individuals in the United States work more hours than in Europe. As this is true for everyone in the labor force, this argument immediately translates into a statement about hours worked for the aggregate economy.

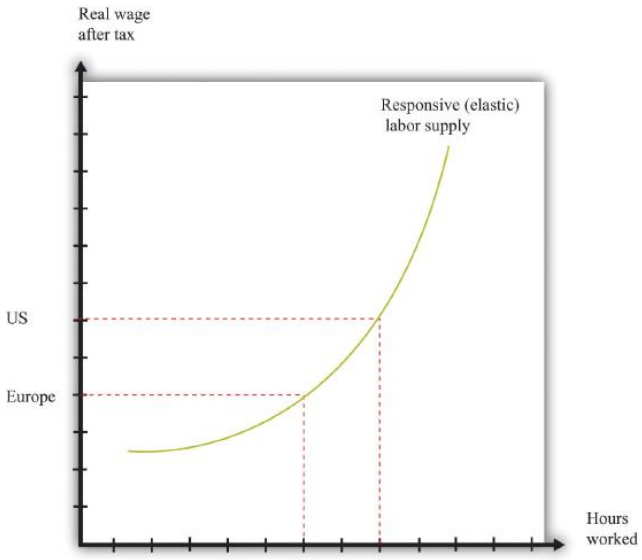


Figure 23.4.4: Differences in Hours Supplied. There are two real wages after taxes shown: one for Europe and one for the United States. These differences in real wages translate into differences in hours worked.

Can the difference in the after-tax real wage explain the observed difference in hours worked? This depends on how responsive labor supply is to changes in the real wage. Figure 23.4.5 shows two labor supply curves. In one case (the solid curve), labor supply is very responsive to changes in the wage. Relatively small differences in taxes then have substantial effects on hours worked. In the other case (the dashed curve), labor supply is not very responsive to the wage. Differences in tax rates are then unlikely to be able to explain the differences in hours worked.

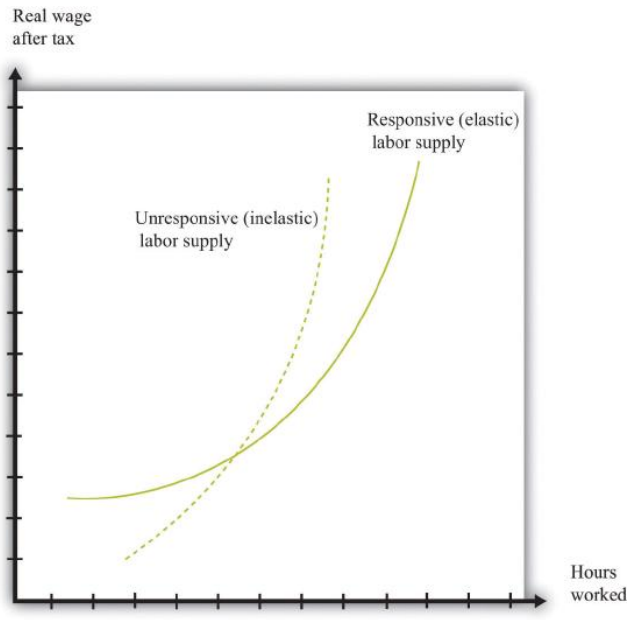


Figure 23.4.5: Responsive and Unresponsive Labor Supply. For the solid labor supply curve, hours worked responds strongly to changes in the real wage after taxes, while for the dashed curve, the response is very weak.

Prescott argues that the difference in taxes between the United States and Europe is enough to account for the differences in hours worked. To make this argument, Prescott holds fixed the labor supply curve (Figure 23.4.4) across countries and asks how much of the observed difference in hours can be explained by tax policy. This is a *movement along* the labor supply curve because the vertical axis measures the after-tax real wage. To support this argument, however, Prescott assumes that labor supply is indeed quite responsive to changes in after-tax wages.

Key Takeaways

- The average hours worked varies over countries. In the United States, the average hours worked are greater than in Europe.
- One way to explain differences in hours worked is through the higher marginal labor income taxes paid in Europe.

Exercises

1. Draw a diagram of the labor market to show how taste differences might explain differences in hours worked across countries.
2. In Figure 23.4.4, why is a tax policy change a movement along the labor supply curve and not a shift in the labor supply curve?

This page titled 23.4: Hours Worked is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

23.5: The Government and the Labor Market 23.5: The Government and the Labor MarketEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What are common forms of government intervention in labor markets?
2. Why do governments intervene in labor markets?

The employment and unemployment experience of Europe is quite different from that of the United States. We have developed some frameworks that help us understand the sources of these differences. But we have not yet really addressed the question at the heart of this chapter: what is the impact of different labor market policies in the two places?

Government interventions in the labor market are commonplace in most European countries. In Europe, there are many examples of restrictions on hiring, firing, the closing of plants, and so forth. There are some restrictions of this kind in the United States as well but not to the extent that we observe in Europe. In part this is because public opinion in Europe is more supportive of such regulations, as compared to the United States. For example, in 2003, the French food producer Danone decided to close two unprofitable factories in France. This news, which would almost certainly have been unexceptionable in the United States, led to massive protests, boycotts, and condemnation by politicians.

Europe is not the only part of the world in which governments intervene directly in labor markets. Labor regulations have recently been under consideration in China as well.Joseph Kahn and David Barboza, "China Passes a Sweeping Labor Law," *New York Times*, World Business, June 30, 2007, accessed June 30, 2011, http://www.nytimes.com/2007/06/30/business/worldbusiness/30chlabor.html.

The new labor contract law, enacted by the Standing Committee of the National People's Congress, requires employers to provide written contracts to their workers, restricts the use of temporary laborers and makes it harder to lay off employees.

Because of China's communist history, most workers are not represented by labor unions. It is the government that steps in to represent workers. The need to do so is enhanced by the increasing share of private rather than publically owned firms in China's economy.

We finish this chapter by considering some of the policies that have been adopted by governments in an attempt to influence the functioning of labor markets. We are interested both in why policymakers think these policies are a good idea and in the effect of these policies on the economy.

Unemployment Insurance

In Figure 23.5.2, we described the flow of workers between situations of employment, unemployment, and out of the labor force. We also argued that having a flexible labor market in which people can change jobs easily may more than compensate for the fact that people may sometimes spend time in unemployment.

But this is abstract economist-speak. People who lose their jobs, even if only temporarily, see their livelihood vanish. The reallocations of jobs that are beneficial to the economy as a whole may be costly, even devastating, to the affected individuals. For this reason, most developed economies have some kind of unemployment insurance to protect their workers. Unemployment insurance means that, if you are unemployed, you will receive some income from the government. Exactly how long you receive this income for and exactly how much you get depends on where you live. Some countries have much more generous unemployment insurance than others. Even if you live in the United States, the amount of insurance varies from state to state.

When it comes to buying car insurance, home insurance, or life insurance, households typically decide for themselves how much insurance to purchase. It is not a decision made by the government. Unemployment insurance is different: it is provided by the government rather than by private companies. This insurance is funded by taxes levied on firms and workers together.

The reason unemployment insurance is provided by the government is because it might be difficult for private firms to provide this coverage. Private insurance companies rely on the fact that not everyone makes claims on insurance at the same time. For example, a provider of home insurance knows that 20 percent of the houses that they insure will not burn down in the same month. But in a recession, the high rate of unemployment means that a lot of people claim benefits at the same time. If private insurers were

providing the benefits, insurance companies might go bankrupt, leaving workers without insurance. The government, by contrast, can use its ability to borrow, so it can finance unemployment insurance in one year from tax receipts it will receive in the future.

In the United States, the amount of insurance you receive typically depends on how much you have earned over the past year. A rule of thumb is that workers get about 25 percent of their wage income paid back through unemployment insurance. Benefits are available for only 26 weeks, although this is usually extended when the economy is in a recession. Other countries have much more generous programs."The Ins and Outs of Long-Term Unemployment," *OECD Employment Outlook 2002*, accessed June 30, 2011, http://www.oecd.org/dataoecd/36/48/17652683.pdf. Table 4.1 provides an extensive cross-country comparison. In Denmark, for example, unemployment benefits are about 90 percent of labor income and can last for up to 4 years.

Unemployment insurance has two main effects. First, and most obviously, this insurance makes it easier for unemployed people to sustain their level of consumption until they regain employment. Thus this form of insurance helps support **consumption smoothing**. Second, unemployment insurance affects the incentives of the unemployed. If individuals know they will receive some income even when they are unemployed, they are more likely to be willing to search extensively for good jobs. Instead of feeling the need to take the first job that comes along, people can wait longer and search longer for a job that is a really good match.

Unemployment insurance therefore contributes to labor market flexibility. It is, however, tricky to decide just how much unemployment insurance should be provided. After all, if unemployment insurance is too generous, then unemployed workers will be tempted to defer getting a new job for a long time—perhaps indefinitely. For this reason, governments usually restrict the period of time for which a worker can collect insurance to provide an incentive for them to search for a job.

Firing Costs

Imagine that you are the human resources (HR) manager of a firm in the United States. Suppose that the demand for your firm's product has declined, so you need to lay off some workers. You will be obliged to provide two weeks' notice to them. In many cases, that will be the end of your firm's obligations, although workers may sometimes be entitled to additional severance payments as part of their employment contracts. In the United States, employment contracts are largely a private matter between a firm and its workers. A firm cannot fire a worker for a discriminatory reason, but otherwise the government stays out of the contractual agreements among workers and firms. According to the Department of Labor, "In general, if the reason for termination is not because of discrimination on these bases, or because of the employee's protected status as a whistleblower, or because they were involved in a complaint filed under one of the laws enforced by the Department of Labor (see Whistleblower and Non-Retaliation Protections), then the termination is subject only to any private contract between the employer and employee or a labor contract between the employer and those covered by the labor contract.""Termination," US Department of Labor, accessed June 30, 2011, www.dol.gov/dol/topic/termination/index.htm.

In other countries, matters are not so simple. Imagine now that you are the HR manager of a firm in Portugal. Your product demand has fallen off, and you want to reduce output. In contrast to the United States, you may not be able to simply lay off workers. In Portugal, and in many other countries, there are numerous laws that make it costly to dismiss workers.

If you want to design a public policy to reduce the unemployment rate, it is tempting to make it harder to fire workers. If it is difficult to fire people, then fewer individuals will move from employment into unemployment. As we discussed earlier, though, spells of unemployment are sometimes necessary if workers are to move from less productive jobs to more productive ones. An increase in firing costs makes the labor market less flexible, so the economy will adjust less effectively to changes in workers' productivities.

There is also a more subtle unintended consequence of firing costs. If it is harder to fire workers, then firms become more reluctant to hire workers. Neither firms nor workers know the true value of a match in advance. When you take a part-time job, your productivity at that job and job satisfaction cannot be known ahead of time. Suppose there was a law that stated that once you accept a job you must stay with that employer for five years. You would certainly become very careful about deciding to accept a job offer. Exactly the same applies to firms. If the cost of laying off a worker is very high, then the firm will simply not hire the worker. A policy designed to promote employment can actively discourage it.

The French government, as we saw at the beginning of the chapter, made an attempt to introduce labor market reforms based on exactly this reasoning and tried to make the argument that we have just outlined to the protesters in the streets. If there were more flexibility in the firm's employment decision, they argued, firms would become more willing to hire young workers. This would help to reduce youth unemployment. The following *New York Times* article tells what happened next.Elaine Sciolino, "Chirac Will Rescind Labor Law That Caused Wide French Riots," *New York Times*, April 11, 2006, accessed June 30, 2011, http://www.nytimes.com/2006/04/11/world/europe/11france.html?_r=1.

President Jacques Chirac crumbled under pressure from students, unions, business executives and even some of his own party leaders on Monday, announcing that he would rescind a disputed youth labor law intended to make hiring more flexible. The retreat was a humiliating political defeat for both Mr. Chirac and his political protégé, Prime Minister Dominique de Villepin […]

It also laid bare the deep popular resistance to liberalizing France's rigid labor market, and makes any new economic reform politically impossible before a new government is in place, and perhaps not even then.

"Dead and buried," is how Jean-Claude Mailly, leader of the leftist union Force Ouvrière, described the fate of the labor law. "The goal has been achieved."

[…]

The new law was intended to give employers a simpler way of hiring workers under 26 on a trial basis without immediately exposing companies to the cumbersome and costly benefits that make hiring and firing such a daunting enterprise. Opposition to the law reflects the deep-rooted fear among the French of losing their labor and social protection in a globalized world.

[…]

In its initial form, the law allowed employers to fire new employees within two years without cause. In the face of mounting pressure, Mr. Chirac watered it down so that employers could subject new employees to only a yearlong trial period, and then would have to offer a reason for any dismissal.

Students and unions, bolstered by support from the opposition Socialists and even some business leaders, had vowed to continue their street protests until the law was rescinded.

The Socialists were quick to proclaim victory on Monday. "This is an unquestionable retreat," Francois Hollande, the leader of the Socialist Party, told reporters. "It is a grand success for the young and an impressive victory for the unity of the unions."

[…]

Restrictions on Hours

Another tempting policy to increase employment is to limit the number of hours an employee can work. Suppose that a firm needs 1,200 hours of labor time a week. If a typical worker works 40 hours per week, then the firm will need to hire 30 workers. But if the government were to legislate a 30-hour workweek, then the firm would need to hire 40 workers instead.

This idea of "spreading work" through restrictions on hours was part of the response in the United States to the Great Depression. During the early 1930s, the US government instituted such restrictions under the heading of the "National Economic Recovery Act." The idea persists to the present day. In France, the government passed a law limiting hours worked to 35 hours per week (for workers at large firms) starting in the year 2000. In Germany, the government operates a policy called Kurzarbeit, whereby it subsidizes firms who retain workers for shorter hours in times of recession.

One problem with such policies is that restrictions on hours reduce the value of a match between a worker and a firm. Consequently, fewer matches will be formed, and more workers will be unemployed. Another problem is that it reduces flexibility in the labor market, which leads to less efficient functioning of the economy.

As a concrete example, consider auto manufacturers in the years following the Great Depression in the United States. This industry had substantial variations in hours worked over the model year. During times of high demand for cars (the spring), factories and their workers were working overtime to meet the increased demand. Restrictions on hours meant that overtime working had to be replaced by increased hiring. Firms that wanted to produce more output had to hire and train new workers. This was costly, so firms sometimes found it was better simply to accept that they would not meet the high demand.

In the case of France's 35-hour workweek, matters were a bit more complicated. The mandated short workweek imposed some rigidity on firms. However, during the negotiations for this change in the laws, French labor unions agreed to some other changes that improved the flexibility of the labor market. France later moved away from the 35-hour workweek by permitting firms and workers to agree to longer work hours if they wish.

Key Takeaways

- Most governments provide workers with unemployment insurance. In many countries, governments also impose costs on firms that fire workers and also restrict hours worked.

- One rationale for intervention by governments is to provide insurance to workers that is not available in private markets. Governments also take action in an attempt to increase employment rates.

Exercises

1. Can a firm in the United States fire a worker without permission of the government?
2. What was one of the arguments in France for restricting the hours worked per week?

This page titled 23.5: The Government and the Labor Market is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

23.6: End-of-Chapter Material 23.6: End-of-Chapter MaterialEdit Update titlecancel

In Conclusion

Europe and the United States differ in many ways. From the perspective of macroeconomists, some of the most striking differences are in the laws governing labor markets.

In the United States, labor markets are relatively flexible. It is relatively easy for firms to hire and fire workers, and it is relatively easy for workers to move between jobs. This brings many benefits to the economy as a whole, the most important being that it helps ensure good and productive matches between workers and firms. It also has some less attractive implications, particularly for workers. Job security is very limited, and workers might find themselves out of a job with very little warning.

In Europe, labor markets tend to be more rigid. We have explored some of the ways in which this is true. Minimum wages are often higher, unemployment insurance is more generous, and the costs of hiring and firing workers are greater. As a consequence, European countries are typically characterized by higher unemployment than the United States. In addition, unemployment duration tends to be longer: workers who become unemployed tend to take longer to find a new job. This makes the labor market a more difficult place for workers who do not have jobs but a better place for those who do have jobs because they typically enjoy higher salaries and greater security.

We have analyzed the differences between these two parts of the world, but we have not explained why these different economies have settled on such different configurations of labor laws. The explanation is not simple and goes well beyond economics into questions of history, politics, and sociology. Still, there is probably some truth in the simplest explanation: voters have different preferences about how their working lives should look. Perhaps voters in Europe prefer a world of greater job security for the employed, even if it comes at the cost of unemployment problems and a less-efficient economy. Perhaps voters in the United States prefer a dynamic economy, even if it comes at the cost of more uncertainty for working people.

Key Links

- US Department of Labor, information on unemployment insurance: http://www.ows.doleta.gov/unemploy/aboutui.asp
- Organisation for Economic Co-operation and Development (OECD) key employment statistics: http://www.oecd.org/document/53/0,3746,en_2825_495670_42788213_1_1_1_1,00.html
- European Union statistics: labor market policies: epp.eurostat.ec.europa.eu/portal/page/portal/labour_market/introduction
- Bureau of Labor Statistics, labor force statistics: http://www.bls.gov/cps/home.htm
- International Labour Organization, labor policies: http://www.ilo.org/empelm/lang--en/index.htm

exercises

1. A *Washington Post* article quoted the following opinion from a French student.Molly Moore, “French Students Hit Streets to Protest New Labor Law,” *Washington Post*, World News, March 17, 2006, accessed July 7, 2011, www.washingtonpost.com/wp-dyn/content/article/2006/03/16/AR2006031601908.html. Do you agree or disagree with these views? Do you think of the labor market experience in your country differently?

   “They’re offering us nothing but slavery,” said Maud Pottier, 17, a student at Jules Verne High School in Sartrouville, north of Paris, who was wrapped in layers of scarves as protection against the chilly, gray day. “You’ll get a job knowing that you’ve got to do every single thing they ask you to do because otherwise you may get sacked. I’d rather spend more time looking for a job and get a real one.”

2. (Advanced) What effect does unemployment insurance have on the savings behavior of employed households? Think about the life-cycle model, discussed in Chapter 28 (and in the toolkit). How would you add the prospect of unemployment to the household’s life-cycle decisions on consumption and saving?
3. Explain how each of the following factors might affect the duration of unemployment for a single unemployed worker: (a) rate of unemployment in the economy, (b) age of the worker, (c) skills of the worker, (d) country of the worker, (e) generosity of unemployment insurance, (f) wealth of the worker, and (g) employment status of the worker’s spouse. What other factors can you contribute to this list?

4. The following table contains information about worker output in two jobs. Explain why worker B has an absolute advantage in both jobs. What is the most efficient assignment? Which worker has a comparative advantage in job 1? Calculate the opportunity cost of assigning the workers to job 2. Which worker has a lower opportunity cost of taking job 2?
5. Consider the following job assignment problem based on the table titled "Output Level per Hour from Assigning Jobs". Here there are three workers, three jobs, and the prospect of not working. In the table, the value of output produced not working can be interpreted as the value of either leisure time or the output produced at home (say, in the garden). Find the optimal assignment of workers to jobs. Should anyone be unemployed? If not, how would you change the table so that someone was not working?
6. Explain why making it easier to fire people might reduce the unemployment rate.
7. Suppose that there is a legal minimum wage, set in nominal terms. Draw a diagram to show how this can lead to unemployment. Now suppose that there is inflation. What happens to the employment rate? What happens to the unemployment rate?

Worker	Job 1	Job 2
A	1	9
B	2	12

TABLE 23.6.1: OUTPUT LEVEL PER DAY IN DIFFERENT JOBS

1.

Worker	Job 1	Job 2	Job 3	Not Working
A	10	12	6	0
B	8	1	1	2
C	6	3	5	3

TABLE 23.6.2: OUTPUT LEVEL PER HOUR FROM ASSIGNING JOBS

Economics Detective

1. Go to the website for the Organisation for Economic Co-operation and Development (http://www.oecd.org/home/0,2987,en_2649_201185_1_1_1_1_1,00.html). Find the latest table reporting unemployment rates in Europe. How is unemployment defined in this table?
2. Find a recent discussion of employment protection laws across countries. In which countries are jobs most regulated? Has this changed much over time? Can you find any evidence relating the measure of employment protection laws with the unemployment experience of the individual countries?
3. Go to the website for the Current Population Survey (http://www.bls.gov/cps). Develop a figure similar to Figure 23.6.2 for the current month. Why do the numbers differ from those reported in Figure 23.6.2? Find a year when the United States was in a recession. What were the rates of job flows like during the recession?
4. Find a discussion of the unemployment insurance that would apply to you if you lost a job where you currently live. Does it matter in your state/country *why* you are not currently employed? In your state/country, do you have to continue to look for a job to receive unemployment insurance? If so, what do you have to do?
5. In Europe, is the amount of unemployment insurance determined by individual countries or by the European Union?
6. Go to the website for the Bureau of Labor Statistics (http://www.bls.gov) and find out what ages are classified as working age in the United States.

This page titled 23.6: End-of-Chapter Material is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

24: Money- A User's Guide

24.1: The Color of Money
24.2: What Is Money?
24.3: Using Money to Buy Goods and Services
24.4: Using Money to Buy Other Monies- Exchange Rates
24.5: Using Money to Buy Assets- Interest Rates
24.6: End-of-Chapter Material

This page titled 24: Money- A User's Guide is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

24.1: The Color of Money 24.1: The Color of MoneyEdit Update titlecancel

You wake up one morning, drag yourself out of bed, and, bleary-eyed, throw on some clothes. You stumble out of your apartment and across the road to your neighborhood coffee shop. “Coffee, please,” you say to the barista, those being the only two words that you are capable of getting out of your system before you get some caffeine into it. She pours a cup of the coffee of the day and places it on the counter in front of you. Just smelling the coffee makes you feel a little bit better already.

“That’ll be a dollar.” You reach into your pocket, pull out a crumpled twenty, hand it to her, and reach for the cup. “I’m sorry,” she says, pulling the cup away from you, “I can’t accept that.” “Why on earth not?” you ask, bemused. “It’s the wrong color,” she says. “You could have used that yesterday, but—look—this is what bills look like now.” And she reaches into her register and shows you a bright purple $20 bill, like the one in the following figure .



Figure 24.1.1: The New $20 Bill. Imagine if you woke up one morning and found that all money was now this color.

In this story, normal green dollar bills were accepted as money yesterday, while purple dollar bills were worthless colored pieces of paper. But today, purple dollar bills are accepted as money, and green dollar bills are just worthless pieces of paper. This sounds absurd. Yet it is not so far from what happened in a dozen different countries on January 1, 2002. If you had awakened in Italy on that day and gone down the street to a neighborhood café, you would have noticed that the simple act of buying coffee had changed from the day before. Your local café still looked as it did on December 31, 2001. But where you had previously paid with notes and coins called Italian lira, you would now pay with a completely new currency called the euro.

The same was true in France, Finland, Germany, Greece, and seven other European countries. On that day, 12 countries all officially gave up their own currencies and instead adopted a common currency—the euro. Admittedly, the transition was not quite as stark as in our story: there was a period of about 2 months in which euros and the old local currencies both circulated. But the essence is the same. At one time, euro notes were just colored pieces of paper that shopkeepers would not accept for transactions. Then, not that long afterward, those colored pieces of paper became valuable, while the old currencies turned into worthless pieces of paper.

This was an amazing event for the international economy. Familiar currencies like the French franc, the German deutschmark, the Greek drachma, and the Spanish peseta simply disappeared. The following figure shows some of these vanished currencies. Some of the world’s largest economies changed their currency. In Chapter 30, we take up another aspect of this event: what it means for a country to disband its central bank and delegate monetary policy to a centralized entity. To make sense of this event, we need to answer a disarmingly simple-looking question, which is the theme of this chapter:

Why do people want to hold apparently worthless pieces of paper?



Figure 24.1.2: Some Vanished Currencies. Here are some of the banknotes that disappeared from circulation in Europe upon the advent of the euro.

Road Map

Understanding what happened in Europe requires us to answer two more basic questions:

What is money? and Why is it valuable?

We begin this chapter by looking at what makes something a money. Surprisingly, this is not straightforward: we will see that money has several attributes, and many different things can act as money. Then we look at what we can do with money. We use money to buy goods and services, we use money to buy other kinds of money, and we use money to buy money in the future.

Before exploring the world of money, we need to make one clarification. In everyday language, if you bought a camera for $200 and sold it for $300, we would say that you made money from the deal. Economists, however, use the term *money* more precisely, in ways that we make clear in this chapter. An economist would say that your resale of the camera earned you income, and you received that income in the form of money.

This page titled 24.1: The Color of Money is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

24.2: What Is Money? 24.2: What Is Money?Edit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What gives money value?
2. What are the functions of money?

Take a look at some currency—a dollar bill, for example. It is nothing more than a piece of paper with writing on it. A very pretty piece of paper, perhaps, with fancy writing and some pictures, but it is still just a piece of paper. Yet people voluntarily give up valuable goods or services in exchange for pieces of paper. This is the mystery of money.

The question motivating this chapter—why do people want money?—is a deep one. That may seem a surprising claim because obviously we all like having money. But questions that seem trivial sometimes provide insights into how the world works. If we can understand why people want these intrinsically worthless pieces of paper, then we can understand why money is valuable. And to understand why people want these pieces of paper, we need to know what people want to do with their money.

The Characteristics of Money

A striking feature of modern developed economies is that people are typically *specialists in production and generalists in consumption*. By this we mean that most of us work at one or at most two jobs, producing (or, more often, helping to produce) a very small number of things—for example, a cattle farmer produces beef, a software designer produces computer code, and a nurse produces health services. However, we all purchase hundreds of goods and services.

There is no law that says that we have to buy goods and services using money. An alternative is to trade goods or services directly for one another. This is called **barter**. We do see some barter in the world. A restaurant may allow its employees free meals at the end of the night, which means that some of the employees' wages effectively takes the form of food. If a car mechanic and a caterer live next door to each other, they may have an informal arrangement whereby the mechanic repairs the caterer's truck in exchange for food for a birthday party.

Sometimes we know exactly where to buy the goods and services that we want. At other times, we go looking—perhaps walking or driving from store to store, perhaps searching using a phone book or the Internet. We do this because we don't know which store has the goods we want in stock; in addition, we might not know the prices that different stores are charging, and we want to hunt around for the best deal.

To understand the role that money plays in an economy, begin by imagining a world where we must search for the goods and services that we want to buy and there is no money, so all trades take place through barter. Imagine, for example, that you are a web designer, and you want to buy a used car. You must look around for someone who has a car for sale. This search takes time: it has an opportunity cost in that you would prefer to spend that time working or enjoying leisure. Eventually, you find someone who has a car that you are interested in buying. But your problems are not over. He has a car for sale, but what can you give him in exchange? You have to hope that he is interested in obtaining some web design services in exchange for the car. Successful barter requires a coincidence of wants: you must have what the other person wants, *and* they must have what you want.

A world of nothing but barter is hard to imagine. Each time you wanted to buy something from a seller in a store, you would have to exchange some good or service for that good. If you went to a café, you might have to wash the dishes in return for a coffee. Professors of economics wanting a meal would have to go from restaurant to restaurant trying to find a chef who wanted to hear an economics lecture. They would probably go hungry. It is easy to see why all societies find some way of making these transactions easier.

If you can carry some kind of money around with you to make purchases like these, life is much easier. You still have to hunt for the goods and services that you want, but you don't have to worry about whether the other party in the transaction wants the product that you sell. Money, therefore, plays a key role in ensuring that trades occur. Trades, in turn, create value in our economy. People are not forced to buy or sell things; they do so only if the trade leaves them better off than they were prior to trading. Money therefore plays a critical role in value creation.

The reason that we rarely see exchange without money is that it is so inefficient. Without money, a coincidence of wants is unlikely, so desirable trades do not occur, and value is not created. With money, transactions are much easier. If you want a meal in

a restaurant, the owner will always serve it to you if you have money. Likewise, you obtain money by working at your job. You don't care what good or service your employer produces; as long as your employer pays you in money, you are happy to supply your labor time to them.

Let us think for a moment about what characteristics this money needs to have:

- Money must be *portable*. If you are going to walk around searching for goods and services, you want to be able to carry money with you. Sacks of coal would not make a very good money.
- Money must be *divisible*. Different goods have different prices, and the money we use must accommodate that. Watches would not make a very good money.
- Money must be *durable*. Daffodils would not make a very good money.

It is easy to list many things that are reasonably portable, divisible, and durable: chocolate chip cookies, cigarettes, and printer paper are just a few examples. These are not typically used as money, although they could be. If you went into a fast-food restaurant, asked for a burger, and then offered to pay using chocolate chip cookies, you can be confident that you would not get the food that you want. That is because there is a fourth characteristic of money that is rather different from the other three.

- Money must be *acceptable*.

Something can function as money only if people are willing to accept it as money. It is not impossible to imagine a world where chocolate chip cookies function as money. If everyone else is willing to accept cookies in payment for goods and services, then you will be willing to do so as well. But if other people accept only printed pieces of paper as money, then you would be foolish to accept chocolate chip cookies for the product that you sell.

Fiat Money

We know of no country, of course, that actually uses chocolate chip cookies for money. In most countries, money takes a particular form called **fiat money**. Fiat money is money that is not backed by any physical commodity, such as gold. Instead, the currency is intrinsically useless pieces of paper that attain value in exchange.

Fiat is a Latin word that means "let it be." Fiat money is money just because the government says so. In a fiat money system, the government does not promise to exchange goods for money. In addition, money is not generally something that we can directly consume: most people would not enjoy eating a dollar bill. So if it doesn't taste good and the government doesn't promise to give you something in exchange for it, what gives fiat money value? Why are we all willing to work hard to get pieces of these—intrinsically worthless—pieces of paper?

The answer is because these pieces of paper are acceptable as money. Other people will accept them, so you and I will as well. To put it another way, *fiat money has value because everyone believes it has value*. Think back to the story with which we opened the chapter. The US economy uses green and white pieces of paper as money. US residents are willing to give up valuable goods and services in exchange for these green and white pieces of paper because they believe that others, in turn, will accept them. Such an arrangement sounds fragile, and it is. If everyone stopped believing that fiat money had value, this would be a self-fulfilling prophecy. Something very much like this happens in the circumstances of very high inflation rates, as explained in Chapter 26.

Suppose the money in an economy changed overnight from green pieces of paper to purple pieces of paper, as we fancifully suggested at the beginning of this chapter. Everyone now works for and accepts the new purple currency. You are forced to follow. It would be foolish for you to work and accept green paper because no one would give you goods or services in exchange. Instead, you demand to be paid in purple paper because that is what you now need to buy goods and services.

Of course, we do not often observe these switches across colors of paper within an economy. People get used to one type of currency, and it is difficult to change everyone's behavior at once. Still, Europe *did*, in effect, switch from green pieces of paper to purple pieces of paper. Sure enough, no one in Europe these days is willing to accept French francs, Portuguese escudos, or Finnish marks. These are the old pieces of paper. Now people will accept only the new pieces of paper.

That conversion was not truly instantaneous. Prior to the changeover to the euro, there was a switch to a dual unit of account: French bank statements in 2001 gave balances in both French francs and euros, for example. Even now, years after the changeover, bills in Europe often still appear in both the old local currency and euros. It was also possible to use the euro as a store of value before the changeover because banks started establishing accounts in euros.

Even though fiat money issued by the government is, in the end, just pieces of colored paper, it typically does have one particular property that stems from the power of the state. The US government states that it will accept dollars in settlement of government

debts—most importantly, tax bills. The government also states that dollars can be used in settlement of private debts. Dollars are legal tender.There is a subtle question here about whether this aspect of money means that even intrinsically worthless currency must always have some value. If people owe debts to the government that are specified in money terms, then they will be willing to pay something for legal tender currency.

The Functions of Money

Thus far, we have thought about money in terms of its characteristics. We can also think about what makes a good or bad money in terms of the functions that it serves.

Medium of Exchange

If you walk into an electronics store and see a camera with a price tag of $500, the store is making an *offer* to you and other customers: if you hand over ten $50 bills, you can have the camera in exchange. Money serves as a **medium of exchange**.

There are other ways to purchase a camera rather than cash. You could write a check, for example, or use a debit card (a card that immediately deducts the $500 from your bank account and pays it into the store's account). The fact that there are different ways of paying for something is a clue that there is, in fact, no single thing that we can call money. Money is anything that does what money does.

Interestingly, one common form of purchase does not involve money at all. If you use a credit card to buy a camera, you do not pay at all at the time of purchase, so no money—by any definition—changes hands. In this case, you receive the camera in exchange for a promise to pay for the camera later. It is only when that promise to pay is fulfilled that you hand over the money for the purchase.

Store of Value

Any medium of exchange must also serve as a **store of value**. This just means that money should keep its value between the time that you receive it (in exchange for goods that you sell or work that you do) and you spend it again.

If an object lost all or most of its value over a short period, then it would not be acceptable in exchange. So something that serves as money must be a store of value. Imagine for a moment an economy in which ice played the role of money. Except on the coldest days, the ice you receive on payday would not last long enough for you to buy anything with it. It would be a terrible store of value and, as a result, would not do a good job of facilitating exchange.

Paper money and coins are not like ice. They are durable and do not dissolve with use. Because of this, you can be confident that the dollar you have in your pocket today will still be a dollar you can spend tomorrow. The fact that people are willing to hold money for long periods of time is indicative of the role of money as a store of value. If money were not a store of value, then all people would want to get rid of cash as soon as they received it. To mix our metaphors: if money were ice, it would become a hot potato.

Being a store of value is more than just a physical property of money. Currency in your pocket can remain there for a long periods of time before disintegrating. So, in a physical sense, that currency retains its worth. But, if prices are increasing, then in terms of what the currency can buy, the money in your pocket is not retaining its value. In times of inflation, money functions less well as a store of value.

Unit of Account

Almost universally, prices are quoted in terms of some currency, such as pesos, dollars, or euros. Goods and services sold in the United States have prices in terms of US dollars. The dollar serves as a **unit of account**. But when the very same goods and services are sold in Europe, they are priced in a different unit of account: euros. This role of money is so familiar as to be mundane, yet our economy simply could not function without a commonly accepted monetary measuring stick. It would be like building a house without an accepted measure of length or running an airline without an accepted measure of time.

The unit that people use to keep account of their monetary transactions varies from country to country. In Mexico, prices are quoted in pesos, in India prices are quoted in rupees, and so on. In most countries, the medium of exchange and the unit of account are the same thing, but this need not be true.

Because the US dollar is known throughout the world, it is often used as a unit of account in unexpected places. Prices of commodities in international transactions may be quoted in terms of the dollar even when the transaction does not directly involve

the United States. Luxury hotels in China and elsewhere sometimes quote prices in US dollars even to guests who are not coming from the United States. In Chapter 18, we discuss both nominal and real gross domestic product (real GDP). Nominal GDP is the value of all the goods and services produced in an economy, *measured in terms of money*. Money is used as a unit of account to allow us to add together different goods and services. Even the concept of real GDP uses money as a unit of account: the difference is that we use money prices from a base year to value output rather than current money prices. As another example, after the changeover to the euro, that currency became the medium of exchange and the "official" unit of account. But many people—at least in terms of their own thinking and mental accounting—continued to use the old currencies. In everyday conversation, people continued to talk in terms of the old currencies for months or even years after the change.On a bike trip in the summer of 2002, one of the authors had lunch in a French country restaurant. Though it was many months after the change to the euro, the menu was still in French francs. An elderly lady running the restaurant painstakingly produced a bill in euros: for each entry (in French francs), she multiplied by the exchange rate (euros to francs) and then added the amounts together. Even today, some bills and bank statements in Europe continue to quote the old currency along with the euro.

Meanwhile, merchants in countries who have not adopted the euro may still quote prices in that currency. In Hungary, the local currency is called the forint. Figure 24.2.1 shows a sign at a restaurant in Budapest, Hungary, advertising goods in both currencies: goulash soup, for example, is sold for 1,090 forint or 4.40 euro. If, as may well be the case, the restaurant is also willing to accept euros in payment, then the euro is also acting as a medium of exchange alongside the forint.



Figure 24.2.1: The Euro as a Unit of Account. A sign at a restaurant in Hungary quotes prices in euros and the local currency (forint).

Source: Image taken by the authors

Key Takeaways

- Fiat money has value because everyone believes it has value.
- The three functions of money are medium of exchange, store of value, and unit of account.

Exercises

1. In what sense are you a specialist in production and a generalist in consumption?
2. Why is money less effective as a store of value when inflation is high?
3. In times of inflation, money is also less effective as a unit of account. Why?

This page titled 24.2: What Is Money? is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

24.3: Using Money to Buy Goods and Services 24.3: Using Money to Buy Goods and ServicesEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What is arbitrage?
2. What is the law of one price?

Having defined money through its characteristics and functions, we now turn to the uses of money. By looking at what we can do with money, we can understand how intrinsically worthless pieces of paper acquire their value.

Let us imagine, then, that you are lucky enough to find a $100 bill on the sidewalk. You have no way of returning it to its rightful owner. What might you do with this money? The first and most obvious answer is that you can use it to buy something you want: you can take the $100 and purchase some goods and services.

The Value of Money

The observation that we use money to buy things tells us more about the value of money. Economists often make a distinction between *real* and *nominal* values; this distinction can be applied to money as well. First, what is the nominal value of money? This is almost a trick question: we are asking, "How many dollars is a dollar bill worth?" The answer, which does not require a doctorate in economics, is that a dollar bill is worth $1.

Nominal variables—those measured in dollars or other currencies—can be converted into real variables—that is, those measured in units of real gross domestic product (real GDP). To convert a nominal variable to a real variable, we simply divide by the price level. For example, if your nominal wage is $20 per hour and the price level is $10 (meaning that a typical unit of real GDP costs this amount), then your real wage is 2 units of real GDP.

Toolkit: Section 31.8

If you want to review the process of correcting for inflation, you will find more details in the toolkit.

Exactly the same principle can be applied to money itself. The real value of a dollar is obtained by dividing one by the price level. Thus

Think of an economy in which real GDP is measured in pizzas and suppose the price level—the price of a pizza—is $10. Then the value of a dollar bill is 1/10 of a pizza.

Although $1 is always worth $1, you are not guaranteed that the dollar bill in your pocket will buy the same amount of goods and services from one day to the next. If your local café increases the price of a cookie from $1.00 to $1.25, then your $1 will no longer buy you a cookie; its value, measured in cookies, has declined. If the price level increases, then the real value of money decreases. For notes and coins to be a good store of value, it must be the case that prices are not increasing too quickly. We discuss this problem in more detail in Chapter 26.

Using Money to Make Money: Arbitrage

An old joke has it that the secret to getting rich is very simple: buy at a low price and sell at a high price. So another use of your $100 would be to buy goods not to consume but to resell—a process known as **arbitrage**.

Suppose you discovered that a particular model of digital camera could be bought much more cheaply in Minneapolis, Minnesota, than in Flagstaff, Arizona. Then you could purchase a large number of cameras in Minneapolis, load them into a suitcase, fly to Flagstaff, and sell them for a profit. If the gap in price were large enough to compensate for your time and travel costs, then this would be a money machine. By buying cameras at a low price and selling them at a high price, you could make as much profit as you wished.

This situation would not persist. You, and other entrepreneurs as well, would start to bid up the price of cameras in Minneapolis. Meanwhile, the increased supply of cameras in Flagstaff would cause prices there to decrease. Before too long, your money

machine would have dried up: the gap between the Flagstaff price and the Minneapolis price would no longer justify the effort.

Arbitrage ensures that the prices of individual goods do not vary too much across different regions of the United States. Taken to its extreme, it would imply that the price level would be the same throughout the country. Economists call this idea the **law of one price**. The law of one price says that different prices for the same good or service cannot persist because arbitrage eliminates such differences. Arbitrageurs would buy the good at the low price and sell it at the high price. Demand would increase in the market where the price was low, causing that price to increase. Supply would increase in the market where the price was high, causing that price to decrease. This process would continue until the prices were equalized across the two markets.

There are, of course, differences in the prices of individual goods and services in different states and different cities. These differences are primarily due to the fact that some items cannot be arbitraged. If cameras are cheaper in Minneapolis than in Flagstaff, then they can be bought and sold as we described. But if apartments in Flagstaff are cheaper than in Minneapolis, it isn't possible to ship them across the country. Likewise services typically cannot be arbitraged. Thus we do not expect the law of one price to be literally true for every good and service. Nevertheless, the law of one price does lead us to expect that the overall price level will not differ too much in different parts of the country.

It can be difficult to apply the law of one price in practice because we have to be careful about what we mean by the "same" product. An apparently identical shirt at two different retailers might not qualify as the same—perhaps one retailer allows goods to be returned, while the other does not allow returns. Identical goods are not the same if they are in different places: a Toyota on a dealership lot in Kentucky is not the same as the identical model car on a lot in Pretoria, South Africa, and so on. In such situations, the law of one price tells us that we should not expect prices of goods to be "too different," depending on the costs of transportation and the other costs of arbitrage.

We said earlier that money makes an economy more efficient because it makes transactions easier. Money makes arbitrage easier as well. Arbitrage would be a less certain way of making money in an economy with barter. First, the lack of a clear unit of account would make arbitrage opportunities less transparent. Second, the lack of a reliable medium of exchange would make arbitrage risky: the person in Flagstaff who wants to buy a digital camera from you might not have anything you want, so you might end up giving up something you own and not getting something you want in return.

Key Takeaways

- Arbitrage is the process of making a profit by buying goods at a low price and selling them at a higher price.
- When arbitrage is possible, we expect the same good to sell at the same price. There are no arbitrage profits to be made when the law of one price holds.

Exercises

1. All else being the same, if the price level increases, what happens to the real value of money?
2. Explain why the law of one price is less likely to hold for a service than for a good.

This page titled 24.3: Using Money to Buy Goods and Services is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

24.4: Using Money to Buy Other Monies- Exchange Rates 24.4: Using Money to Buy Other Monies- Exchange RatesEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What is the difference between the nominal exchange rate and the real exchange rate?
2. How is the law of one price related to the nominal exchange rate?

As we all know, there are multiple currencies in the world. These are most often associated with a single country: the yen in Japan, the yuan in China, the peso in Mexico, and so on. Sometimes many countries will use the same money, with the leading example being the use of the euro by the member countries of the European Union (http://ec.europa.eu/economy_finance/euro/index_en.htm). Sometimes multiple currencies are in use in a single place: when you land at a major European airport, such as Frankfurt, Germany, or Amsterdam, the Netherlands, you will see that you can buy a cup of coffee at the airport using many different currencies. Likewise, the US dollar is freely accepted in some countries in addition to the local currency, British pounds formerly were freely accepted in Ireland, and so on.

If you happened to find your $100 right before going on a trip to another country, you might decide to use it to buy the money of that country. For example, if you were about to take a trip to Canada, you could take the bill into a bank or a foreign exchange merchant and exchange it for Canadian dollars. If you want to buy goods and services in Canada, you need Canadian dollars because they are the medium of exchange in that country.

When you make such an exchange, you buy the local currency using your home currency. If you travel from the United States to Europe, you buy euros using dollars. The price you pay is the *dollar price of the euro*: the amount in dollars you must pay to obtain 1 euro. This is completely analogous to using a dollar to buy a bottle of soda, when you pay the dollar price of soda.

In practice, it is often unnecessary to carry out a physical exchange of notes and coins. In most countries, you can go to an automated teller machine (ATM) and withdraw local currency directly. Your bank deducts the equivalent sum in your home currency from your bank account. You are still carrying out an exchange, of course, but it is hidden from view, and you will see it only when you look at your next statement. The same is true if you make a purchase using a credit card.

Just as a US resident traveling to Europe wishes to buy euros with dollars, a visitor to the United States from, say, Holland will need to buy dollars with euros. The price she pays is the euro price of the dollar: the number of euros needed to obtain $1. The price of one currency in terms of another is called an **exchange rate**.

Toolkit: Section 31.20

If you want to review the definition of an exchange rate, you will find more details in the toolkit.

If we think of two currencies—euros and dollars, for example—then there are two exchange rates to keep in mind: the price of euros in dollars and the price of dollars in euros. (You might suspect, correctly, that these two prices are linked; we return to this shortly.) In a world of 3 currencies, each has a price in terms of the other two currencies, so there are 6 (= 3 × 2) different prices. And in a world of 100 currencies, then for each one, there are 99 prices for the other currencies. So there are 100 × 99 = 9,900 prices to quote! A Zambian traveling to Armenia wants to know about the kwacha price of drams, a Malaysian traveling to Oman is interested in the ringgit price of rials, and so on.

Foreign Exchange Markets

Imagine a series of three visitors traveling from the United States to Europe. First, we have someone arriving on vacation. Chances are that she will want to exchange dollars for euros to have money to spend on hotels, meals, and so on. She also buys souvenirs in Europe—goods that she imports back to the United States. Our second visitor spends a lot of time in Europe for work purposes. He might open a bank account in, say, Germany. If he wanted, he could use this bank account to keep some of his wealth in Europe. He would buy euros with his dollars, deposit these euros in the bank to earn interest, and then—at some point in the future—he would take his money out of the bank in Germany and exchange the euros for dollars. (Later, we will consider how you can decide if this is a good investment strategy. For now, our point is that this type of financial investment is another source of demand for euros.)

Our third visitor to Europe is a professional wine buyer who wants to purchase wine to sell in a US restaurant. She travels to the wine-growing regions of Europe (France, Spain, Italy, Germany, Portugal, etc.) and must exchange dollars for euros to pay for her purchases.

Our three visitors represent a microcosm of the transactions that take place in the foreign exchange market every day. Households and firms buy euros to pay for their imports of goods and services (souvenirs, wine, etc.). Many different goods and services are produced in Europe and sold in the United States. Some are imported by retailers, others by specialist import-export firms, and still others by individuals, but in all cases there is an associated purchase of euros using dollars.

The demand for euros also arises from financial investment by households, firms, and financial institutions. For example, a wealthy private investor in the United States may purchase stock issued by a company in Europe. To buy that stock, the US investor sells dollars and buys euros. In practice, such transactions are typically carried out by financial institutions that undertake trades on behalf of households and firms.

Most exchanges of dollars for euros do not actually entail someone traveling to Europe. Think about the foreign currency needs of a large multinational firm that produces goods and services in Europe but sells its output in the United States. The company naturally needs euros to pay workers and suppliers in Europe. Since it sells goods and thus earns revenues in dollars, the company must convert from dollars to euros very frequently. But you will not see the company's chief financial officer in an airport line to exchange money. Instead, such currency operations are conducted through financial institutions, such as commercial banks.

Because of all these transactions, there are very active and sophisticated markets in which currencies are traded. We can represent these markets using the familiar **supply-and-demand** framework. Figure 24.4.1 shows a picture of the market where euros are bought and sold. Buyers from the United States buy euros with dollars, and European traders sell euros in exchange for dollars.Of course, it is not literally the case that everyone who is buying is from the United States and that everyone selling is from Europe. If you have dollars, you can buy euros; if you have euros, you can sell them for dollars. But it is simpler to explain if we think of Europeans selling euros and Americans buying them. The supply and demand curves refer to the object being traded—euros. Thus the quantity of euros is shown on the horizontal axis. The price on the vertical axis is in dollars.

This market is just like any other you encounter. The demand curve is downward sloping: as the price of euros increases, the quantity of euros demanded decreases. This is the **law of demand** at work. As the price of euros increases, people in the United States will find that goods and services produced in Europe are more expensive. For example, suppose that 1 euro costs $1, and a Mercedes automobile costs EUR 50,000.There is an established set of three-letter symbols for all the currencies in the world. Euros are denoted by EUR, US dollars are denoted by USD, Australian dollars are denoted by AUD, and so forth. In this book we use the familiar $ symbol for US dollars and the three-letter symbols otherwise. A list of the currency codes can be found at http://www.xe.com/iso4217.php. Then its cost in dollars is $50,000. Now imagine that euros become more expensive, so that EUR 1 now costs $2. You now need $100,000 to buy the same Mercedes in Europe. So an increase in the price of euros means that Americans choose to buy fewer goods and services produced in Europe. Exactly the same logic tells us that an increase in the price of the euro makes European assets look less attractive to investors. A German government bond, a piece of real estate in Slovenia, or a share in a Portuguese firm might look like good buys when the euro costs $1 yet seem like a bad idea if each euro costs $2.

The supply curve also has a familiar upward slope. As the price of euros increases, more people in Europe sell their euros in exchange for dollars. They do so because with the higher dollar price of euros, they can obtain more dollars for every euro they sell. This means that they can buy more US goods and services or dollar-denominated financial assets.

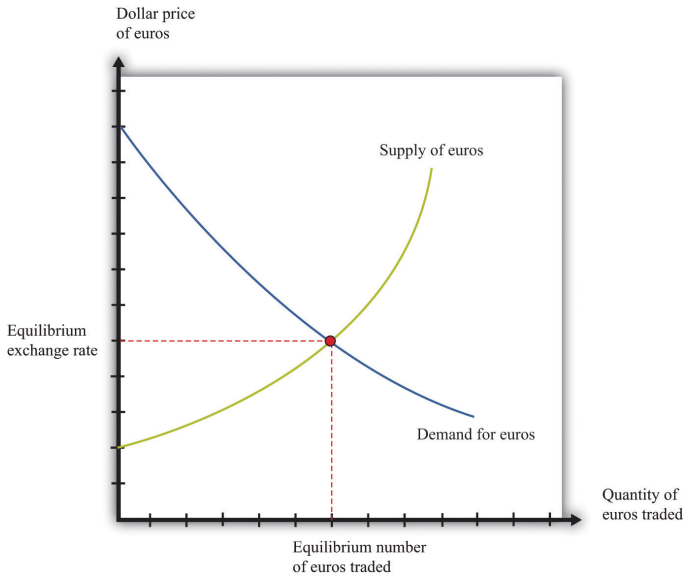


Figure 24.4.1: The Market for Euros. This diagram shows the foreign exchange market in which euros are bought and sold. As the price of euros (in dollars) increases, more euros are supplied to the market, but fewer euros are demanded.

The price where supply equals demand is the **equilibrium** exchange rate. (The market also shows us the equilibrium number of euros traded, but here we are more interested in the price of the euro.)

> **Toolkit: Section 31.9**
>
> The foreign exchange market is an example of a market that we can analyze using the tool of supply and demand. You can review the supply-and-demand framework and the meaning of equilibrium in the toolkit.

Arbitrage with Two Currencies

So far, we have talked about buying foreign currencies to purchase either assets or goods and services. Another reason to buy foreign currencies is in the hope that you could make money by trading them. Let us think about how you might try to make money in the foreign exchange market. You might start with some dollars and exchange them for euros. Then you could take those euros and exchange them for dollars again. Is it possible that, by doing this, you could end up with more money than you started with? Could you buy euros cheaply and then sell them at a high price, thus making a profit?

Begin by supposing that dollars and euros are only two currencies in the world, and there are only two economies: the United States and Europe (a shorthand for "those European countries that use the euro"). Imagine that there are two separate markets: in the euro market, the price of 1 euro is $2; in the dollar market, the price of one dollar is EUR 1. With these two prices, there is money to be made by buying and selling currencies. Start with 1 euro. Sell that euro in the market for euros and obtain $2. Use those dollars to buy euros in the market for euros and obtain 2 euros. Now we are talking business: you started with 1 euro, made some trades, and ended up with 2 euros.

There is, of course, a catch. The prices that we just suggested would not be consistent with equilibrium in the foreign exchange markets. As we have just seen, there is a simple recipe for making unlimited profit at these prices, not only for you but also for everyone else in the market. What would happen? Everyone would try to capitalize on the same opportunity that you saw. Those with euros would want either to sell them in the euro market—because euros are valuable—or to use them to buy dollars in the dollar market—because dollars are cheap. Those with dollars, however, would not want to buy expensive euros in the euro market, and they would not want to sell them in the dollar market. Hence, in the euro market, the supply of euros would shift rightward, and the demand for euros would shift leftward. The forces of supply and demand would make the dollar price of euros decrease. In the dollar market, the supply of dollars would shift leftward, and the demand for dollars would shift rightward, causing the euro price of dollars to increase.

The mechanism we just described is arbitrage at work again. The arbitrage possibility between the dollar market for euros and the euro market for dollars disappears when the following equation is satisfied:

$$price\ of\ euro\ in\ dollars \times price\ in\ dollar\ in\ euros\ = 1.$$

When this condition holds, there is no way to buy and sell currencies in the different markets and make a profit. As an example, suppose that EUR 1 costs $2 and $1 costs EUR 0.5. These prices satisfy the equation because $2 \times 0.5 = 1$. Imagine you start with $1. If you use it in the dollar market for euros to buy euros, then you will have EUR 0.50. If you then use these in the euro market for dollars to buy dollars, you will get $2 for each euro you supply to the market. Since you have half of a euro, you will end up with $1, which is what you started with. There is no arbitrage opportunity.

By now you have probably realized that there is a close connection between the market for euros and the market for dollars (where dollars are bought and sold using euros). Whenever someone buys euros, they are selling dollars, and whenever someone sells euros, they are buying dollars. In our two-country, two-currency world, the market for euros and the market for dollars are exactly the same market, just looked at from two different angles.

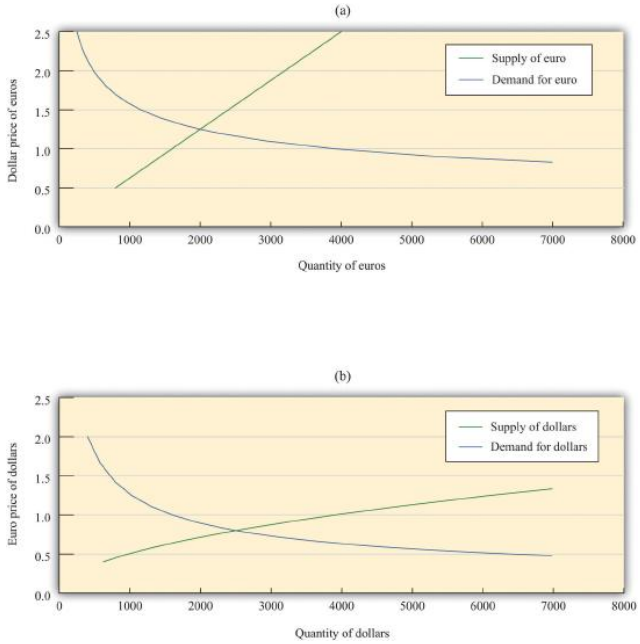
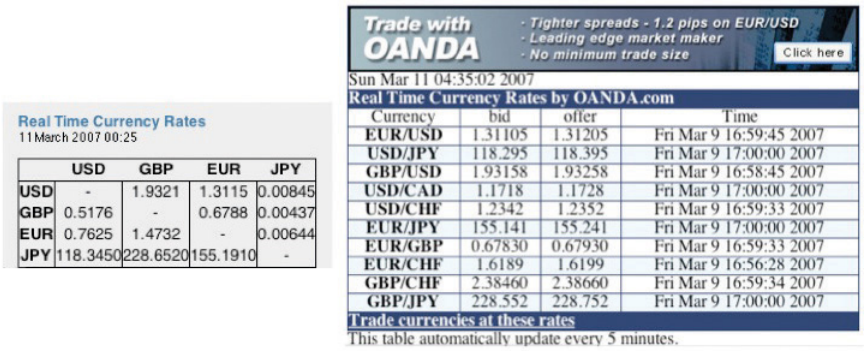


Figure 24.4.2: The Market for Euros and the Market for Dollars. Here we show the market where euros are bought and sold with dollars (a) and the market where dollars are bought and sold with euros (b). Because of arbitrage, *these are just two different ways of looking at the same market.*

We illustrate this in Figure 24.4.2. In part (a) of Figure 24.4.2, we show the market where euros are bought and sold, and in part (b) of Figure 24.4.2 the market where dollars are bought and sold. The supply curve for dollars is just the demand curve for euros, and the demand curve for dollars is the same as the supply curve for euros. For example, suppose 1 euro costs $2. From part (a), we see that, at this price, people would supply EUR 3,200. In other words, there are individuals who are willing to exchange EUR 3,200 for $6,400. If we think about this from the perspective of the market for dollars, these people would demand $6,400 in the market when $1 costs EUR 0.50—and, indeed, we see that this is a point on the demand curve in part (b). The market is in equilibrium when EUR 1.00 costs $1.25, or equivalently when $1 costs EUR 0.80. At this exchange rate, holders of dollars are willing to give up $2,500, and holders of euros are willing to give up EUR 2,000.

Arbitrage with Many Currencies

We live in a world with many different currencies, not just two. Figure 24.4.3 shows some exchange rates from http://www.oanda.com, a site that provides current and historical data on exchange rates and that is also an online market where you can trade currencies. So, on March 11, 2007, just after midnight, the price of a euro in dollars was 1.3115. At the same time, the price of a dollar in British pounds was 0.5176.



Real Time Currency Rates
11 March 2007 00:25

	USD	GBP	EUR	JPY
USD	-	1.9321	1.3115	0.00845
GBP	0.5176	-	0.6788	0.00437
EUR	0.7625	1.4732	-	0.00644
JPY	118.3450	228.6520	155.1910	-

Trade with OANDA
- Tighter spreads - 1.2 pips on EUR/USD
- Leading edge market maker
- No minimum trade size
Click here

Sun Mar 11 04:35:02 2007

Real Time Currency Rates by OANDA.com

Currency	bid	offer	Time
EUR/USD	1.31105	1.31205	Fri Mar 9 16:59:45 2007
USD/JPY	118.295	118.395	Fri Mar 9 17:00:00 2007
GBP/USD	1.93158	1.93258	Fri Mar 9 16:58:45 2007
USD/CAD	1.1718	1.1728	Fri Mar 9 17:00:00 2007
USD/CHF	1.2342	1.2352	Fri Mar 9 16:59:33 2007
EUR/JPY	155.141	155.241	Fri Mar 9 17:00:00 2007
EUR/GBP	0.67830	0.67930	Fri Mar 9 16:59:33 2007
EUR/CHF	1.6189	1.6199	Fri Mar 9 16:56:28 2007
GBP/CHF	2.38460	2.38660	Fri Mar 9 16:59:34 2007
GBP/JPY	228.552	228.752	Fri Mar 9 17:00:00 2007

Trade currencies at these rates

This table automatically update every 5 minutes.

Figure 24.4.3: Exchange Rates. These tables come from http://www.oanda.com. *The table on the left shows exchange rates among four currencies and the table on the right shows the rates at which you can actually conduct trades at this site.*

If you look at the table on the left side of Figure 24.4.3, you see that it provides both the dollar price of the euro and the euro price of the dollar (and similarly for the other currency pairs). Tables such as this one have already built in the arbitrage condition, so you cannot keep buying and selling the same currency in exchange for dollars and make money.

When there are multiple currencies, we can imagine more complicated trading strategies. As an example, consider the following string of transactions.

1. Take a dollar and use it to buy euros.
2. Take the euros and buy Japanese yen.
3. Take the yen and buy dollars.

If you end up with more than $1, then there are profits to be made buying and selling currencies in the manner outlined here. Can you make a profit this way? The answer, once again, is no. If you could, then the markets for foreign currency would not be in equilibrium: everyone would buy euros with dollars, sell them for yen, and then sell the yen for dollars. Once again, exchange rates would rapidly adjust to remove the arbitrage opportunity.

To verify this, let us go through this series of transactions using Figure 24.4.3. One dollar will buy you EUR 0.7625. Now take these and use them to buy yen. You will get $0.7625 \times 155.1910 = JPY118.3331$. Now, use these yen to buy dollars, and you will get $118.3331 \times 0.00845 = \$0.9999$. You start with $1; you end with $1 (give or take a rounding error).

These calculations assume that there are no costs to trading foreign currencies. In practice, there are costs involved in these exchanges. A traveler arriving at an airport in need of local currency does not see rates posted as in the left-hand table in Figure 24.4.3. Instead, they see something that looks like the right-hand table, where rates are posted in two columns: bid (buying) and offer (selling). The bid is a statement of how much the currency seller is willing to pay in local currency for the listed currency. The offer column is the price in local currency at which the seller is willing to sell to you. Naturally, the offer price is bigger than the bid: the seller buys currencies at a low price and sells them at a high price. The difference between the bid and offer prices is called the *spread*. The existence of the spread means that if you try to buy and sell currencies with the dealer, you will actually lose money. At the same time, the spread creates a profit margin for the dealer and thus pays for the service that the dealer provides.

Arbitrage with Goods and Currencies

We have talked about arbitrage with goods and arbitrage with foreign currencies. We can also put the two together to study the prices of goods that are traded across international borders. Arbitrage of goods from one country to another is a bit more complicated because it involves buying and selling currencies as well as goods. To see how this works, imagine you are going on a trip to Europe. You are allowed two suitcases filled with belongings free of charge on the airplane. What about filling a suitcase full of new blue jeans, transporting them to Europe, and then selling them there? Could you make money that way?

Suppose that the dollar price of 1 euro is $1.50. Further, suppose that the price of a pair of blue jeans is $70.00 in the United States and EUR 50.00 in Paris. Consider the following sequence of actions.

1. Take $70 out of your pocket and buy a pair of blue jeans.
2. Travel with these blue jeans to Paris.
3. Sell the jeans for euros.
4. Buy dollars with your euros.

The question is whether you can make money in this way. The answer is given by how many dollars you will have in your pocket at the end of these steps. When you sell the jeans in Paris, you will have EUR 50.00. If the dollar price of euros is $1.50, then by selling the jeans in Paris you will get 50 × $1.50 = $75. This is a profit of $5 for each pair of jeans—you are in business.

Once again, the opportunity for arbitrage suggests that this situation is unlikely to persist. Entrepreneurs will buy jeans in the United States, take them to Paris, and sell them there. Market forces in three different markets will work to eliminate the profit. First, the activity of arbitrageurs will increase the demand for jeans in the United States, causing the US price of jeans to increase. Second, the increased supply of jeans in Paris will cause the price there to decrease. And third, there will be an increased supply of euros in the foreign exchange market, which will cause the euro to depreciate. This is shown in Figure Runaway argument for \eatSpaces?

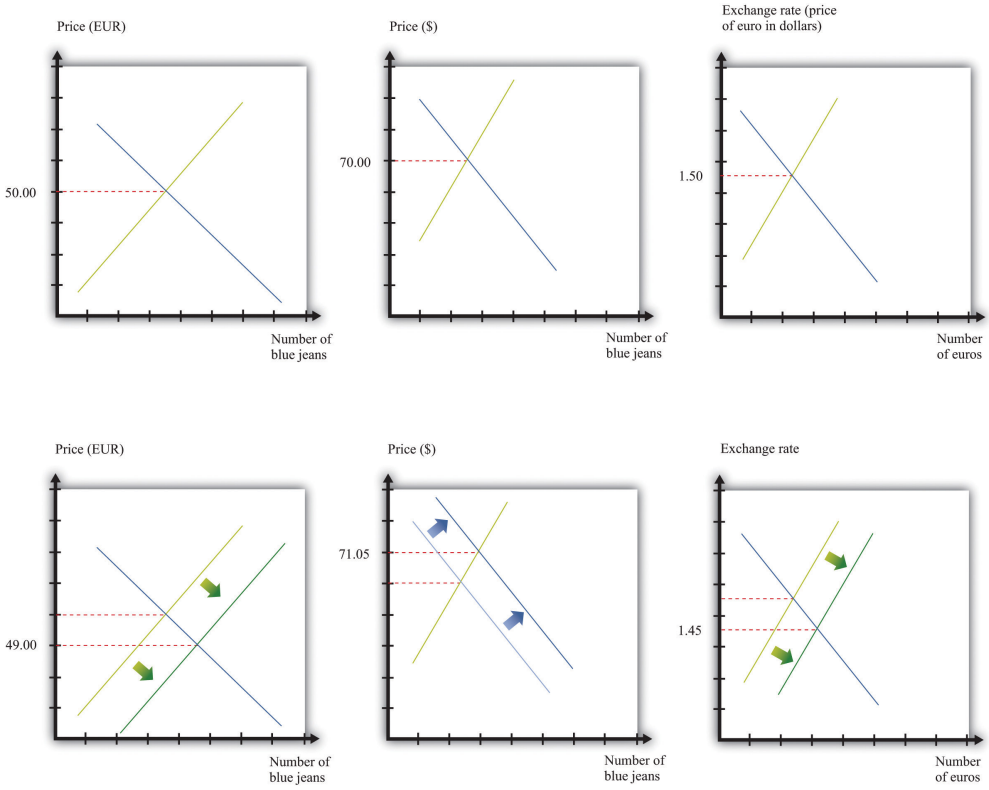


Figure 24.4.4: International Arbitrage Restores the Law of One Price. When blue jeans cost $70 in the United States and EUR 50 in France, and the exchange rate is $1.50 per euro, arbitrageurs can make a profit by importing blue jeans to Europe from the United States.

These price changes continue until there are no profits to be made by arbitrage. Exactly how much of the adjustment will take place in each market depends on the slopes of the supply and demand curves. In Figure 24.4.4, we have drawn the new equilibrium as follows: blue jeans cost EUR 49 in Europe and $71.05 in the United States; and the exchange rate is $1.45 per euro. At these prices,

$$price\ of\ blue\ jeans\ in\ dollars\ =\ price\ of\ blue\ jeans\ in\ euros \times price\ of\ euro\ in\ dollars,$$

and there is no longer any possibility of arbitrage. This is another illustration of the law of one price. If we were literally talking just about arbitrage in blue jeans, most of the adjustment would take place in the markets for blue jeans in the United States and Europe, and there would be a negligible effect on the exchange rate. But if the same kinds of arbitrage opportunities exist for lots of goods, then there will be an impact on the exchange rate as well.

For **tradable goods**, the law of one price says that the

$$dollar\ price\ of\ good\ =\ euro\ price\ of\ good \times dollar\ price\ of\ euro.$$

When this condition holds, there are no arbitrage profits to be gained by purchasing the good with dollars, selling it for euros, and then buying dollars with euros. Likewise, if this condition holds, there are also no arbitrage profits from purchasing the good with euros, selling it for dollars, and then buying euros with dollars. In general, we expect that such arbitrage will occur very quickly. There are no profits to be made from arbitrage when the law of one price holds.

The *Economist* has kept track of the price of a McDonald's Big Mac in a number of countries for many years, creating something they call the "the Big Mac index." Table 24.4.1 "The *Economist*'s Big Mac Index, July 2011" contains some of their data. The last column of Table 24.4.1 "The *Economist*'s Big Mac Index, July 2011" gives the price of a Big Mac in each selected country in July 2011, converted to US dollars at the current exchange rate. That is, the last column is calculated by dividing the local currency price (the second column) by the exchange rate (the third column). A Big Mac costs $4.07 in the United States but more than twice as much in Norway. China is a real deal at only $1.89.

Table 24.4.1: The *Economist*'s Big Mac Index, July 2011

Country	Local Currency Price of Big Mac	Local Currency Price of a Dollar	Price in US Dollars
United States	USD 4.07	1	4.07
Norway	NOK 45	5.41	8.31
Euro Area	EUR 3.44	0.70	4.93
Czech Republic	CZK 69.3	17.0	4.07
China	CNY 14.7	6.45	1.89

Source: "The Big Mac Index: Currency Comparisons, to Go," *Economist* online, July 28, 2011, accessed August 2, 2011, http://www.economist.com/blogs/dailychart/2011/07/big-mac-index.

The price differentials in this table violate the law of one price: there is (apparently) profit to be made by buying Big Macs at a low price and selling them at a high price. Applying the principle of arbitrage, we should all be flying to China, buying Big Macs, traveling to Norway, and selling them on the streets of Oslo. Of course, there are a few small problems with this scheme, such as the following:

- It is expensive to fly back and forth between China and Norway.
- There is a limited capacity for transporting Big Macs on the airplane.
- The quality of the Big Mac might deteriorate while it is being transported.
- You might not be permitted to import meat products from China into Norway.
- You might have to pay taxes when you bring Big Macs into Norway.
- It might be tough to open a McDonald's in Oslo.

This long list easily explains the deviations from the law of one price for Big Macs. Similar considerations explain why the law of one price might not hold for other goods. The law also does not apply to services, such as tattoos, since these cannot be imported and exported. The law of one price is most applicable to goods that are homogeneous and easily traded at low cost. Economists use the law of one price as a guide but certainly do not expect it to hold for all products in all places.

Using the Law of One Price to Understand the Exchange Rate

There is another way to interpret the finding that Big Macs do not cost the same in each country. The *Economist* uses this information to draw conclusions about the values of different currencies and how these values are likely to change over time.

From this perspective, the Big Mac is more expensive in Europe than in the United States because dollars are cheap in Europe. Put differently, we say that the dollar is **undervalued** relative to the euro. If the price of a dollar in euros were 0.85 instead of 0.70, then a Big Mac would cost the same in the United States and Europe. Completely equivalently, we can say that that the euro is **overvalued** relative to the dollar. With this in mind, we might expect the undervalued dollar to increase in value relative to the euro. That is, we would expect the price of a dollar in euros to increase. Similarly, we would conclude that the Norwegian kroner is overvalued relative to the dollar, the Chinese yuan is undervalued, and the Czech Koruna is neither overvalued nor undervalued.

To see how this works more generally, look back at our arbitrage condition for blue jeans. If we divide both sides by the price of blue jeans in euros, we get

This equation says that, according to the law of one price, the dollar price of the euro should equal the dollar price of blue jeans divided by the price of blue jeans in euros. This is exactly the kind of calculation that underlies the Big Mac index, only with blue jeans instead of Big Macs. Equivalently, the law of one price says that the

Suppose we think about this equation applying (approximately) to all goods and services. We can then get a better prediction of the exchange rate by looking at a general price index in each country:

Because of all the reasons why the law of one price does not literally hold, economists certainly do not expect this equation to give an exact prediction of the exchange rate. Nevertheless, it can provide a useful indication of whether a currency is undervalued or overvalued.

A currency is undervalued if, following this equation, its price is too low compared to the ratio of price levels in the two countries. A currency is overvalued if, following this equation, its price is too high compared to the ratio of price levels in the two countries. As in our discussion of the euro, if a currency is overvalued, then we would expect its value to decrease over time. This is called a **depreciation** of the currency. Likewise, we would expect the price of an undervalued currency to increase over time. This is called an **appreciation** of the currency.

The market forces behind these currency movements come from the buying and selling of currencies for trading purposes. If the Chinese yuan is undervalued, goods produced in China will be relatively cheap in US dollars. The demand for Chinese exports will be high, and this will lead to a large demand for the yuan. Eventually the dollar price of the yuan will increase—that is, the yuan will appreciate, and the dollar will depreciate.

Changes in the Exchange Rate

Even though the law of one price does not literally hold for all goods and services, it reminds us that the value of $1 in the United States is linked to its value in the rest of the world. As a result, we expect that price level *changes* are likely to lead to changes in the exchange rate. We see this more clearly if we write our previous equation in terms of growth rates. Using the formula for growth rates, we find the following:

$$\begin{aligned} growth\ rate\ of\ price\ of\ dollar\ in\ euros\ &= \ growth\ rate\ of\ price\ of\ European\ bundle\ of\ goods \\ &- growth\ rate\ of\ price\ of\ US\ bundle\ of\ goods. \end{aligned}$$

Toolkit: Section 31.21

The formulas for using growth rates can be found in the toolkit.

If the bundle of goods in each country corresponds roughly to the goods in the **Consumer Price Index (CPI)**, then the growth rate of these prices corresponds to the **inflation rate**. The growth rate of the exchange rate is just another term for the percentage appreciation of the currency. Thus we get the following:

$$percentage\ appreciation\ of\ the\ dollar \approx European\ inflation\ rate - US\ inflation\ rate.$$

So, if the inflation rate in the United States is higher than it is in Europe, we expect the euro price of the dollar to decrease. We expect depreciation of the dollar if US inflation exceeds European inflation. Inflation reduces the real value of money domestically; it will also tend to reduce the value of money in terms of what it can purchase in the rest of the world. This makes sense. If our currency is becoming less valuable at home, then we should also expect it to become less valuable in the rest of the world.

The Real Exchange Rate

The law of one price is connected to another measure of the exchange rate—the **real exchange rate**. This exchange rate is a measure of the price of goods and services in one country relative to another *when prices are expressed in a common currency*. It is about exchanging goods, rather than money, across countries.

The real exchange rate between the United States and Europe is given as follows:

You can think of the real exchange rate as the number of units of European gross domestic product (GDP) you can get for one unit of US GDP.Let us check the units of the real exchange rate. The US price level over the European price level is in dollars/euros: it is the price of a unit of US real GDP divided by the price of a unit of European real GDP. The nominal exchange rate is measured in euros per dollar. Thus the units are as follows:The dollars and the euros cancel out in this expression, so the real exchange rate is just a number. For example, if the price level in the United States is $1,600, the price level in Europe is EUR 400, and the price of dollars in euros is EUR 0.5, then the real exchange rate is as follows:

One unit of US GDP will get you two units of European GDP.

The real exchange rate is intimately linked to the law of one price. The easiest way to see this is to suppose that we measure US real GDP and European GDP in the same units: that is, suppose we use the same bundle of goods in each case. We know that the law of one price should hold for tradable goods—that is, goods for which arbitrage is possible and practical. If every good that went into GDP were tradable, then the law of one price would hold for every good, and the real exchange rate would equal 1. If the real exchange rate was not 1, you could make arbitrage profits by buying and selling "units of GDP."

As before, suppose the US price level is $1,600, the European price level is EUR 400, and the nominal exchange rate (dollars per euro) is 0.5. Imagine that US GDP and European GDP measure the same bundle of (tradable) goods. Then you could take $800 and buy EUR 400. With these euros, you could buy a basket of goods in Europe. You could sell this basket in the United States for $1,600. The law of one price is violated. We would expect the following:

- Prices in the United States would increase.
- Prices in Europe would decrease.
- The nominal exchange rate would depreciate (the dollar would become less valuable).

Because arbitrage is not possible for all goods and services, we do not expect—nor do we observe—the real exchange rate to be exactly one. But this benchmark is still useful in understanding movements in the real exchange rate.

The Real Exchange Rate in Action

The real exchange rate matters because it is the price that is relevant for import and export decisions. Suppose you are trying to decide between buying a mobile phone manufactured in the United States and one manufactured in Finland. If the dollar appreciates against the euro, then the US phone retailer needs fewer dollars to purchase euros, so Finnish phones will be cheaper in US stores. If prices decrease in Finland, the imported phone again becomes relatively cheaper. If prices increase in the United States, the US phone will be more expensive. In other words, increasing prices in the United States, decreasing prices in Finland, and appreciation of the dollar all make you more likely to buy the imported phone rather than the domestically produced phone.

More generally, anything that causes the real exchange rate to increase will make imports look more attractive compared to goods produced in the domestic economy. Examined from the point of view of Europe, the same increase in the real exchange rate makes US goods look more expensive relative to goods produced in Europe, so Europeans will be likely to import fewer goods from the United States. An increase in the real exchange rate therefore leads to an increase in US imports and a decrease in US exports—that is, it leads to a decrease in **net exports**.

The real exchange rate can and does vary substantially over time. Argentina in the 1990s provides a nice illustration of real exchange rates in action. We discuss this in more detail in Chapter 26. Argentina had a **currency board** during this period. Under a currency board, a country maintains a fixed exchange rate by backing its currency completely with another currency. Although Argentina did have its own currency (the Argentine peso), each peso in circulation was backed by a US dollar held by the Argentine central bank. You could at any time exchange pesos for dollars at a nominal exchange rate of 1.

Figure 24.4.5 shows what happened to prices in Argentina and the United States over this period. Look at 1992–95. Both countries had some inflation. But prices were increasing faster in Argentina than in the United States. The real exchange rate (Argentina–United States) is given by

because the price of the peso in dollars was 1. Therefore the real exchange rate appreciated as Argentine inflation outpaced US inflation.

The appreciation of the real exchange rate meant that Argentine goods became more expensive in other countries, so Argentine exports became less competitive. (The problem was compounded by the fact that the US dollar [and hence the peso] also appreciated against the currencies of neighboring countries such as Brazil.) Without the currency board, it would have been possible for the nominal exchange rate (price of the peso in dollars) to decline, offsetting the effects of the inflation rate. Instead, this appreciation of the real exchange rate ended up causing substantial economic problems in Argentina in the 1990s. In the second half of the decade, the real exchange rate began to depreciate because the inflation rate in Argentina was lower than in the United States. The appreciation at the start of the decade had been so large, however, that the real exchange rate in 1999 was still higher than it had been in 1992.

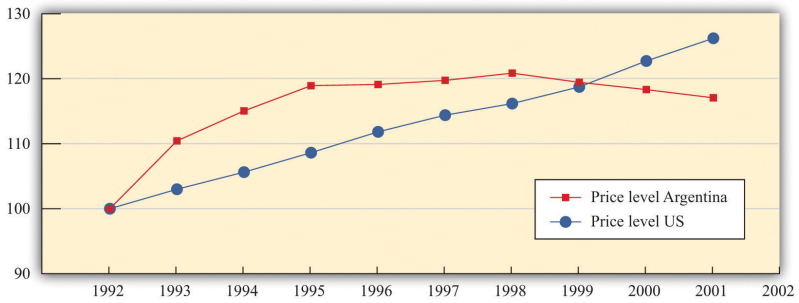


Figure 24.4.5: The Real Exchange Rate in Argentina. Argentina's real exchange rate appreciated between 1992 and 1995 because the nominal (US dollar–Argentine peso) exchange rate was constant and equal to one, and the price level increased more rapidly in Argentina than in the United States.

Source: CPI inflation figures from International Monetary Fund World Economic Outlook database (www.imf.org/external/pubs/ft/weo/2006/01/data/dbcselm.cfm?G=205) and Bureau of Labor Statistics (www.bls.gov).

If countries want to have a permanently fixed exchange rate, there is an option that is more radical than a currency board. Countries can decide to adopt a common currency, like the European countries that adopted the euro. There are several reasons why countries might decide to take such a course of action. The first advantage of a common currency is that it enhances the role of money as a medium of exchange. There is no longer a need to exchange one currency for another, making it easier to trade goods and services across countries. People do not have to deal with the inconveniences of exchanging currencies: individuals do not have to exchange cash at airports, and firms do not need to manage multiple currencies to conduct international business. In the jargon of economics, a single currency removes transaction costs. These costs might be individually small, but they can add up when you consider just how many times households and firms needed to switch from one of the euro area currencies to another.According to studies supporting a common currency, these gains from reduced transactions costs were substantial. One of the key analyses was the Delors report. A summary of that report is available at "Phase 3: the Delors Report," European Commission, October 30, 2010, accessed August 22, 2011, http://ec.europa.eu/economy_finance/euro/emu/road/delors_report_en.htm. A complete report on the history of the euro is available at "One Currency for One Europe: The Road to the Euro," European Commission, 2007, accessed August 22, 2011, http://ec.europa.eu/economy_finance/publications/publication6730_en.pdf.

One way to picture this advantage is to imagine the reverse. Suppose, for example, that each state in the United States decided to adopt its own currency. Trade across state lines would become more complicated and more costly. Even more starkly, imagine that your hometown had its own currency, so you had to exchange money whenever you traveled anywhere else.

A second advantage of a single currency is that it makes business planning easier. A firm in Belgium can write a contract with another firm in Spain without having to worry about the implications of currency appreciation or depreciation. Thus an argument for the move to a single currency was that such a change was likely to encourage trade among countries of the European Union. Again, imagine how much more complicated business would be in the United States if each state had its own freely floating currency.

Finally, a common currency enhances capital flows. Just as it is easier for businesses to trade goods and services, it is also easier for investors to shift funds from country to country. With a common currency, investors do not have to pay the transactions costs of converting currencies, and they no longer face the uncertainty of exchange rate changes. When capital flows more easily across borders, investment activity is more productive, enhancing the growth of the countries involved.

Key Takeaways

- The nominal exchange rate is the price of one currency in terms of another. The real exchange rate compares the price of goods and services in one country to the cost of these goods and services in another country when all prices are in a common currency.
- From the law of one price, a tradable good in one country should have the same price as that same good in another country when the goods are priced in the same currency. This means that the exchange rate is equal to the ratio of the prices expressed in the two different currencies. Put differently, by the law of one price, the real exchange rate between tradable goods should be 1.

Exercises

1. If the price of a euro was $2 and the price of a dollar was 1 EUR, how would you make a profit?
2. If goulash sells for either 1,090 forint or 4.40 euro, what is the price of the forint in terms of the euro? Do the two prices of cabbage quoted in Figure 24.4.1 yield a different euro price for the forint? Is there an arbitrage possibility here (or elsewhere on the menu)?

This page titled 24.4: Using Money to Buy Other Monies- Exchange Rates is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

24.5: Using Money to Buy Assets- Interest Rates 24.5: Using Money to Buy Assets- Interest RatesEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What is the difference between nominal interest rates and real interest rates?
2. What is the yield curve?
3. What is the Fisher equation?

We have now discussed how you could use your $100 to buy goods and services or the money of another country. You can also use your money to buy money in the future. When we say this, we are simply describing a familiar transaction in an unfamiliar way: we are talking about saving. If you put money in a bank, then you are buying money in the future with money you give up today. When you save in this way, you become a participant in the **credit markets (or loan markets)**.

Toolkit: Section 31.24

A credit market (or loan market) brings together suppliers of credit, such as households who are saving, and demanders of credit, such as businesses and households who need to borrow. You can review the credit market in the toolkit.

Arbitrage with Credit and Assets

Suppose you do not want to spend your $100 until next year. You could just put the money under your mattress, but a better option is to find some way of getting more than $100 next year. One way to do this is to lend your money to someone else. For you, this might simply mean taking it to your bank and putting it in your savings account. When you do that, you are making a loan to the bank. Of course, the bank probably will not leave the money in its vault; it will lend that money to someone else. Banks and other financial institutions act as intermediaries between those who want to save and those who want to borrow.

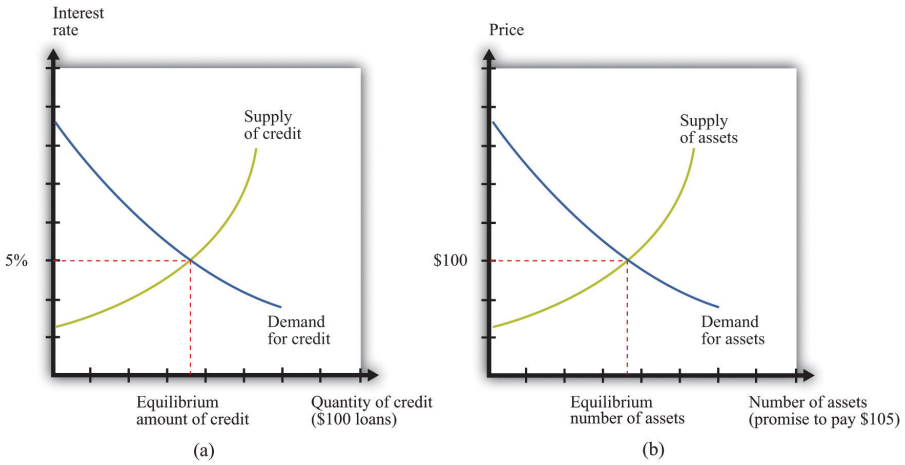


Figure 24.5.1: The Credit Market and the Asset Market. Credit markets and asset markets are two ways of looking at the same market: the market for $100 loans (a) with an equilibrium interest rate of 5 percent is the same as the market for an asset that promises to pay $105 in a year's time (b).

The credit market brings together the suppliers and demanders of credit, and the **nominal interest rate** is the price that brings demand and supply into balance. The supply of credit increases as the interest rate increases: as the return on saving increases, households will generally save more and thus supply more funds to the credit market. The demand for credit decreases as the interest rate increases: when it is expensive to borrow, households and firms will borrow less. At the equilibrium interest rate, the quantity of credit supplied and the quantity of credit demanded are equal. This is shown in part (a) of Figure 24.5.1.

There is another way to look at credit markets. Borrowers get money today in exchange for a promise to pay money later. Lenders purchase those promises by giving up money today. Instead of asking how much the interest rate is for a given $100 loan, we could ask what people would be willing to pay today for the right to receive $105 in a year's time.

The market for the promise to pay $105 in a year is illustrated in part (b) of Figure 24.5.1. The units on the horizontal axis are $105 payments. These are assets: buyers are purchasing a piece of paper that is a promise to deliver $105 in a year's time. The price on the vertical axis is the current price of that asset.

The nominal rate of return on an asset is the amount that you obtain, in percentage terms, from holding the asset for a year. In the case of the simple one-year asset we are considering, the return is given as follows:

We can also rearrange this to give us the price of the asset:

Notice what happens when we look at the market in this way. Buyers have become sellers, and sellers have become buyers. Borrowers are sellers: they sell the promise to pay. Lenders are buyers: they purchase the promise to pay. If we are looking at the same group of buyers and sellers as before, then the current equilibrium price of this asset would be $100.

The nominal interest rate and the nominal rate of return defined through these two markets must be the same. If not, there would be an arbitrage possibility. Imagine, for example, that the interest rate is 5 percent but the price of the asset is $90. In this case, the rate of the return on asset is which is 22.2 percent. So you could make a lot of money by borrowing at a 5 percent interest rate and then purchasing the promises to pay $110 at price of $90.

If you could do this, so also could many major financial institutions—except that they would operate on a much larger scale, perhaps buying millions worth of assets and borrowing a lot in credit markets. So the demand for credit would shift outward, as would the demand for assets. This would cause the interest rate to increase and the asset price to increase, so the rate of return on the asset decreases. This would continue until the arbitrage opportunity disappeared.

In summary, we would say there is no arbitrage opportunity when the

$$nominal\ rate\ of\ return\ on\ asset = nominal\ interest\ rate.$$

The rate of return on the asset, in other words, is equivalent to the interest rate on the asset. Equivalently this means that

In the second line we replaced $(1 + nominal\ interest\ rate)$ with the **nominal interest factor**. The two are equivalent, but sometimes we find it more convenient to work with interest factors rather than interest rates.

The argument that we have just made should seem familiar. It is analogous to the argument for why there cannot be distinct dollar-euro and euro-dollar markets; they are just ways of looking at the same asset. Likewise, we can think of the sale of any asset as equivalent to borrowing, while for any example of credit we can also think of there being an underlying asset.

Different Assets

Very often economists and others talk about "the" interest rate, as if there were just a single asset in the economy. In fact, there are many different assets that you could buy with your $100, each with an associated interest rate. The following are various assets that you might purchase:

- **Currency and coin**. To begin with, your $100 is itself an asset. If you put the money under a mattress and retrieve it after a year, it is very easy to calculate the nominal interest rate on $100. If you give up $100 today, you will get exactly $100 back next year. The nominal interest rate is zero.
- **Bank deposits.** Bank deposits are also an asset. If you put your money in the bank, you are extending credit to the bank. Depending on the type of bank account, you may or may not earn interest on your deposits.
- **Foreign exchange.** The money of other countries is likewise an asset. You can take dollars today and use them to purchase, say, euros or Japanese yen (JPY). Even in this case, there is a rate of return. For example, suppose that today you can buy JPY 100 with $1. Suppose also that in a year's time, there are JPY 90 to the dollar. Then with your JPY 100, you can buy $1.11 (100/90 = 1.11). You obtained a nominal rate of return of 11 percent.
- **Gold and other precious metals.** You could take your $100 and use it to buy gold. Unless you are a dentist or a jeweler, you will not have any direct use for the gold; you simply keep it and resell it at some future date. The rate of return on gold is purely a matter of what happens to the price of gold. If the price of gold (in dollars) increases by 10 percent, then you get a 10 percent rate of return.
- **Government bonds.** A government bond is also a loan contract; if you buy a government bond, you are extending credit to the government. The bond is a promise to pay a certain amount at some future date. Because the loan will be paid off a number of years in the future, it is slightly more complicated to calculate the interest rate.
- **Shares.** Another example of an asset is a share in a company, such as Dell Inc. If you purchase a Dell share, you have bought the right to a share in Dell's profits. In this case, you expect not only one payment at a specified future date but also a sequence

of payments whenever Dell pays out dividends. Notice that there is also a lot of uncertainty here: you do not know, when you purchase the share, how big these payments will be. The implied interest rate is therefore uncertain as well.

- **Real estate.** If you purchase a house, you own yet another kind of asset. The value of a house comes from the fact that people can live in it. If you rent your house out, then it gives you a flow of income, much like a share in a company. If you live in your house, then you consume that flow of services, but we still think of the house as an asset because at any time you can sell your house and transfer that flow of services to someone else.

We could list many more assets, but you should be getting the general idea. Most of these assets are more complicated than the simple one-year credit contract with which we began. For one thing, they often involve a whole stream of repayments at different dates, rather than just one repayment. For another, the amounts of these payments may be uncertain.

In 24.2 Section "What Is Money?", we pointed to the different characteristics and functions of money. For most of us, the word *money* conjures up images of currency and coins. But some of the other assets that we listed also can perform more or less well as money. For example, bank deposits in a checking account or with a linked debit card are portable, durable, divisible, and widely acceptable and serve as a medium of exchange. In general, any asset that performs the functions of money *is* money. Gold can be used as a store of value and perhaps also as a unit of account, but it is not often used as a medium of exchange. There are many different assets in the world, and they vary in the extent to which they perform these different functions and thus how good they are as money.

Arbitrage with Different Assets: The Term Structure of Interest Rates

We just observed that there are many different assets and thus many interest rates in the economy. But these interest rates are all linked to each other because the same people (particularly banks and other financial institutions) trade in many different markets.

One way in which assets differ is in terms of their **maturity**. To see how the returns on assets of different maturity are linked, consider two government bonds of different maturities: one-year bonds and two-year bonds. Here are two different ways you could save for two years.

1. Buy a one-year government bond. Collect the payment at the end of the year and then reinvest that money in another one-year bond.
2. Buy a two-year government bond.

There are three interest rates relevant to your choice. The first one is the current interest rate on a one-year bond. The second is the interest rate on a one-year bond *next* year. The third interest rate is the annualized nominal interest rate on a two-year government bond. An **annualized interest rate** is the interest rate earned each year on a loan that lasts many years, and the annualized interest factor is (1 + the annualized interest rate). For example, suppose that the annualized rate on a two-year loan is 6 percent. Then you would earn 6 percent per year for two years, and

$$repayment\ after\ two\ years = \$100 \times 1.06 \times 1.06 = \$112.36.$$

As you might expect, these three interest rates are connected, and we can understand how by again thinking about arbitrage. If you purchase the two-year government bond return, you get

$$100 \times (annualized\ nominal\ interest\ \ factor\ on\ two - year\ bond)^2$$

Conversely, if you purchase the two one-year bonds, you get

$$100 \times (nominal\ interest\ \ factor\ this\ year) \times (expected\ nominal\ interest\ \ factor\ next\ year).$$

Notice that we have referred to next year's interest factor as "expected." This is because when you make your decision, you do not know what the interest rate will be.

When

$$\begin{aligned}&(annualized\ nominal\ interest\ \ factor\ on\ two - year\ bond)^2\\ &= nominal\ interest\ \ factor\ this\ year \times expected\ nominal\ interest\ \ factor\ next\ year,\end{aligned}$$

the two transactions have the same return. Once again, we can appeal to an arbitrage argument to say that we expect this equation to hold. There is one twist, however. When you make this decision, you do not know for sure what the interest rate will be on one-year bonds next year. You have to make a guess. Thus this arbitrage involves some risk.

This relationship is an example of the **term structure of interest rates**, which describes the relationship between the actual and expected returns on assets that are identical except for their maturities. A version of the relationship applies to not only assets of one-year and two-year maturity but also assets of all maturities.

From the term structure of interest rates, we learn something very significant: *if the annual one-year interest rate is below the annual rate on a two-year loan, then interest rates are expected to increase in the future*. For example, if the annual one-year interest rate is 5 percent and the annual rate on two-year loans is 6 percent, this means both borrowers and lenders expect one-year interest rates to be higher than 6 percent next year. (If desired, you can calculate exactly what the expected rate is by using the previous equation.)

We can see the connection between assets of different maturities by looking at the **yield curve**.For more details and additional graphs, see J. Huston McCulloch, "The US Real Term Structure of Interest Rates: With Implicit Inflation Premium," updated October 30, 2009, accessed August 22, 2011, http://www.econ.ohio-state.edu/jhm/ts/ts.html. The yield curve shows the current annual return for assets of different maturities. Figure 24.5.2 shows the yield curve for US Treasury securities in 2011.This is an average of rates in 2011 for US Treasury securities of different maturities from http://www.econstats.com/r/rusa_ew6.htm. On the horizontal axis of the yield curve is the number of years to maturity of the asset. On the vertical axis is the current annual yield on the asset. Notice that the yield curve is upward sloping: the longer the maturity, the higher the annual interest rate. This is generally what we observe, although sometimes the yield curve is *inverted*, meaning that higher maturity debt has a lower interest rate.

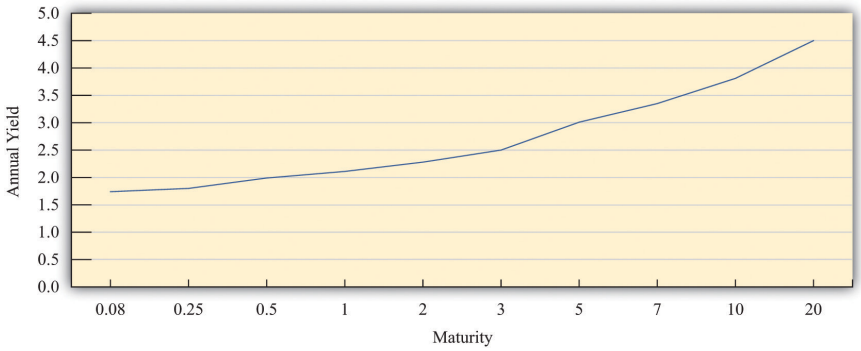


Figure 24.5.2: *The Yield Curve*

All assets are linked, not just government bonds of different maturities. Suppose that the interest rate on one-year government bonds increases. To buy these bonds, financial institutions will start selling other assets—not only bonds at other maturities but also stocks, holdings of foreign currencies, and so on. As they sell those other assets, their prices will decrease, and their rate of return will increase. An increase in the interest rate on one-year treasuries therefore increases interest rates on other assets. Thus different interest rates typically move together, and it is usually not too misleading, at least for the purposes of macroeconomics, to think about there being a single interest rate in an economy.

Arbitrage with Assets and Goods: The Real Interest Rate

The exchanges we have described so far have all been in terms of dollars. The interest rates paid on such exchanges are nominal interest rates. In a world where prices are increasing, however, the nominal interest rate does not represent the true cost of borrowing and lending.

To see why, begin by recalling that the inflation rate is defined as the percentage change in the price level. This means that the price level next year is equal to the price this year multiplied by $(1 + inflation\ rate)$.If this is not clear to you, write out the inflation rate as follows: Then add one to both sides and multiply by the price level this year. Now imagine that two individuals, Bert and Ernie, want to write a credit contract. Bert wants to borrow some money to buy a pizza. The price of a pizza this year is $10, so Ernie lends Bert $10, and they agree on a nominal interest rate for this credit arrangement. This means that next year he will repay $\$10 \times (1 + nominal\ interest\backslash\mathrm{rate})$.

We could also imagine that Bert and Ernie decide to write a different kind of contract to guarantee a return in terms of pizzas. Because this rate of return is specified in terms of goods rather than money, it is a **real interest rate**. Ernie agrees to give Bert (enough dollars to buy) 1 pizza this year in return for being repaid (enough dollars to buy) $(1 + real\ interest\ rate)$ pizzas next year. Ernie lends Bert $10 as before (the equivalent of 1 pizza). To repay this loan next year, Bert must give Ernie enough money to buy (1 + real interest rate) pizzas. The price of a pizza has increased to $\$10 \times (1 + inflation\ rate)$, so Bert must give Ernie $\$10 \times (1 + real\ interest\ rate) \times (1 + inflation\ rate)$.

If you have worked through this chapter carefully, you probably know what is coming next. Because of arbitrage, we know that these two contracts must be equivalent:

$$1 + nominal\ interest\ rate = (1 + real\ interest\ rate) \times (1 + inflation\ rate).$$

As an approximation, this equation implies that theTo see this, multiply out the right-hand side and subtract $1 from each side to obtain nominal interest rate = real interest rate + inflation rate + real interest rate × inflation rate. Now, if the real interest rate and the inflation rate are small numbers, then when we multiply them together, we get a very small number that can be safely ignored. For example, if the real interest rate is 0.02 and the inflation rate is 0.03, then their product is 0.0006, and our approximation is about 99 percent accurate.

$$nominal\ interest\ rate \approx real\ interest\ rate + inflation\ rate.$$

This relationship is called the **Fisher equation**.

Toolkit: Section 31.8

Nominal interest rates and real interest rates are related by the Fisher equation. To convert from nominal interest rates to real interest rates, we use the following formula:

$$real\ interest\ rate \approx nominal\ interest\ rate - inflation\ rate.$$

If you want to know more about the Fisher equation, you can look in the toolkit.

For example, if a loan has a 12 percent interest rate and the inflation rate is 8 percent, then the real return on that loan is 4 percent. Since the nominal interest rate and the inflation rate are easily observed by most of us, we can use the Fisher equation to calculate the real rate of interest. We use the Fisher equation whenever we see a nominal interest rate and wish to convert it to a real interest rate. Just as it is the real exchange rate that matters for people trading goods and assets between countries, so it is the real interest rate that ultimately matters to borrowers and lenders in the economy.

In macroeconomics, we often look at the credit market for the entire economy, where savings and investment are matched in the economy as a whole. The price in this market is the real interest rate. The response of savings and investment to the real interest rate is shown in Figure 24.5.3. Once we know the equilibrium real interest rate, we calculate the implied nominal interest rate using the Fisher equation.

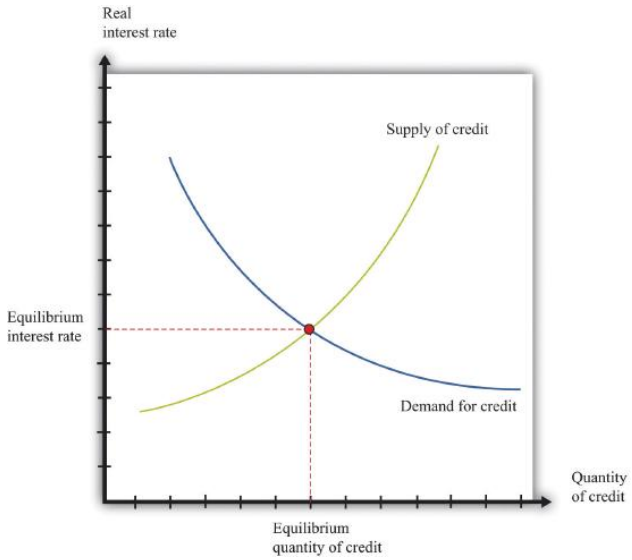


Figure 24.5.3: The Credit Market. Adjustment of the real interest rate ensures that the flows in and out of the financial sector balance.

The (net) supply of loans in the domestic credit market comes from three different sources:

1. The **private savings** of households and firms
2. The savings or borrowing of governments
3. The savings or borrowing of foreigners

Households will generally respond to an increase in the real interest rate by reducing current consumption relative to future consumption. Households that are saving will save more; households that are borrowing will borrow less. Higher interest rates also encourage foreigners to send funds to the domestic economy. Government saving or borrowing is little affected by interest rates.

National savings are defined as private savings plus government savings (or, equivalently, private saving minus the **government deficit**). The total supply of savings is therefore equal to national savings plus the savings of foreigners (that is, borrowing from other countries). The matching of savings and investment in the aggregate economy is described as follows:

$$investment = national\ savings + borrowing\ from\ other\ countries$$

or

$$investment = national\ savings - lending\ to\ other\ countries.$$

This is the same thing as saying that the flows in and out of the financial sector in the circular flow must balance.

The demand for loans comes from firms who borrow to finance investment. As the real interest rate increases, investment spending decreases. For firms, a high interest rate represents a high cost of funding investment expenditures. This is evident if the firm borrows to purchase capital. It is also true if the firm uses internal funds (retained earnings) to finance investment since the firm could always put those funds in an interest-bearing asset instead.

Toolkit: Section 31.27

The toolkit provides more detail on the flows in and out of the financial sector.

Arbitrage with Assets and Currencies: Uncovered Interest Parity

If you are like most people, you do not own assets in another country. You may own multiple assets—a savings account that pays you some interest every month, perhaps a certificate of deposit, or shares of some company—but the chances are that all your financial assets are denominated in a single currency. In fact, there is no reason why you should not own assets denominated in other currencies, such as euros, or pesos, or British pounds. You might consider opening a bank account in another country. Or you might even consider other financial investments in another country, such as purchasing a share in an international mutual fund, buying shares of a foreign company, or buying the debt of a foreign government.

Most of us do not know exactly how to go about making such investments. In fact, they are easy to carry out if you make use of the services of professional financial advisers. In any case, we are not really interested in the mechanics of foreign investment here. We want to answer a more fundamental question: how do you know if buying foreign assets would be a good idea? Consider the choice between two investment strategies.

1. Investing in the United States
    - Deposit $100 in a US bank.
    - Wait for a year.
2. Investing in Europe
    - Take $100 and use it to buy euros.
    - Deposit the euros in a European bank.
    - Wait for a year.
    - Withdraw the deposit and interest and use it to buy dollars.

To decide which is the better strategy, you need to determine how much you will earn in each case.

It is straightforward to determine how much you will get with the first option: you will get your $100 plus the interest payments. For example, if the interest rate at the US bank is 10 percent, then after a year you will earn $10 interest for a total of $110.

What about the second strategy? How many dollars will you have if you deposit money in the European bank? This is a bit more complicated. First, you buy euros with your $100. Second, you deposit these euros in a European bank and earn interest. Third, at the end of the year, you withdraw your euros from the bank and sell them for dollars. For example, suppose that the current dollar price of euros is $1.25 and the interest rate paid on deposits in Europe is 5 percent. Suppose you also expect that the price of a dollar in euros will be EUR 0.70 in a year's time. With the second investment strategy,

- You take your $100 and buy EUR 80.

- You put these EUR 80 in the European bank for a year, giving you EUR 84 at the end of the year.
- You take these EUR 84 and use them to purchase $120.

The second strategy therefore earns you more than the first strategy. It would be better to invest in Europe compared to the United States. Moreover, a slight variation on this strategy seems like it is a money machine. Consider the following.

- Borrow $100 from a US bank for one year.
- Take the $100 and use it to buy euros.
- Deposit the euros in a European bank.
- Wait for a year.
- Withdraw the deposit and interest and use it to buy dollars.
- Repay the dollar loan plus interest.

Using the same interest rates and exchange rates as previously, this transaction works as follows: you borrow $100, obtain $120 at the end of the year, pay back $110 to the bank, and end up with $10 profit.

To evaluate this arbitrage possibility, you need to know (1) the current dollar price of euros, (2) the annual return on deposits in Europe, and (3) the price of a dollar in euros a year from now. Look carefully at the language we used. You need to *know* "the euro price of dollars a year from now." But when we went through the example, we said "you *expect* that the price of dollar in euros will be EUR 0.70 in a year's time." As with the term structure of interest rates, there is some risk involved here. You cannot know the future exchange rate with certainty. This strategy entails a gamble about the future exchange rate. Still, if everyone has the same guess about the future exchange rate as you do, then such a situation could not last. Everyone would pursue the same strategy: borrow in the United States, buy euros, invest in Europe, and convert back in a year's time. What would happen?

- The demand for credit would increase interest rates in the United States.
- The demand for euros would increase the dollar price of euros.
- The extra supply of savings in Europe would drive down the interest rate in Europe.
- Investors might anticipate the extra demand for dollars in a year's time and expect the euro price of dollars to increase.

These forces would all tend to eliminate the profit opportunity.

So when do we expect this arbitrage opportunity to disappear? It disappears when investors expect to make the same profit whether they invest in Europe or the United States. The condition for this is as follows:

The left side is the return on investing in the United States. The terms on the right give the return on investing in Europe. When this condition holds, the returns on deposits in US and European banks are the same. This condition is called **uncovered interest parity**.

Because we do not know the price of euros next year for sure, this equation does not hold exactly when we look at actual data from the past. That is, the actual exchange rates combined with the actual returns on deposits do not quite satisfy this equation. This does not contradict the theory. Hindsight is perfect. The important point is that if people hold similar beliefs, then uncovered interest parity will hold ahead of time.

Using Uncovered Interest Parity to Understand the Exchange Rate

We can rearrange the uncovered interest parity condition as follows:

Written this way, the equation tells us that beliefs matter. Suppose everyone in the market believes that the dollar will depreciate relative to the euro in the future: that is, everyone expects a decrease in the euro price of the dollar. This makes investment in euro-denominated assets a better deal since we will get a lot of dollars per euro in the future. Investors will respond by selling dollars now to buy euros. This increase in the supply of dollars will cause the current euro price of dollars to decrease.

Thus we see that if everyone expects the euro price of dollars to decrease in the future, then the euro price of dollars will decrease today. When we talk about the market for currencies, demand and supply today depend on what households and firms think about the future exchange rate.

We can also rearrange the equation to see what it tells us about exchange rate beliefs:

If the interest rate in Europe is greater than the interest rate in the United States, then the condition tells us that investors must be expecting the dollar to appreciate.

Key Takeaways

- The nominal interest rate is the return on an asset in terms of money. The real interest rate is the return on an asset measured in terms of goods.
- The yield curve describes the relationship between the (annual) return on an asset and its maturity. Normally, the yield curve is upward sloping: assets with a longer maturity have a higher annual return.
- The Fisher equation links the real interest rate to the nominal interest rate. The real interest rate is approximately equal to the difference between the nominal interest rate and the inflation rate.

Exercises

1. If the nominal interest rate is 5 percent and the inflation rate is 3 percent, what is the real interest rate?
2. Can the real interest rate ever be negative?
3. What are the risks involved in investing in a foreign bank?

This page titled 24.5: Using Money to Buy Assets- Interest Rates is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

24.6: End-of-Chapter Material 24.6: End-of-Chapter MaterialEdit Update titlecancel

In Conclusion

We began this chapter with a deceptively simple question about money: why do people want it? To answer that question, we first looked at what money is. We discovered that money is an asset that has certain defining characteristics, such as portability, divisibility, and durability. Most importantly of all, though, we said that money must have acceptability. What turns an asset into a money, ultimately, is the simple fact that enough people are willing to treat it as such.

If we look through history, we find that many different things have served as money in different places and at different times. As well as the familiar notes and coins, these include seashells, stones, cigarettes, cans of food, gold, and silver. These could successfully function as money because they were acceptable as money in their particular context.

We then imagined that you were lucky enough to find a $100 bill on the sidewalk and explored the various things that you could do with this money, including buying goods and services, buying other currencies, and buying assets. As we did so, we explored a number of different arguments, all based on arbitrage, that help us to understand the relationships between interest rates, exchange rates, asset prices, and inflation rates. We argued that arbitrageurs will step in when there are easy profit opportunities. Arbitrage does not say that riskless profit opportunities cannot exist. It says that they will not persist. If a riskless profit opportunity were to exist, then people would very quickly take advantage of it and, by so doing, eliminate it.

Expressed more metaphorically, economists often say that there are no $100 bills lying on the ground waiting to be picked up. It is not that it is impossible for someone to drop a $100 bill, but if one person has dropped a large bill, someone else will almost certainly pick it up very quickly. There is an immediate and powerful lesson of arbitrage, one that you should bear in mind throughout life. If someone tells you of a surefire way to make easy money, beware!

Key Links

- History of currencies: http://www.frbsf.org/currency/index.html
- Value of a dollar: www.minneapolisfed.org/community_education/teacher/calc
- Euro overview: http://ec.europa.eu/economy_finance/euro/index_en.htm
- The *Economist*: Big Mac Index: http://www.economist.com/markets/bigmac
- Exchange rates: www.oanda.com
- US Department of the Treasury Daily Treasury Yield Curve: http://www.treasury.gov/resource-center/data-chart-center/Pages/index.aspx

exercises

1. Suppose you go to a local café to order a drink. Instead of paying with the currency used in your home country, imagine you try to pay with the currency of another country. What do you think the response would be at the café? Why? What could you do to convince them to accept foreign currency at a local café? Imagine that you are at the border of two countries, say in a café near the US border with Canada. Do you think you could use Canadian currency in a US café near the border?
2. When you are traveling in a foreign country and want to use your debit card, what type of fees do you pay to withdraw money in foreign currency? Usually fees take two forms: a fixed fee, say $5, for any size transaction or a fee that is proportional to the amount you withdraw. If you want to make a large withdrawal, which type of fee do you prefer? If the fee is fixed, will this create an incentive to make more or fewer withdrawals? What does the fixed fee do to the size of the withdrawal you make?
3. Suppose the dollar price of euros is $10 and the euro price of dollars is EUR 1. Explain how you could make a profit in this market. What would you buy and what would you sell? Can this be an equilibrium in the foreign exchange market? Show that there are no arbitrage profits if the dollar price of the euro is $1.25 and the euro price of the dollar is EUR 0.80.
4. (Advanced) Using the relationship

$$price\ of\ euro\ in\ dollars \times price\ of\ dollar\ in\ euros = 1,$$

how would you draw the supply and demand curves and depict equilibrium in the market for dollars and the market for euros?

5. Look at the left-hand table in Figure 24.6.3. How are the numbers on the bottom left connected to the numbers on the top right? The diagonal has been left blank. What number could go on the diagonal?
6. Look at the left-hand table in Figure 24.6.3. Suppose you start with GBP 100. Convert those pounds into euros and then convert the euros into dollars. How many dollars would you get? How many pounds do you get if you then convert your dollars into pounds?
7. Perform the same exercise as in Question 6 but use the table on the right-hand side of Figure 24.6.3. How many pounds do you end up with?
8. If the nominal interest rate is 5 percent in France and 3 percent in Europe, according to uncovered interest parity, what do investors think is going to happen to the euro-dollar exchange rate?
9. If the real interest rate is 2 percent in China and 6 percent in India, and investors are not expecting any change in the rupee-yuan exchange rate, then what can you conclude about inflation rates in China and India?
10. Explain how inflation reduces the real value of a currency both domestically and in other countries.

Economics Detective

1. Think of a "basket of goods" you buy often. It should include at least four items (for example, an espresso, a CD, a hamburger, and a copy of *Newsweek*). E-mail a friend in another country to find the prices for that same basket of goods. Check the exchange rate between your country's currency and that in your friend's country. Contrasting the prices in the two countries, look for violations of the law of one price. Is there some way you could make some profit?
2. Check rental car rates across two countries. (This is easy to do online at large car rental companies.) Make sure you choose the same car and insurance options. How might you explain the differences in these rates? Are there arbitrage profits for you to make?
3. Find an issue of *the Economist* from the period in the 1990s when Argentina was pegging the peso to the US dollar through a currency board, and look up the Big Mac index. What was the exchange rate then? What was the price of a Big Mac in Argentina during that period? Compare the peso prices of Big Macs and dollars between the two time periods.
4. Which countries use the kwacha, the dram, the ringgit, the leke, the baht, and the rial?
5. Suppose you want to convert some US dollars to euros, deposit them in a bank in Italy for one year, and then convert your euros to dollars. Search the Internet to determine how you could arrange now to buy dollars with euros in one year's time. What price would you have to pay for dollars?
6. Go to www.oanda.com and look at the latest exchange rate data. Find two currencies that have recently appreciated relative to the dollar and two currencies that have recently depreciated relative to the dollar.
7. Find a currency that has appreciated relative to the dollar since March 2007. Can you discover any explanations about why there was this change in the exchange rate?
8. Find data on the dollar price of the euro starting from the inception of the euro. Find periods when the dollar was appreciating relative to the euro. Find periods when the dollar was depreciating relative to the euro.
9. Who holds US government debt? What type of foreign exchange supported this?
10. Use www.minneapolisfed.org/community%5Feducation/teacher/calc to calculate the value of a dollar at different points in time. What would a dollar in 1955 buy today?
11. When you deposit money in the United States, you receive deposit insurance. If you deposit money in a bank in Italy or Japan or Mexico, will you receive deposit insurance? How does the existence of this insurance influence your decision about making deposits in foreign banks?
12. Call your bank to ask a hypothetical question: What will you have to do to deposit a large euro check in your dollar account? What will the bank charge you for this transaction? Are these costs proportional to the size of the euro check or is the cost a fixed number?

Spreadsheet Exercises

1. Suppose there are three currencies: dollars, pesos, and yuan. Write a spreadsheet program to find the dollar price of yuan given the dollar price of a peso and the peso price of the yuan such that there are no arbitrage profits to be made.
2. Suppose there are two countries: the United States and Mexico. Write a spreadsheet program to determine the interest rate on deposits in Mexico given the interest rate on deposits in the United States, the current exchange rate, and the expected future exchange rate so that there are no arbitrage profits to be made. All else being the same, how does a change in the US deposit rate affect the current exchange rate?

This page titled 24.6: End-of-Chapter Material is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

25: Understanding the Fed

25.1: Money and Power
25.2: Central Banks
25.3: The Monetary Transmission Mechanism
25.4: Monetary Policy, Prices, and Inflation
25.5: Monetary Policy in the Open Economy
25.6: The Tools of the Fed
25.7: The Fed in Action
25.8: End-of-Chapter Material

This page titled 25: Understanding the Fed is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

25.1: Money and Power 25.1: Money and PowerEdit Update titlecancel

In August 2011, these 10 individuals were among the most powerful people in the world.

- Ben S. Bernanke
- William Dudley
- Elizabeth Duke
- Charles Evans
- Richard Fisher
- Narayana Kocherlakota
- Charles Plosser
- Sarah Raskin
- Daniel Tarullo
- Janet L. Yellen

You may not have heard any of these names before. It is certainly unlikely that you have heard of more than one or two of these individuals. Yet they decide how easy or difficult it will be for you to get a job when you graduate. They decide how expensive it is for you to buy a car. They decide how many pesos you get for a dollar if you travel from the United States to Mexico. They decide if the Dow Jones Industrial Average is going to increase or decrease. They decide whether the stock markets in Tokyo, London, Hong Kong, and Frankfurt are going to increase or decrease. They decide the cost of your vacation abroad and the cost of the clothes that you buy at home.

So who are they?

They are the members of a group called the Federal Open Market Committee (FOMC). They are responsible for setting monetary policy in the United States. Of course, they do not literally decide all the things we just mentioned, but their decisions do have a major influence on everything we listed. This chapter is about what these people do and why their choices matter so much for our day-to-day life. We begin with an example of this group at work.

FOMC Policy Announcement: February 2, 2005

For immediate release

The Federal Open Market Committee decided today to raise its target for the federal funds rate by 25 basis points to 2-1/2 percent.

The Committee believes that, even after this action, the stance of monetary policy remains accommodative and, coupled with robust underlying growth in productivity, is providing ongoing support to economic activity. Output appears to be growing at a moderate pace despite the rise in energy prices, and labor market conditions continue to improve gradually. Inflation and longer-term inflation expectations remain well contained.

The Committee perceives the upside and downside risks to the attainment of both sustainable growth and price stability for the next few quarters to be roughly equal. With underlying inflation expected to be relatively low, the Committee believes that policy accommodation can be removed at a pace that is likely to be measured. Nonetheless, the Committee will respond to changes in economic prospects as needed to fulfill its obligation to maintain price stability.

Voting for the FOMC monetary policy action were: Alan Greenspan, Chairman; Timothy F. Geithner, Vice Chairman; Ben S. Bernanke; Susan S. Bies; Roger W. Ferguson, Jr.; Edward M. Gramlich; Jack Guynn; Donald L. Kohn; Michael H. Moskow; Mark W. Olson; Anthony M. Santomero; and Gary H. Stern.

In a related action, the Board of Governors unanimously approved a 25-basis-point increase in the discount rate to 3-1/2 percent. In taking this action, the Board approved the requests submitted by the Boards of Directors of the Federal Reserve Banks of Boston, New York, Philadelphia, Cleveland, Richmond, Atlanta, Chicago, St. Louis, Minneapolis, Kansas City, Dallas, and San Francisco.Federal Open Market Committee, “Press Release,” *Federal Reserve*, February 2, 2005, accessed July 20, 2011, http://www.federalreserve.gov/boarddocs/press/monetary/2005/20050202/default.htm.

This FOMC statement is from February 2005. We have deliberately chosen a statement from a few years ago because we want to begin with monetary policy prior to the economic crisis that began in 2008. This policy statement contains all the essential elements of monetary policy in normal times.

The 12 people listed in the second-to-last paragraph of this announcement were the FOMC members in February 2005. (These names are different from those we named at the start of the chapter because the composition of the FOMC changes over time.) The president of the United States was not one of them. And none of them are members of Congress. You did not vote for any of them. None of the three main branches of the US government (executive, legislative, or judicial) is involved in the setting of US monetary policy. The FOMC is part of a government body called the US Federal Reserve Bank, commonly known as the Fed. The Fed is *independent*: decisions made by the Fed do not have to be approved by other branches of the government.

In this statement we find the following phrases:

- "The Federal Open Market Committee decided today to raise its target for the federal funds rate by 25 basis points to 2-1/2 percent."
- "The Committee perceives the upside and downside risks to the attainment of both sustainable growth and price stability for the next few quarters to be roughly equal."
- "In a related action, the Board of Governors unanimously approved a 25-basis-point increase in the discount rate to 3-1/2 percent."

The first phrase indicates an *action* undertaken by the Fed: it changed its "target" for something called the "federal funds rate." This is a particular interest rate related to the rate banks pay each other for loans. Although you will never borrow to buy a car or a house at this rate, the interest rates you confront are heavily influenced by the federal funds rate. For example, over the past few years, the federal funds rate has decreased from 5.25 percent in 2006 to a value of 0.25 percent at the time of writing (mid-2011). Over this same period of time, rates on other types of loans, including mortgages and car loans, decreased as well. For example, typical car loan rates were about 7–8 percent in 2006 and about 3–4 percent in mid-2011. In this way, the actions of the Fed affect all of us.

The second phrase contains the FOMC's assessment of the state of the economy, expressed in terms of two *goals*: economic growth and the stability of prices. The Fed is charged with the joint responsibility of stabilizing prices and ensuring the full employment of economic resources. The final statement details another action with respect to a different interest rate, called the discount rate.

The FOMC issues statements such as this on a regular basis. Our goal in this chapter is to equip you with the knowledge to understand these statements, which will in turn help you make sense of the discussions of the Fed's actions on television or in the newspapers. We want to answer the following questions:

What does the Federal Reserve do? And why are its actions so important?

Road Map

The FOMC statement reveals that, to understand the Fed, we need to know both the goals and the tools of the Fed. From the statement, we learn that the goals of the Fed are *sustainable growth* and *stable prices*. The Fed cannot do much to affect the long-run growth rate of the economy, but it can and does try to keep the economy close to potential output. At the same time, it tries to ensure that the overall price level does not change very much—in other words, it tries to keep inflation low. The Fed pursues these goals by means of several tools that it has at its disposal. The FOMC statement informs us that these tools include two different interest rates.

We begin with a little bit of background information. We briefly explain what the Federal Reserve does, and we note that other monetary authorities are similar, although not identical, in terms of goals and behavior. Because we have seen that the Fed's actions frequently revolve around interest rates, we make sure that we know exactly what an interest rate is.

We then get to the meat of the chapter, which discusses the workings of monetary policy. We explain how the Fed uses its tools to affect the things it ultimately cares about. Broadly speaking, we can summarize the cyclic behavior of the Fed as follows:

- The Fed observes current economic conditions.
- The Fed decides on policy actions.
- These policy actions affect real GDP (gross domestic product) and inflation.
- The Fed observes the new economic conditions.

There is a long chain of connections between the Fed's tools and the ultimate state of the economy. To make sense of what the Fed does, we follow these connections, step by step. As we do so, we create a framework for understanding the effects of monetary policy, called the monetary transmission mechanism. We must also look at the connection in the other direction: how does the state of the economy influence the Fed's decisions? Figure 25.1.1, which we use as a template for the chapter,

summarizes the interaction between the monetary transmission mechanism and the behavior of the Fed. We conclude the chapter by looking at the tools of the Fed in more detail and by discussing some historical episodes through the lens of monetary policy.

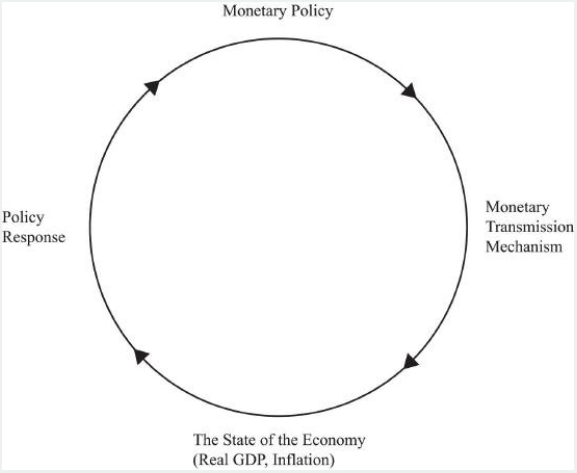


Figure 25.1.1: The Links between Monetary Policy and the State of the Economy. The Federal Reserve looks at current economic conditions and decides on a policy response. This policy affects the state of the economy. The Fed then observes the new economic conditions and decides on a new policy response, and so forth.

This page titled 25.1: Money and Power is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

25.2: Central Banks 25.2: Central BanksEdit Update titlecancel

learning ojbectives

After you have read this section, you should be able to answer the following questions:

1. When and why was the Federal Reserve System created in the United States?
2. What are the connections between the Federal Reserve System and the executive and legislative branches of the US government?
3. How does our study of monetary policy apply to other central banks around the world?

We start our discussion with institutions.

The Federal Reserve

The Federal Reserve System was formally established in an act of Congress on December 23, 1913, called the Federal Reserve Act (http://www.federalreserve.gov/aboutthefed/fract.htm). The stated purpose of the act was as follows: "To provide for the establishment of Federal reserve banks, to furnish an elastic currency, to afford means of rediscounting commercial paper, to establish a more effective supervision of banking in the United States, and for other purposes."The Federal Reserve Act is found at "Federal Reserve Act," Board of Governors of the Federal Reserve System, accessed September 20, 2011, http://www.federalreserve.gov/aboutthefed/fract.htm, and the structure of the Federal Reserve System is presented at "The Structure of the Federal Reserve System," The Federal Reserve Board, accessed September 20, 2011, http://www.federalreserve.gov/pubs/frseries/frseri.htm. The Federal Reserve System is built around a 7-member Board of Governors together with 12 regional banks. The members of the board are appointed by the president and approved by Congress to serve for 14 years. The FOMC, which is instrumental in the conduct of monetary policy, has 12 members.

Although the president and Congress play a role in the appointment of members of the Fed, their direct control stops there. The Fed is an independent body. The executive and congressional branches of the government have no formal input into the determination of monetary policy. Congressional control is limited to the fact that the chair of the Fed is required to report to Congress periodically and to Congress eventually having the power to change the laws governing the Fed's conduct.

The goals of the Fed are specified in the section of the Federal Reserve Act titled "Monetary Policy Objectives": "The Board of Governors of the Federal Reserve System and the Federal Open Market Committee shall maintain long run growth of the monetary and credit aggregates commensurate with the economy's long run potential to increase production, so as to promote effectively the goals of maximum employment, stable prices, and moderate long-term interest rates."“Federal Reserve Act: Monetary Policy Objectives," Federal Reserve, December 27, 2000, accessed August 6, 2011, http://www.federalreserve.gov/aboutthefed/section2a.htm. These objectives provide guidance to the Fed: it is required to pay attention to the level of economic activity ("maximum employment") and to the level of inflation ("stable prices"). Exactly how the Fed promotes these goals—and chooses among them if necessary—is not specified. In some cases, the different aims of the Fed may conflict. For example, promoting employment may not be consistent with low inflation. The February 2, 2005, statement explicitly notes the balance between these goals.

The Fed has three main ways of affecting what goes on in the economy. The first was alluded to, although not mentioned by name, in the February 2, 2005, policy announcement. It is called open-market operations and represents the main way that the Fed influences interest rates. A second tool—the discount rate—was mentioned explicitly in the policy announcement. The third tool—reserve requirements—was not mentioned on February 2, 2005, but is nonetheless an important weapon in the Fed's arsenal. Later on in this chapter, we examine the tools of the Fed in detail. For the moment, it is enough to know that the Fed affects the economy through changes in interest rates.

Central Banks in Other Countries

Our discussion in this chapter applies to not only the United States but also other countries. Wherever there is a currency, there is a monetary authority—a central bank—charged with the control of that currency. For example, in Europe, the *European Central Bank* (ECB; http://www.ecb.int/home/html/index.en.html) dictates monetary policy for all those countries that use the euro. In Australia, the *Reserve Bank of Australia* (RBA; http://www.rba.gov.au) manages monetary policy.

Different central banks do not all function in exactly the same way. To illustrate, here are policy announcements from the *Bank of England* (BOE; www.bankofengland.co.uk/publications/news/2006/078.htm), the *Central Bank of Egypt* (CBE; www.cbe.org.eg/public/PRESS_Release_For_Monetary_Policy/2011/MPC_PressRelease_09_06_2011_E.pdf), and the RBA (http://www.rba.gov.au/media-releases/2011/mr-11-09.html). The details of the announcements are not critical. However, all have a "Monetary Policy Committee" rather than an FOMC. The different banks target slightly different interest rates: the BOE targets the "Bank rate paid on commercial bank reserves"; the CBE refers to overnight deposit and lending rates, the "7-day repo," and the discount rate; and the RBA refers to the "cash rate." You do not need to worry about exactly what these different rates are. All three banks are looking at the overall state of the economy, in terms of both output and inflation, and are setting interest rates to pursue broadly similar goals.

News Release: Bank of England Raises Bank Rate by 0.25 Percentage Points to 4.75 Percent, 3 August 2006 "News Release," Bank of England, August 3, 2006, accessed July 20, 2011, www.bankofengland.co.uk/publications/news/2006/078.htm.

The Bank of England's Monetary Policy Committee today voted to raise the official Bank rate paid on commercial bank reserves by 0.25 percentage points to 4.75 percent.

The pace of economic activity has quickened in the past few months…As a result, over the past few quarters GDP [gross domestic product] growth has been at, or a little above, its long-run average and business surveys point to continued firm growth.…

CPI [Consumer Price Index] inflation picked up to 2.5 percent in June, and is expected to remain above the 2.0 percent target for some while. Higher energy prices have led to greater inflationary pressures, notwithstanding muted earnings growth and a squeeze on profit margins.…

Against the background of firm growth, limited spare capacity, rapid growth of broad money and credit, and with inflation likely to remain above the target for some while, the Committee judged that an increase of 0.25 percentage points in the official Bank rate to 4.75 percent was necessary to bring CPI inflation back to the target in the medium term.

Press Release, June 9, 2011: The Central Bank of Egypt Decided Not to Raise Its Policy Rates "Press Release," Central Bank of Egypt, June 9, 2011, accessed July 20, 2011, www.cbe.org.eg/public/PRESS_Release_For_Monetary_Policy/2011/MPC_PressRelease_09_06_2011_E.pdf.

In its meeting held on June 9, 2011, the Monetary Policy Committee (MPC) decided to keep the overnight deposit and lending rate unchanged at 8.25 and 9.75 percent, respectively, and the 7-day repo at 9.25 percent. The discount rate was also kept unchanged at 8.5 percent.

Headline CPI increased by 0.20 percent in May [month to month] following the 1.21 percent in April, bringing the annual rate down to 11.79 percent from 12.07 percent registered in April. …

Meanwhile, real GDP contracted by 4.2 percent in 2010/2011 Q3 which marks the first negative year-on-year growth since the release of quarterly data in 2001/2002. …

Against the above background, the slowdown in economic growth should limit upside risks to the inflation outlook. Given the balance of risks on the inflation and GDP outlooks and the increased uncertainty at this juncture, the MPC judges that the current key CBE [Central Bank of Egypt] rates are appropriate.

Media Release Number 2011-09: Statement by Glenn Stevens, Governor: Monetary Policy Decision Glenn Stevens, "Media Release," Reserve Bank of Australia, June 7, 2011, accessed July 20, 2011, http://www.rba.gov.au/media-releases/2011/mr-11-09.html.

At its meeting today, the Board decided to leave the cash rate unchanged at 4.75 per cent.

The global economy is continuing its expansion, led by very strong growth in the Asian region, though the recent disaster in Japan is having a major impact on Japanese production, and significant effects on production of some manufactured products further afield. Commodity prices have generally softened a little of late, but they remain at very high levels, which is weighing on income and demand in major countries and also pushing up measures of consumer price inflation. …

Growth in employment has moderated over recent months and the unemployment rate has been little changed, near 5 per cent. Most leading indicators suggest that this slower pace of employment growth is likely to continue in the near term…

CPI inflation has risen over the past year, reflecting the effects of extreme weather and rises in utilities prices, with lower prices for traded goods providing some offset. The weather-affected prices should fall back later in the year, though substantial rises in

utilities prices are still occurring. The Bank expects that, as the temporary price shocks dissipate over the coming quarters, CPI inflation will be close to target over the next 12 months.

At today's meeting, the Board judged that the current mildly restrictive stance of monetary policy remained appropriate. In future meetings, the Board will continue to assess carefully the evolving outlook for growth and inflation.

In this chapter, we talk, for the most part, about the Federal Reserve. We focus on the United States principally because we do not want to get too bogged down in learning the different languages used by different central banks. From looking at the statements of the Fed, the BOE, the CBE, and the RBA, we see that the terminology of monetary policy varies greatly from country to country, the names of the key interest rates differ, and so forth. The underlying principles of monetary policy are largely the same in all countries, however.

Key Takeaways

1. The Federal Reserve System of the United States was created in 1913. A key motivation for the creation of a central bank was to manage the stock of currency and thus influence the state of the aggregate economy, particularly output and prices.
2. In the United States, the central bank is independent. Decisions about monetary policy are made within the Federal Reserve System. Members of the Board of Governors of the Federal Reserve System are nominated by the president and approved by the Senate.
3. There are central banks around the world, conducting monetary policy with similar tools and with the same basic model of the aggregate economy in mind.

Exercises

1. What is the input of the US president in determining monetary policy?
2. By learning about how the Federal Reserve System in the United States conducts monetary policy, what can we learn about other countries?

This page titled 25.2: Central Banks is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

25.3: The Monetary Transmission Mechanism 25.3: The Monetary Transmission MechanismEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What is the link between the actions of the Fed and the state of the economy?
2. What interest rate does the Fed target?
3. What components of aggregate spending depend on the interest rate?

The actions of monetary authorities, such as the Fed and other central banks around the world, influence interest rates and thus the levels of employment, output, and prices. The links between a central bank's actions and overall economic performance are far from straightforward, however. The process is summarized by the **monetary transmission mechanism** (shown in Figure 25.3.1), which is the heart of this chapter. The monetary transmission mechanism is more than just some theory that economists have devised to try to make sense of monetary policy. It summarizes *how the Fed thinks about its own actions*.

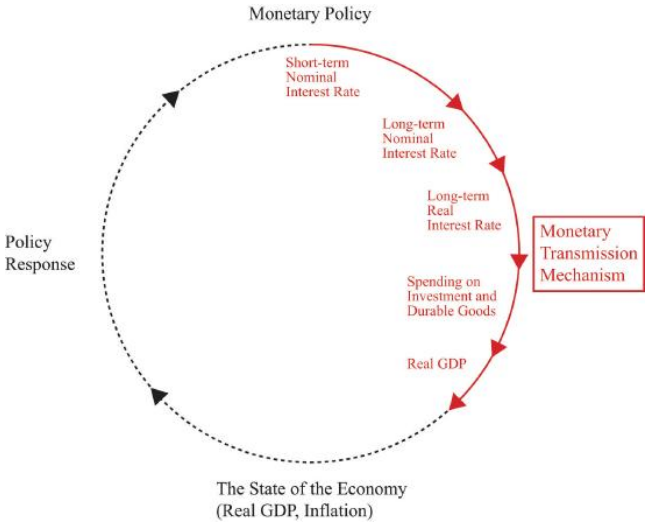


Figure 25.3.1: The Monetary Transmission Mechanism. The Fed targets a short-term nominal interest rate. Changes in this rate lead to changes in long-term real interest rates, which affect spending on investment and durable goods, ultimately leading to a change in real GDP.

The monetary transmission mechanism explains how the actions of the Federal Reserve Bank affect aggregate economic variables, and in particular real gross domestic product (real GDP). More specifically, it shows how changes in the Federal Reserve's target interest rate affect different interest rates in the economy and thus influence spending in the economy. Through open-market operations, the Fed targets a *short-term* **nominal interest rate**. Changes in that interest rate in turn affect *long-term* nominal interest rates. Changes in long-term nominal rates lead to changes in long-term **real interest rates**. Changes in long-term real interest rates affect **investment** and **durable goods** spending. Finally, changes in spending affect real GDP. We will examine every step of this process.

This chapter focuses on the effects of Fed actions, but essentially the same analysis applies to the study of monetary policy in other countries. The channels of influence are to a large degree independent of which country we study, although the magnitudes of the policy effects might differ across countries. Monetary policy differs across countries more through the targets set by different central banks than through the transmission mechanism.

How Well Can the Fed Meet Its Target?

On February 2, 2005, the Federal Open Market Committee (FOMC) decided to increase the target federal funds rate to 2.5 percent. The word *target* is critical here. If you listen to television news, you might get the impression that the Fed sets interest rates. It does not. It influences them, with greater or lesser success at different times.

Figure 25.3.2 shows the monthly target and actual federal funds rate between 1971 and 2008. From this picture, it is evident that the target and actual federal funds rates move together. We can conclude that the first stage of the monetary transmission mechanism is reliable. The Fed can influence the federal funds rate. So far so good—at least for this period of time. As we shall see later, when we consider more recent events, the Fed was much less successful in targeting the federal funds rate during the periods of financial distress in 2007 and 2008.

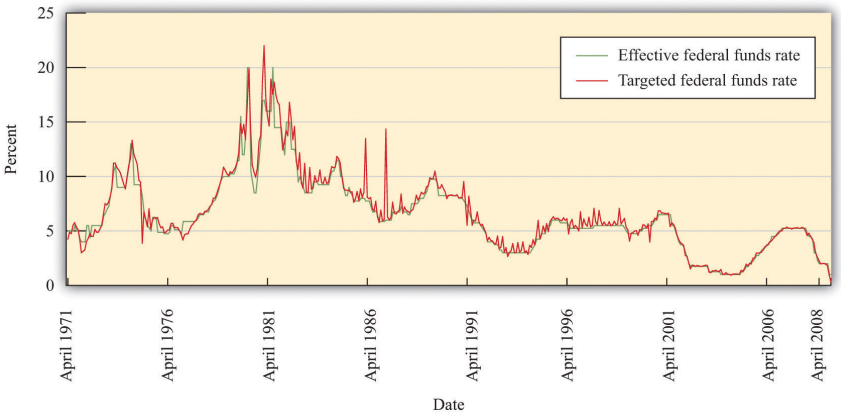


Figure 25.3.2: Target and Actual Federal Funds Rate, 1971–2005. The target and actual federal funds rates move closely together.

From Short-Term Interest Rates to Long-Term Interest Rates

The next question is, do movements in the federal funds rate lead to corresponding movements in long-term interest rates? By "long-term," we mean the rates on assets that have a maturity of at least 1 year and, in particular, assets that have a maturity of 5 years, 10 years, or even longer. **Arbitrage** among different assets means that annual interest rates on assets with different **maturities** are linked. As a result, the actions of the Fed to influence short-term rates also affect long-term rates.

Figure 25.3.3 shows the relationship between the federal funds rate and longer-term interest rates. Broadly speaking, these long rates move with the federal funds rate. But it is also clear that the longer the horizon on the debt, the less responsive is the interest rate to movements in the federal funds rate.

This is one of the difficulties faced by the Fed: it can target short-term rates very accurately, but its influence on long-term rates is much less precise. Since—as we shall see—many economic decisions depend on long-term rates, the Fed's ability to influence the economy is imperfect. Some writers have suggested that the Fed is an all-powerful organization that can move the economy around on a whim. There is no doubt that the Fed wields a great deal of power over the economy. Nevertheless, the Fed's influence is substantially limited by the fact that it cannot control long-term interest rates with anything like the same precision that it brings to bear on the federal funds rate.

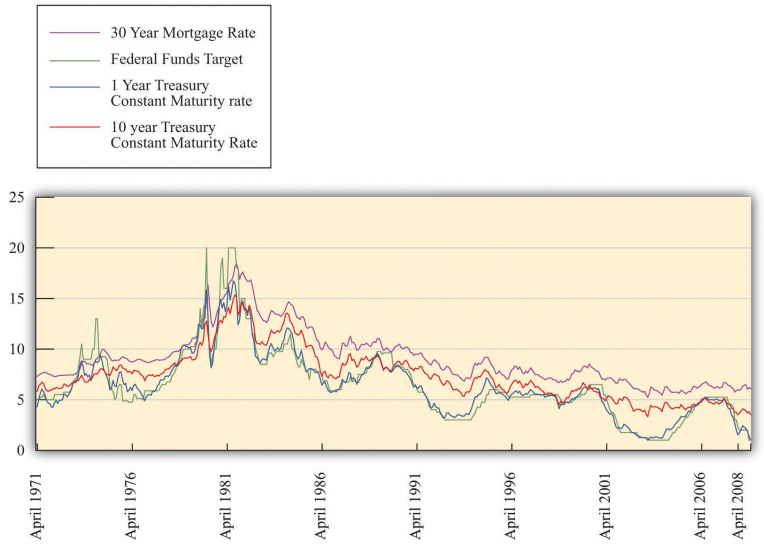


Figure 25.3.3: Short-Term and Long-Term Interest Rates. The Fed's ability to influence long interest rates is much more limited than its ability to affect short rates.

From Nominal Interest Rates to Real Interest Rates

So far in this section, we have been considering nominal interest rates, but we know that the decisions of firms and households are based on real interest rates. The link between nominal and real interest rates is given by the **Fisher equation**:

$$real\ interest\ rate \approx nominal\ interest\ rate - inflation\ rate.$$

To use this relationship, we simply subtract the inflation rate from the nominal interest rate. So if the nominal interest rate were 15 percent, as it was in the early 1980s, and the inflation rate were 12 percent, then the real interest rate would be 3 percent. But if the inflation rate were higher—say, 18 percent—then the real interest rate would be minus 3 percent.

Toolkit: Section 31.25

The toolkit reviews the derivation of the Fisher equation.

Figure 25.3.4 shows the nominal and real rates of return for a one-year Treasury bond. Because inflation is positive, the nominal interest rate exceeds the real rate. The figure shows that the nominal and real rates typically move closely together. In the early 1980s, for example, the real interest rate was negative. Presumably when households lent money in the early 1980s, they did not expect a negative return on their saving but instead expected that the nominal interest rate would exceed the inflation rate. From that perspective, the negative real interest rate is a consequence of higher than anticipated inflation.

The Fed's ability to influence longer-term nominal rates through its influence on the federal funds rate apparently extends to the real interest rate as well. The connection is not perfect, however. On some occasions, movements in nominal rates are decoupled from movements in real rates.

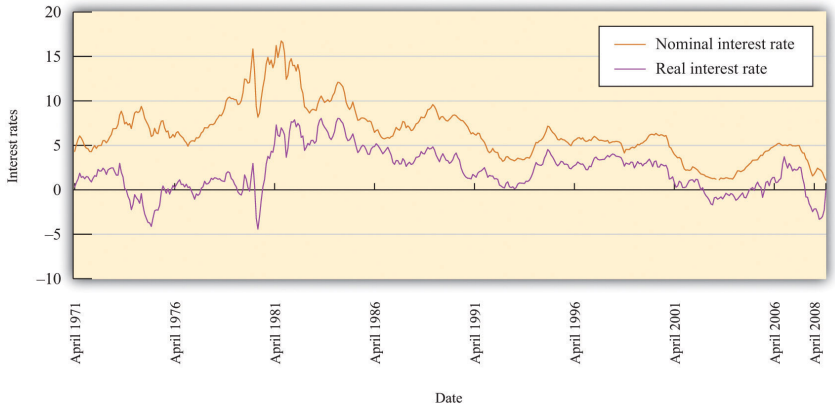


Figure 25.3.4: Real and Nominal Interest Rates. Changes in nominal interest rates generally lead to changes in real interest rates, but the link between the two is imperfect.

From Real Interest Rates to Spending on Durable Goods

Real rates of interest influence spending by both households and firms. The main categories of purchases that are affected by interest rates are as follows:

- Investment spending by firms
- Housing purchases by households
- Durable goods purchases by households

What do these have in common? In each case, the purchase yields a flow of benefits that extends over some significant period of time. If a firm builds a new factory or purchases a new piece of machinery, it typically expects to be able to use that plant and equipment for years or decades. When a household buys a new home, it expects either to live there for a long time or else to sell it to someone else who can live there. If you buy a durable good such as a new car or a refrigerator, you expect to obtain the benefits of that purchase for several years.

Figure 25.3.5 shows the relationship between the real interest rate and spending on durable goods. The higher the real interest rate is, the lower is the amount of spending on durable goods. Of course, the relationship need not be a straight line; we have just drawn

it this way for simplicity. As you might imagine, monetary policymakers are very interested in the exact form of this relationship. They want to know exactly how big a change in durable goods spending is likely to follow from a given change in interest rates.

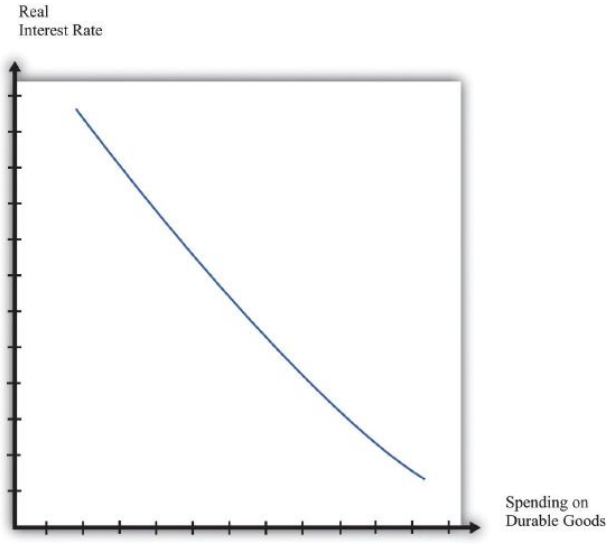


Figure 25.3.5: Real Interest Rates and Spending on Durable Goods. At higher interest rates, firms are less likely to borrow for investment projects, and households are less likely to borrow to purchase housing and durable goods such as new cars. Thus spending on durable goods is lower at higher interest rates and vice versa.

Discounted Present Value and Spending on Durable Goods

To understand in more detail why interest rates affect spending on durable goods, consider the purchase of a machine by a firm. Firms carry out such investment spending because they expect the machine to yield a flow of profits not only in the present but also for several years into the future. A machine is a capital good; it is used in the production of other goods and is *not* used up during the production process. The fact that the returns from the machine accrue over several years is what we mean by the term *durable*.

It is not correct to simply add profit flows in different years because a dollar today is usually worth more than a dollar next year. Why? If you take a dollar today and put it in a savings account at the bank, you will get your dollar plus interest back next year. If the interest rate is 10 percent, then $1 this year is worth $1.10 next year. Turning it around, $1 next year is worth only about 91 cents this year (because 1/1.1 = 0.91).

The technique for adding together flows of resources in different periods is called **discounted present value**. To work out whether a given investment is profitable, a firm must calculate the value, in today's terms, of the flows of profits that it expects to receive. It then compares this to the cost of the investment. If the discounted present value of the profits exceeds the cost, the firm will undertake the investment.

Toolkit: Section 31.5

You can review discounted present value in the toolkit.

Table 25.3.1 "Return on Investment" illustrates a simple investment decision. In year 1, you pay for a machine, and it yields some profit in that year. The next year, the machine yields further profit. Suppose you, as a manager of a firm, must decide whether or not to buy this machine. How do you make this decision? In the first year, you pay $970 for the machine and earn only $500 back, so you are down $470. In the second year, you will earn an additional $500—but you have to wait a year to get this money. Think of the profits in year 2 as being in real terms—that is, already corrected for inflation.

Table 25.3.1: Return on Investment

Year	Payment for Machine	Real Profit from Machine
1	970	500
2	0	500

To decide about the purchase of the machine, you need to know the interest rate. If the interest rate were zero, the calculation would be easy. You could just add together the profit flows in the 2 years, observe that $1,000 is more than the $970 that you have to pay for the machine, and conclude that the purchase is a good idea. Now suppose the real interest rate is 5 percent, which means that the real interest factor is 1.05. Then the discounted present value of the profit flows from the machine is given by the following equation:

In this case, the purchase is still a good idea. You will earn $976 in present value terms, exceeding the $970 that you have to pay, so you still come out ahead.

But what if that the real interest rate is 10 percent? In this case,

This is less than the amount that you had to pay for the machine. The investment no longer looks like a good idea. The higher the interest rate, the more we must *discount* future earnings, so the less likely it is that a current investment will be profitable.

In most real-life cases, the flow of profits extends for several years, so the discounted present value calculation is somewhat harder. (Still, even harder calculations can be done easily with a calculator or a spreadsheet.) Our example may be simple, but it illustrates our key point. As the real interest rate decreases, the discounted present value of profits from a machine increases. In the economy, there are at any given time many possible investment opportunities. Some have higher profit flows than others. At lower interest rates, more machines will be profitable to purchase: investment increases as the real interest rate decreases.

Households purchase homes and durable consumption goods, such as cars and household appliances. If the household borrows to make such purchases (through mortgages, car loans, or other personal loans), then exactly the same logic applies. Higher interest rates will tend to deter the household from these purchases, whereas lower interest rates will encourage purchases.Even if the household uses its own accumulated saving to buy the durable good, there is an *opportunity cost* of using these funds: it could have put the money in the bank instead. The higher the real interest rate, the better it looks to put money in the bank. Households usually have some choice about when exactly to purchase such goods. If interest rates are high this year, it probably makes sense to put off that purchase of a new washing machine until next year, when rates might be lower.

The effect of an increase in the real interest rate on spending on durable goods is captured in Figure 25.3.6.

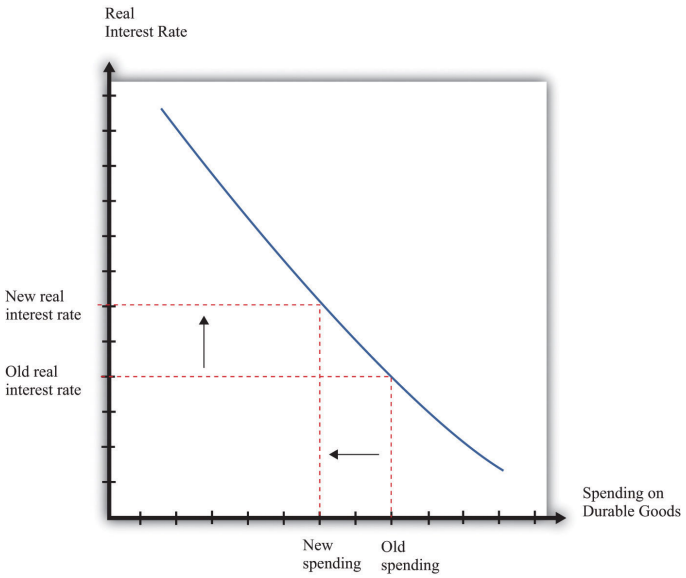


Figure 25.3.6: *The Relationship between the Real Interest Rate and Spending on Durable Goods*. When the real interest rate increases, spending on durable goods decreases.

From Spending on Durable Goods to Real GDP

Look again at Figure 25.3.1. We have so far explored the links from the Fed's decision on a target to spending on durable goods and net exports. Now we examine how changes in spending affect total output in the economy.

The **aggregate expenditure model** allows us to see how changes in aggregate spending translate into changes in GDP, at a given price level. The idea underlying the aggregate expenditure model is that, by the rules of national income accounting, real GDP must equal both production and spending. If spending increases, then it must be the case that production increases as well. The key diagram of the aggregate expenditure model is shown in Figure 25.3.7.

Variations in the real interest rate influence the level of aggregate spending through the level of autonomous spending (the intercept term). To see why, recall that total spending is the sum of consumption, investment, government purchases, and net exports. The intercept term of the expenditure relationship includes all the influences on spending *other than output*. Thus any changes in consumption, investment, or net exports that are *not* induced by changes in output show up as changes in the intercept term. In particular, if an increase in interest rates causes firms to cut back on their investment spending, then the planned spending line shifts downward.

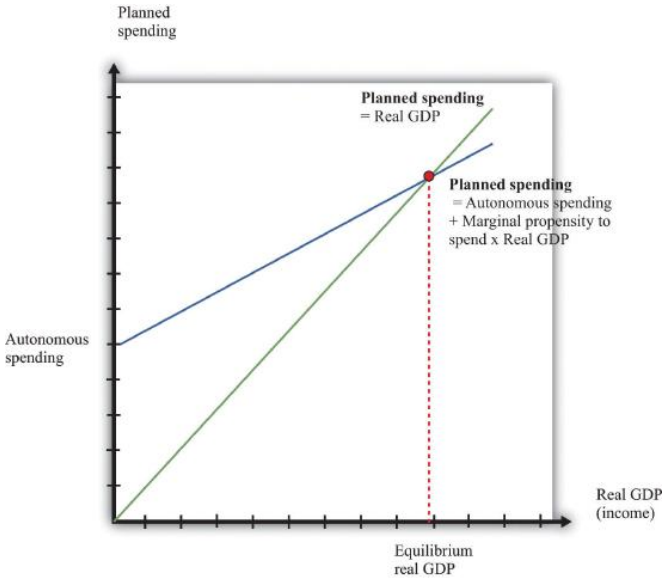


Figure 25.3.7: Aggregate Spending Depends Positively on Income. The economy is in equilibrium when spending equals real GDP.

We saw in Figure 25.3.6 that, as the real interest rate increases, the level of spending on durables decreases. This leads to a decrease in spending, given the level of income, and thus a decrease in the intercept of the spending line, as shown in Figure 25.3.8. The magnitude of the reduction in spending—that is, the shift downward in the spending line—will depend on the sensitivity of durable spending to real interest rates. *The more sensitive durable spending is to changes in the real interest rate, the larger the shift in the spending line will be when the real interest rate changes.*

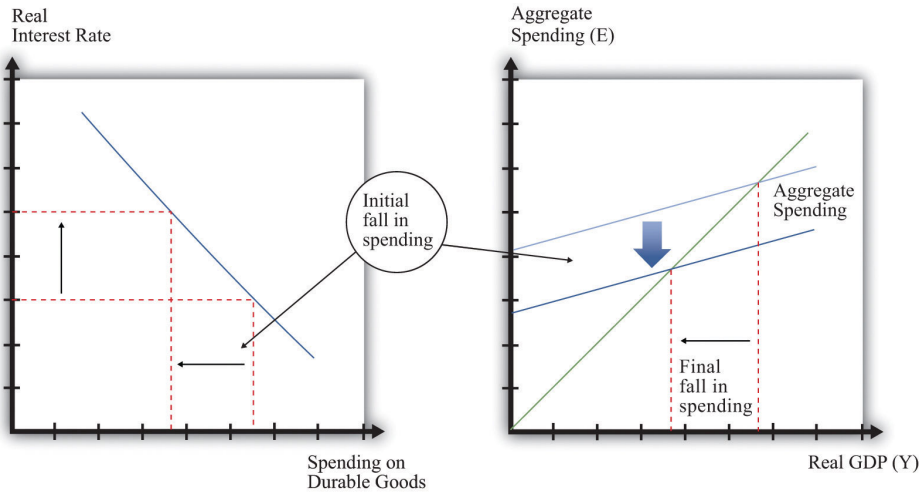


Figure 25.3.8: Increases in Real Interest Rates Reduce Real GDP. As a consequence of increases in real interest rates, aggregate spending decreases.

The initial reduction in spending induced by the increased real interest rate is then magnified by the **multiplier** process. The reduction in durable spending leads to a contraction in output. The resulting decrease in income leads households to spend less, leading to further contractions in output and income. In the end, the overall reduction in output exceeds the initial reduction in spending. This is visible in Figure 25.3.8 from the fact that the horizontal difference between the old and new equilibrium points is larger than the vertical shift in the spending line.

Toolkit: Section 31.30

You can review the aggregate expenditure model and the multiplier in the toolkit.

The Real Interest Rate–Real GDP Line

We can summarize much of the monetary transmission mechanism by means of a relationship between real interest rates and real GDP, as shown in Figure 25.3.9. After we work through all the connections from real interest rates to the various components of spending and real GDP, we find that there is a level of real GDP associated with each real interest rate. The higher the interest rate, the lower is real GDP.

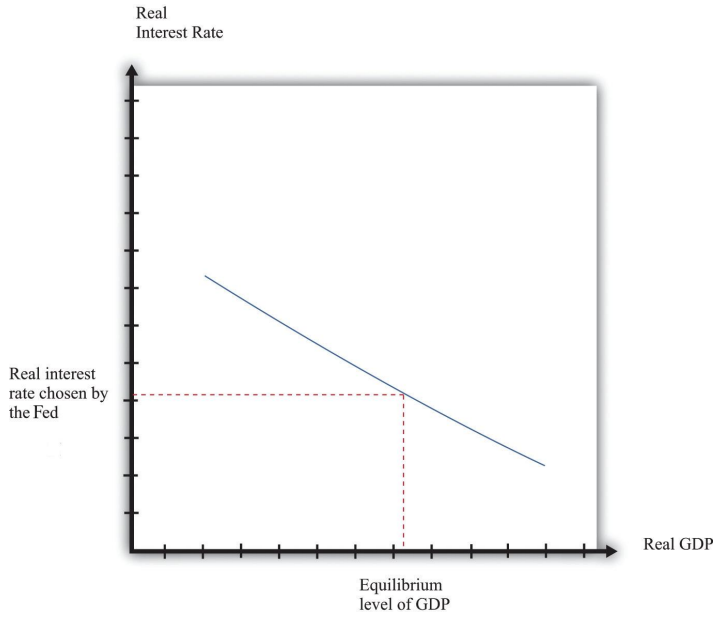


Figure 25.3.9: The Relationship between the Real Interest Rate and Real GDP. This picture summarizes several steps in the monetary transmission mechanism to show the relationship between real interest rates and real GDP.

As the monetary authority changes the real interest rate, the economy moves along this curve. So, for example, a reduction in the real interest rate leads to increased spending on durables, which, through the **multiplier** process, increases aggregate output. The shape of the curve tells us something about the Fed's ability to influence the economy. Suppose that (1) durable spending is very sensitive to the real interest rate and (2) the multiplier is large; then imagine that the Fed cuts interest rates. Firms and households both respond to this change. Firms decide to carry out more investment: they buy new machinery, open new plants, and so forth. Households, attracted by the low interest rates, borrow to buy new cars and new homes. As a result, durable spending increases substantially. Furthermore, this increase in spending leads to higher income and thus to further increases in spending by households. The end result is a large increase in real GDP. In this case, the curve is flat.

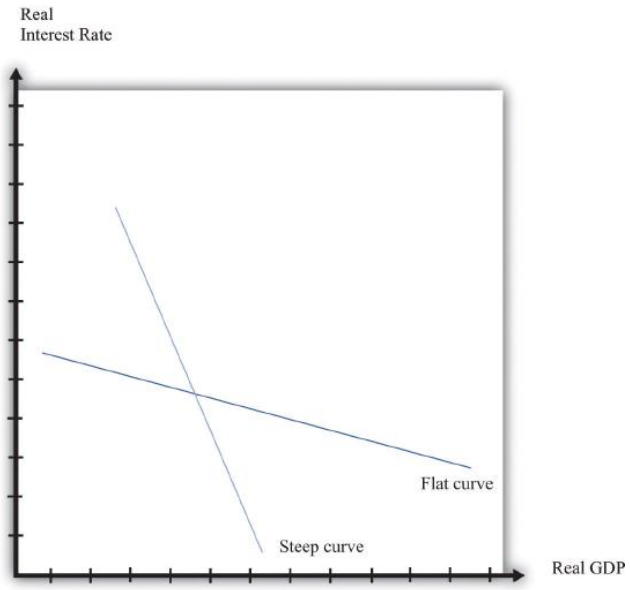


Figure 25.3.10: The Fed's Influence on the Economy Depends on the Real Interest Rate–Real GDP Relationship. When the curve is flat, the Fed is able to have a big influence on the economy. When the curve is steep, it is harder for the Fed to affect economic activity.

Figure 25.3.10 shows both this case and the case where it is harder for the Fed to influence the economy. If spending on durable goods is not very responsive to changes in the real interest rate and the multiplier is small, then changes in interest rates end up having only a small effect on real GDP. In the diagram, this shows up as a steep curve. The Fed's ability to use the monetary transmission mechanism to its advantage requires good knowledge of the shape of this relationship between interest rates and output.

Key Takeaways

1. The monetary transmission mechanism describes the links between the actions of the Fed and the state of the aggregate economy.
2. The Fed targets a short-term nominal interest rate called the federal funds rate. The Fed does not set this rate directly but rather uses its tools to influence this interest rate.
3. The main components of spending that depend on the real interest rate are spending by households on durable goods and investment. When these components of spending are sensitive to the interest rate, then the Fed can influence the economy through small variations in its target federal funds rate.

Exercises

1. Which interest rate determines investment spending—the real interest rate or the nominal interest rate?
2. Some newspapers state that the Fed sets the interest rate. Is that right?

This page titled 25.3: The Monetary Transmission Mechanism is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

25.4: Monetary Policy, Prices, and Inflation 25.4: Monetary Policy, Prices, and InflationEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. How do prices adjust in the economy?
2. What are the effects of monetary policy on prices and inflation?
3. What is the Taylor rule?

We now understand the effect of an interest rate increase on output. According to the monetary transmission mechanism, we expect that this will result in lower spending and a lower real gross domestic product (GDP). Remember, though, that the Fed is also charged with worrying about prices and inflation. Look back at the Federal Open Market Committee (FOMC) announcement with which we opened the chapter. Much of that announcement concerns inflation, not output. It states that "inflation and longer-term inflation expectations remain well contained," that "underlying inflation [is] expected to be relatively low," and that "the Committee will respond to changes in economic prospects as needed to fulfill its obligation to maintain price stability. "Federal Open Market Committee, "Press Release," *Federal Reserve*, February 2, 2005, accessed July 20, 2011, http://www.federalreserve.gov/boarddocs/press/monetary/2005/20050202/default.htm.

The statements by the Bank of England, the Central Bank of Egypt, and the Reserve Bank of Australia likewise betray a strong concern with inflation. The policy of many central banks is directed toward the inflation rate. This policy, appropriately called inflation targeting, focuses the attention of the monetary authority squarely on forecasting inflation and then controlling inflation through its current policy choices.

Price Adjustment and Inflation

The **inflation rate** is defined as the *growth rate* of the overall **price level**. In turn, the price level in the economy is based on the prices of all the goods and services in an economy. From one month to the next, some prices increase, others decrease, and still others stay the same. The overall inflation rate depends on what is happening to prices on average. If most prices are increasing and few are decreasing, then we expect to see inflation.

A complete explanation of inflation requires an understanding of all the decisions made by managers throughout the economy as they decide whether to change the prices of the goods and services that they sell. Some managers might find themselves facing increasing costs and strong demand for their product, so they would choose to increase prices. Others might have decreasing costs and weak demand, so they would choose to decrease prices. The overall inflation rate depends on the aggregation of these decisions throughout the economy and is summarized in a **price adjustment equation**. The price adjustment equation is shown in Figure 25.4.1.

Toolkit: Section 31.31

The net effect of all the price-setting decisions of firms yields a price adjustment equation, which is as follows:

$$inflation\ rate = autonomous\ inflation - inflation\ sensitivity \times output\ gap.$$

The price adjustment equation summarizes, at the level of the entire economy, all the decisions about prices that are made by managers throughout the economy. It tells us that there are two reasons for increasing prices. The first is that there may be underlying (autonomous) inflation in the economy, even when it is at potential output. This depends, among other things, on the inflation rate that firms anticipate. The second reason for increasing prices is if the output gap is negative. The output gap is the difference between potential output and actual output:

$$output\ gap = potential\ real\ GDP - actual\ real\ GDP.$$

A positive gap means that the economy is in recession—below potential output. If the economy is in a boom, then the output gap is negative.

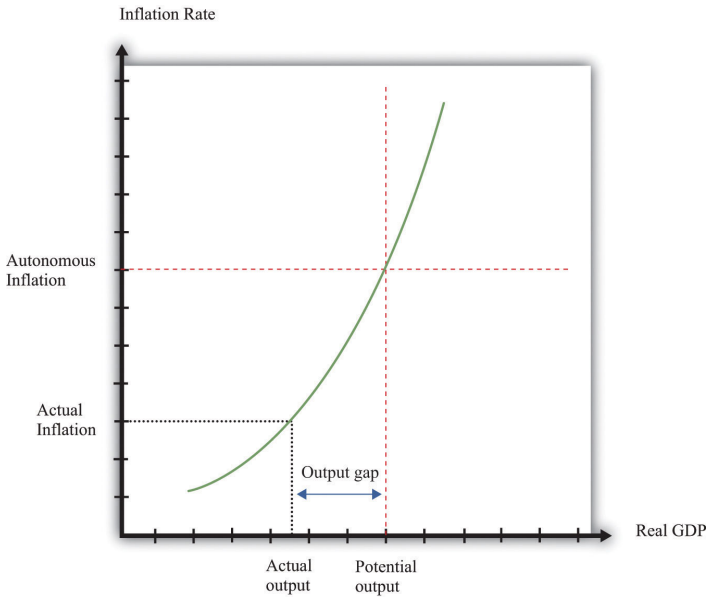


Figure 25.4.1: Price Adjustment. The price-adjustment equation tells us that when real GDP is below potential output, the output gap is positive, and the actual inflation rate is below its autonomous level. The opposite is true if real GDP is above potential output.

The output gap matters for inflation because as GDP increases relative to potential output, labor and other inputs become scarcer. Firms see increasing costs and increase their prices as a consequence. The second term of the price adjustment equation shows that when real GDP is above potential output (the output gap is negative), there is upward pressure on prices in the economy. The inflation rate exceeds autonomous inflation. By contrast, when real GDP is below potential output (the output gap is negative), there is downward pressure on prices. The inflation rate is below the autonomous inflation rate. The "inflation sensitivity" tells us how responsive the inflation rate is to the output gap.

If the output gap were the only factor affecting prices in the economy, then we would often expect to see **deflation**—decreasing prices. In particular, we would see deflation whenever the economy was in a recession. Although the United States and some other economies have occasionally experienced deflation, it is relatively rare. We can conclude that there must be factors other than the output gap that cause inflation to be positive.

Autonomous inflation is the inflation rate that prevails in the economy when the economy is at potential output (the output gap is zero). In the United States in recent decades, the inflation rate has been positive but low, meaning that prices have been increasing on average but at a relatively slow rate. Autonomous inflation is typically positive because most economies have some growth of the overall money supply in the long run. A positive output gap then translates not into deflation but simply into an inflation rate below the level of autonomous inflation. Thus in the FOMC statement with which we opened this chapter, the discussion is not about how contractionary policy will cause deflation; it is about how this policy will moderate the inflation rate. Positive autonomous inflation means that firms will typically anticipate that their suppliers or their competitors are likely to increase prices in the future. A natural response is to increase prices, so actual inflation is positive.

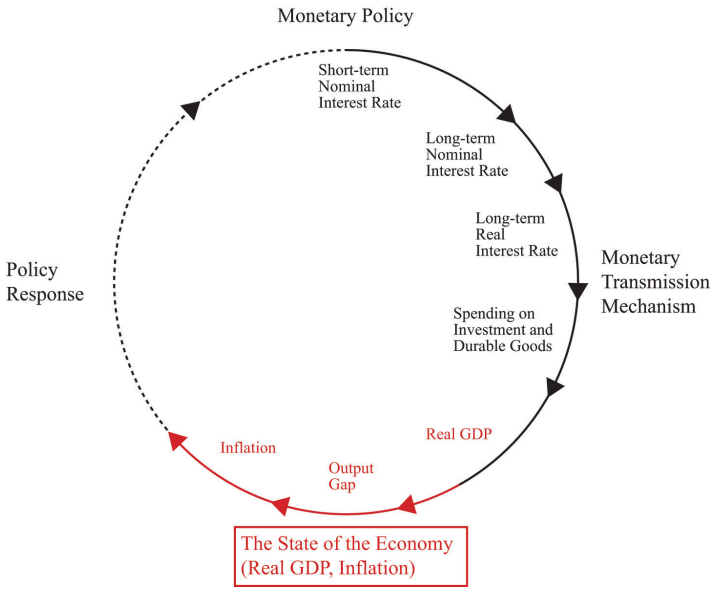


Figure 25.4.2: Interactions among Interest Rates, Output, and Inflation

The Effect of an Increase in Interest Rates on Prices and Inflation

The monetary transmission mechanism teaches us that an increase in real interest rates reduces spending and hence leads to a reduction in real GDP. In the (very) short run, the reduction in spending translates directly into a decrease in real GDP because prices are fixed. The reduction in GDP increases the output gap in the economy. Our price adjustment equation tells us in turn that this will tend to reduce the inflation rate in the economy.

Some firms will then adjust prices very quickly to the changing economic conditions. We do not think that the price level in the economy is literally fixed—unable to move—for any significant period of time. That said, some firms are likely to keep their prices unchanged for several months, even in the face of changing economic conditions. Thus the adjustment of prices in the economy takes some time. It will be months, perhaps years, before all firms have adjusted their prices.

In summary, an increase in interest rates leads to a gradual reduction in the inflation rate in the economy. Contractionary monetary policy leads to a reduction in economic activity and, over time, lower inflation. US monetary policy in the early 1980s provides a good illustration. At the start of that decade, the inflation rate was over 10 percent. To reduce inflation, the Fed, under Chairman Paul Volcker, conducted a contractionary monetary policy that sharply increased real interest rates. The immediate result was a severe recession, and the eventual result was a reduction in inflation, just as the model suggests.

Closing the Circle: From Inflation to Interest Rates

We have now traced the effects of monetary policy from interest rates to spending to real GDP to inflation. The effects of monetary policy do not stop there. Instead, as inflation adjusts in response to monetary policy, there is a feedback to interest rates through monetary policy itself. This is shown in Figure 25.4.3.

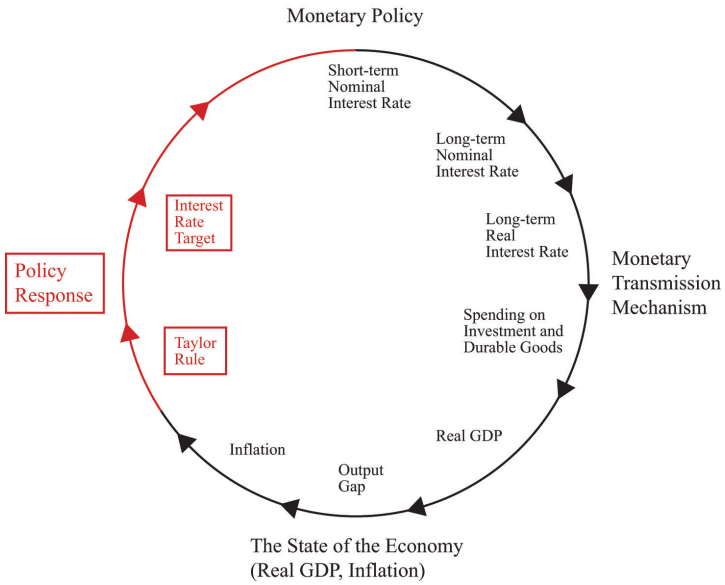


Figure 25.4.3: Completing the Circle of Monetary Policy. We close the monetary policy circle by observing that the Fed's policies depend on the state of the economy.

Observers of the Fed's behavior over the past 20 or so years have argued that the Fed generally follows a rule that makes its choice of a target interest rate somewhat predictable. The rule that summarizes the behavior of the Fed is sometimes called the **Taylor rule**; it is named after John Taylor, an economist who first characterized Fed behavior in this manner.Comments on John Taylor's career and his contributions to monetary economics by Fed Chairman Ben Bernanke are available at "Opening remarks to the Conference on John Taylor's Contributions to Monetary Theory and Policy, Federal Reserve Bank of Dallas, Dallas, Texas," Federal Reserve, October 12, 2007, accessed September 20, 2011, http://www.federalreserve.gov/newsevents/speech/bernanke20071012a.htm. The Taylor rule stipulates a relationship between the target interest rate and the state of the economy, typically represented by both the inflation rate and some measure of economic activity (such as the gap between actual and potential GDP). Usually, we think that the monetary authority operates with a lag so that the interest rate the monetary authority sets at a point in time reflects the output gap and inflation from the recent past. According to the Taylor rule, the Fed will increase real interest rates when

- inflation is greater than the target inflation rate,
- output is above potential GDP (a negative output gap).

Conversely, the Fed will decrease real interest rates when

- inflation is less than the target inflation rate,
- output is below potential GDP (a positive output gap).

The Fed will want to increase interest rates and thus "put the brakes on the economy" when inflation is high and when they think that real GDP is above its long-run level (potential output). The Fed will want to decrease interest rates when inflation is relatively low and the economy is in a recession.

An example of a Taylor rule is shown in Figure 25.4.4. The vertical axis is the real interest rate target of the Fed, and the horizontal axis is the inflation rate. As the inflation rate increases, the Fed, according to this rule, then increases the interest rate.

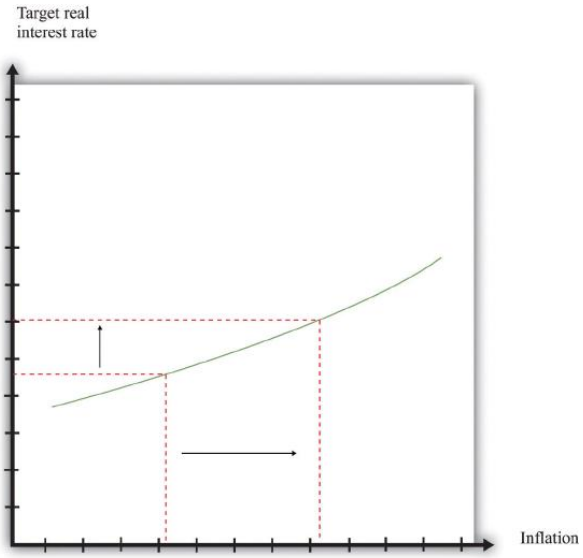


Figure 25.4.4: The Taylor Rule. The monetary policy rule shows how the Fed adjusts real interest rates in response to changes in inflation rates. As inflation increases, the monetary authority targets a higher real interest rate.

The different pieces of the Taylor rule can be in conflict. For example, the Fed may face a situation where inflation is relatively high, yet the economy is in recession. The precise specification of the rule then provides guidance as to how the Fed trades off its inflation and output goals. The rule is largely descriptive: it summarizes in a succinct manner the actions of the Fed. In doing so, it allows individuals to predict with some accuracy what actions the Fed is likely to take in the future.

The Taylor rule describes Fed policy in terms of the *real* interest rate. We know, however, that the Fed actually targets a nominal rate. This has a surprising implication when we examine how the Fed responds to inflation. Suppose the Fed is currently meeting its target inflation rate—say, 3 percent—and the federal funds rate is currently 5 percent. The real interest rate is therefore 2 percent (remember the Fisher equation). Now suppose the Fed sees that inflation has increased from 3 percent to 4 percent. The increase in the inflation rate has the effect of decreasing the real interest rate—again, this comes directly from the Fisher equation. The real interest rate is now only 1 percent. Yet the Taylor rule tells us that the Fed wants to *increase* the real interest rate. To do so, it must increase nominal interest rates by *more* than the increase in the inflation rate. In our example, the inflation rate increased by one percentage point, so the Fed will have to increase its target for the federal funds rate by more than one percentage point—perhaps to 6.5 percent.

The Taylor rule completes the circle of monetary policy. As indicated by Figure 25.4.3, the monetary policy rule links the state of the economy, represented by the inflation rate and the output gap, to the interest rate. There is usually a lag in the response of the Fed to the state of the economy. So, for example, the decision made at the FOMC meeting in February 2005 reflected information on the state of the economy through the end of 2004, at best.

In Summary: The Three Key Pieces of the Monetary Transmission Mechanism

We now have the three pieces we need to understand the relationship between monetary policy, inflation, and real GDP:

1. The Taylor rule linking the real interest rate to the inflation rate (Figure 25.4.4)
2. The inverse relationship between the real interest rate and real GDP (Figure 25.4.9)
3. The price adjustment process (Figure 25.4.1)

Together, these three pieces paint a complete picture of the monetary policy process. The top left panel in Figure 25.4.5 is taken from Figure 25.4.4 and shows a positive relationship between inflation and the real interest rate. The top right panel in Figure 25.4.5 is taken from Figure 25.4.9 and shows the relationship between real GDP and the interest rate. As shown in the figure, the higher the real interest rate, the lower real GDP is. As a reminder, higher real interest rates lead to lower aggregate spending. Finally, from the price-setting equation, changes in real GDP lead to changes in the inflation rate. We showed this previously in Figure 25.4.1, and it appears in the bottom right panel of Figure 25.4.5. If real GDP decreases, the output gap increases, and the inflation rate decreases.

We can use Figure 25.4.5 to summarize the conduct of monetary policy. In this diagram, we see the Taylor rule in action: the Fed sees high inflation and so increases the real interest rate.

- Start at the top right panel with “Last Period’s Interest Rate.” The panel shows us the level of real GDP that resulted from the interest rate choice. The bottom right panel then shows the inflation rate that came from the price adjustment equation. Point A therefore shows the state of the economy last period—that is, it shows last period’s inflation and last period’s real GDP. This is the information that the Fed uses when making its decision for this period.
- Given last period’s inflation rate, the top left panel shows us the value of the real interest rate that the Fed wants to choose this period. The Fed therefore sets a new target for the federal funds rate. This increases real interest rates, both short term and long term, which in turn leads to a decrease in durable goods spending.
- From the top right panel we can see that the Fed has chosen a higher interest rate than last period, which means that there is a decrease in real GDP.
- Decreased real GDP causes the inflation rate to decrease, as we see in the bottom right panel.
- Coming up to its next meeting, the FOMC again looks at the current state of the economy (point B), and the process begins again.

We have simplified the discussion here in two ways. First, we neglected the fact that the output gap also enters into the Taylor rule. The basic idea remains the same in that more complicated case. Second, we did not discuss autonomous inflation. Autonomous inflation, remember, captures managers’ expectations of future inflation and future demand conditions. It, too, will tend to change over time. Theories of autonomous inflation are a subject for more advanced courses in macroeconomics.

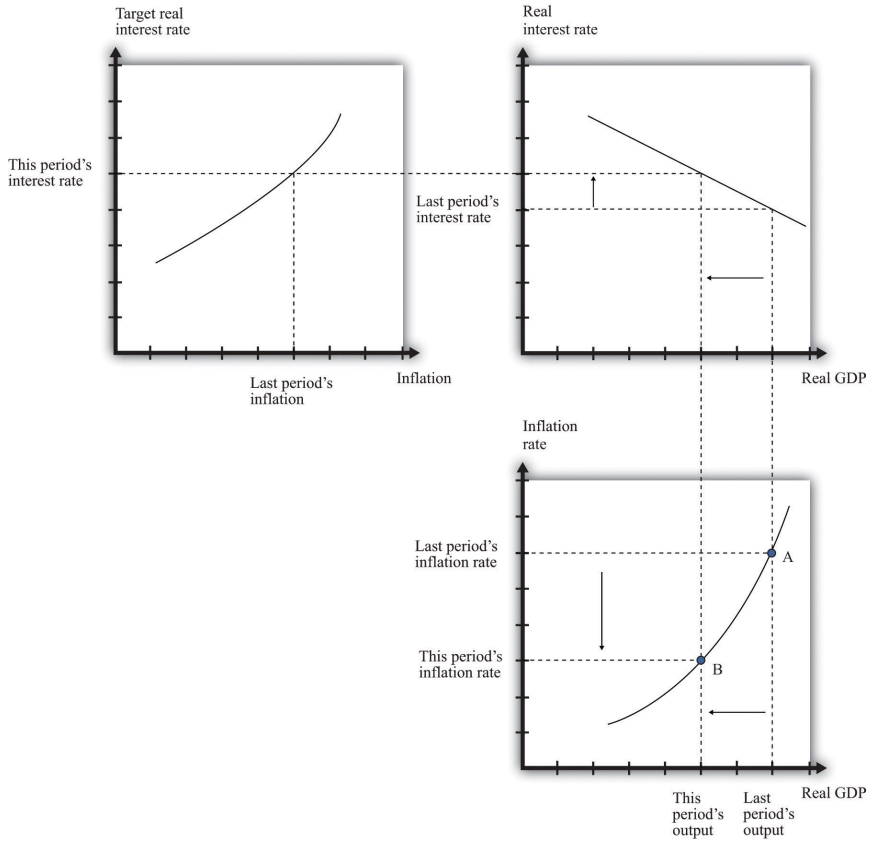


Figure 25.4.5: The Adjustment of Inflation over Time. Last period the economy was at point A, with high output and high inflation. Because inflation is too high, the Fed increases the real interest rate (top left). This reduces this period’s output (top right), which in turn leads to a reduction in the inflation rate (bottom right). The economy ends up at point B.

Key Takeaways

1. The price adjustment equation describes the dependence of price changes (inflation) on the output gap, given the autonomous inflation rate.
2. Given prices, monetary policy influences the output gap. Over time, prices adjust in response to the effects of monetary policy on the output gap.
3. The Taylor rule describes the dependence of the interest rate targeted by the Fed on the inflation rate and the output gap.

checking you understanding

1. Describe why a reduction in the target interest rate will ultimately lead to higher inflation.
2. If the economy is in a recession, what should happen to the target interest rate according to the Taylor rule?

This page titled 25.4: Monetary Policy, Prices, and Inflation is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

25.5: Monetary Policy in the Open Economy 25.5: Monetary Policy in the Open EconomyEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. How does monetary policy operate in an open economy?
2. How does monetary policy in other countries influence the US economy?

Monetary policy has international implications as well. Changes in interest rates lead to changes in supply and demand in the **foreign exchange market**. Chapter 24 explains this connection. In turn, changes in exchange rates affect exports and imports and influence the overall demand for goods and services. Among other things, this means that the monetary policy of other countries will have an effect on your own country. So if you live in Europe, you are not immune to Federal Open Market Committee (FOMC) actions. And if you live in the United States, you are not immune to the actions of the European Central Bank (ECB).

The Monetary Transmission Mechanism in the Open Economy

The key element in the monetary transmission mechanism is the ability of the central bank to influence the real interest rate. Changes in real interest rates lead to changes in spending on durable goods, which are a component of aggregate expenditures. But there is also another channel of influence. If the Fed cuts interest rates, for example, then the demand for dollars to invest in US asset markets will be reduced. This will reduce the foreign currency price of dollars. The weaker dollar means that goods produced in the United States are cheaper, so US exports will increase, and US imports will decrease. Thus changes in interest rates lead to changes in **exchange rates**, which in turn lead to changes in **net exports**. Net exports are also a component of aggregate expenditures. This is illustrated in Figure 25.5.1.

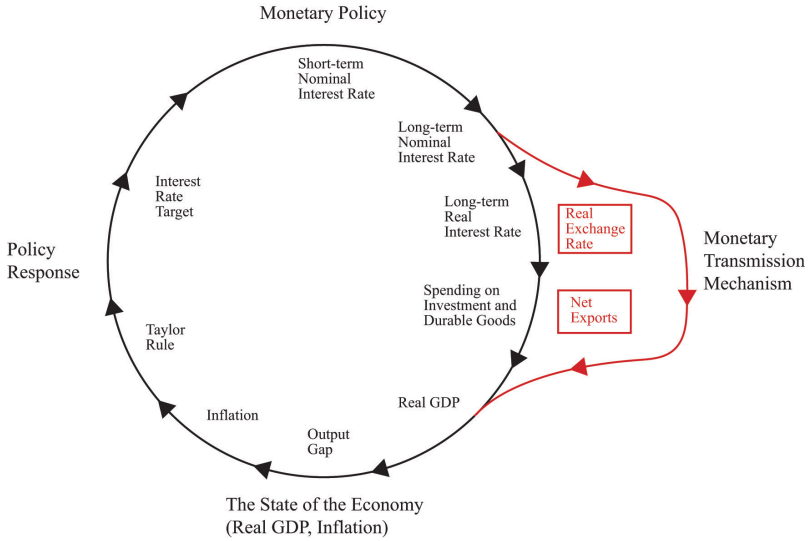


Figure 25.5.1. There is an additional channel of the monetary transmission mechanism that operates through the exchange rate. Changes in interest rates lead to changes in exchange rates, which in turn lead to changes in net exports. This channel reinforces the effect operating through interest rates.

Even when we include this channel, it is just as easy to understand the monetary transmission mechanism as it was before. When interest rates are cut, there is an increase both in spending on durables and net exports. Both channels lead to higher aggregate spending and thus higher output.

Toolkit: Section 31.20

You can review the workings of the foreign exchange market and the definition of the exchange rate in the toolkit.

Monetary Policy in the Rest of the World

The United States does not exist alone in the world economy. US financial markets are influenced by events in other countries, such as the actions of the ECB. Likewise, citizens in Europe are influenced by monetary policy in the United States.

Suppose the ECB cuts interest rates in Europe. As in the United States, the typical mechanism for this would be a purchase of debt issued by European governments. An increase in the price of this debt is equivalent to a decrease in interest rates. If nothing else happens, this decrease in European interest rates gives rise to an arbitrage opportunity. Investors want to move funds to the United States to take advantage of the higher interest rates. There is an increased demand for US assets and hence an increased demand for dollars. Interest rates in the United States decrease, which tends to increase durable goods spending and stimulate the US economy. Against that, the higher value of the dollar leads to fewer exports from the United States and more imports into the United States, so US net exports will decrease.

Completely analogously, monetary policy in the United States influences interest rates in other countries. If the Fed undertakes an open market sale of US government debt, for example, interest rates will increase in other countries as well as in the United States.

The US Federal Reserve and the ECB are big players in world financial markets. Their actions move world interest rates and world currency markets. There are other countries that are relatively small in the world economy. For example, suppose the Central Bank of Iceland increases interest rates in that country. The mechanisms that we have explained still apply: investors will find Icelandic assets more attractive, and there will be an increased demand for the Icelandic krona. However, the flows of capital into Iceland will be negligible in terms of the world economy. They will not have any noticeable effect on interest rates in Europe or the United States.

Key Takeaways

1. In an open economy, interest rate changes induced by monetary policy influence exchange rates and thus net exports.
2. Actions by monetary authorities in other countries influence the net exports of the United States through exchange rate changes and through the level of aggregate spending on the United States by households in other countries.

Exercises

1. If the Fed increases its target value for the federal funds rate, what happens to the value of the dollar?
2. If the ECB increases its target interest rate, what happens to US net exports?

This page titled 25.5: Monetary Policy in the Open Economy is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

25.6: The Tools of the Fed 25.6: The Tools of the FedEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What do banks do?
2. What are the tools of the Fed?

We have not yet said very much about exactly *how* the Fed changes interest rates. The Fed has three major tools at its disposal: open-market operations, the reserve requirement, and the discount rate. We discuss these in turn. Monetary policy operates through the Fed's interactions with the banking system, so we first must make sure we understand what banks do in the economy.If your find this material interesting, a course on *Money and Banking* will delve much further into the details of how banks operate and how they interact with the monetary authority. Throughout this discussion, we use the credit market to think about how the Fed operates.

Toolkit: Section 31.24

You can review the workings of the credit market in the toolkit.

What Do Banks Do?

Financial markets (that is, banks and other financial institutions) provide the link between savings and investment in the economy. A bank is a profit-making entity that takes in deposits from households and firms and makes loans to firms, households, and the government.

Banks can be fragile institutions. The fragility of banks is discussed in more detail in Chapter 22. They must ensure that their depositors are not worried that the bank might go out of business, taking their money with it. Banks do many things to ensure that their customers have confidence in them. Perhaps the most important is that they keep a certain amount of their assets in a very liquid form, such as cash. This means that if a depositor comes in to withdraw his or her money, the bank will be able to meet that demand. These liquid deposits are called the **reserves** of the bank.

Most banks in the United States are members of the Federal Reserve System. This membership comes with a responsibility to hold some fraction of deposits on reserve. This is called a **reserve requirement**. Current reserve requirements are at "Reserve Requirements," Federal Reserve, accessed September 20, 2011, http://www.federalreserve.gov/monetarypolicy/reservereq.htm#table1. Reserve requirements limit the amount of deposits that banks are able to loan out to firms and households. Suppose a bank has $1,000 on deposit and the reserve requirement is 10 percent. Then the bank must hold at least $100 on reserve and can loan out at most $900. We say "at least $100" since the bank is free to hold more than 10 percent on reserve. In uncertain times, when a bank is unsure how many depositors are likely to want to withdraw their money, the bank may choose to keep reserves above and beyond the level required by the Fed.

What does a bank do if it finds itself with insufficient reserves on a given day to meet its reserve requirements? The answer is that it borrows—either from other banks or from the Federal Reserve itself. Because the Federal Reserve can influence the interest rates at which banks borrow, it can influence the behavior of banks.

Open-Market Operations

In the memo with which we opened the chapter, the Federal Open Market Committee (FOMC) decided to increase the target federal funds rate to 2.5 percent. But what exactly does this mean, and how did the Fed accomplish it? The federal funds rate is the interest rate in a particular market—the market where banks make overnight loans to each other. Overnight loans, as the name suggests, are assets that have a very short time to maturity (one day). The interest rate on these loans is therefore one of the "shortest" interest rates in the economy, which is why it is targeted by the Fed. The interest rate is so named because the loans are made using the funds that banks have available in their accounts at the Federal Reserve.

The Federal Reserve does not participate directly in this market. It influences the federal funds rate by buying and selling in a different market—the market for short-term government debt. These purchases and sales are called **open-market operations**.Section 14 of the Federal Reserve Act describes open-market operations. Let us examine how this works. The effect of

open-market operations can be seen in the market for government debt. Part (a) of Figure 25.6.1 shows the supply and demand of this asset. The horizontal axis shows the quantity of assets (think of this as the amount traded on a given day), and the vertical axis shows the price of those assets. The participants in this market are financial institutions and others who hold, or want to hold, bonds as part of their portfolio of assets. Current owners will be willing to sell bonds if their price is sufficiently high. Conversely, if the price of bonds decreases, more people will want to purchase them. The same institution could be either a supplier or a demander, depending on the price. It is perfectly possible that a financial institution would want to buy bonds if their price were low and sell them if their price were high.

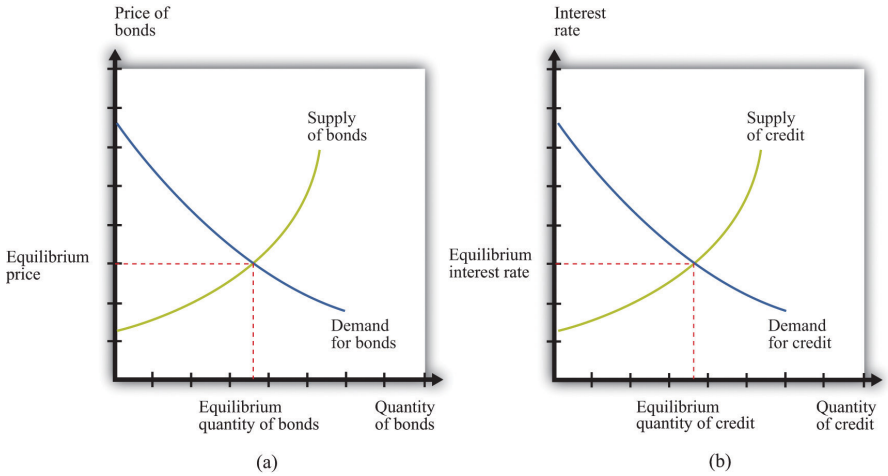


Figure 25.6.1: The Market for Government Bonds. (a) The price of bonds is determined by supply and demand. (b) These same transactions are represented in a credit market, which is another way of looking at exactly the same market.

Part (b) of Figure 25.6.1 shows the equivalent representation of this as a credit market. When the Fed buys bonds, it is making a loan. When the government or private investors sell bonds to the Fed, they are borrowing from the Fed. The crossing of the supply and demand curves tells us the equilibrium price of government bonds. It also tells us how many bonds changed hands that day, but our interest here is in what is happening to prices.

Now suppose the Federal Reserve steps into this market and buys some government bonds. This increases the demand for bonds, so the price of bonds will increase. This is shown in part (a) of Figure 25.6.2. Part (b) of Figure 25.6.2 shows the same action viewed through the lens of a credit market. Conversely, if the Fed decides to sell some of its stock of government bonds, the supply of bonds will shift out, and the price of bonds will decrease (see Figure 25.6.3).

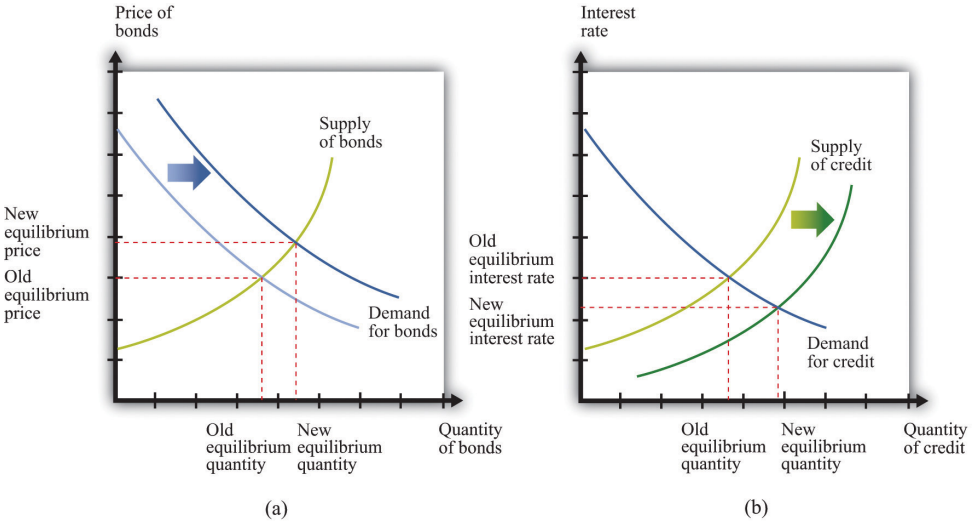


Figure 25.6.2: Intervention by the Federal Reserve. When the Federal Reserve conducts an expansionary open-market operation, it purchases bonds (a) or, equivalently, supplies more credit (b). The price of bonds increases, or, equivalently, the interest rate decreases.

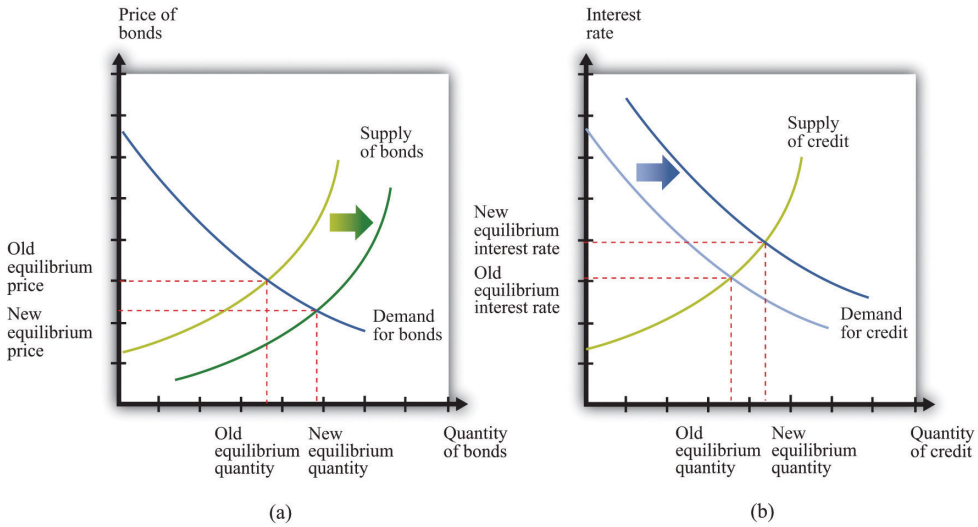


Figure 25.6.3: Intervention by the Federal Reserve. When the Federal Reserve conducts a contractionary open-market operation, it sells bonds (a) or, equivalently, demands more credit (b). The price of bonds decreases, or, equivalently, the interest rate increases.

Thus the Federal Reserve, by buying or selling government bonds in this market, has the ability to influence the price of bonds. This means that it can affect the interest rate on those bonds.

From this relationship, we know the following:

- If the Fed buys bonds, then the price of bonds increases, and interest rates decrease.
- If the Fed sells bonds, then the price of bonds decreases, and interest rates increase.

The Fed's actions in this market have an effect on interest rates in other markets, as banks and other financial institutions adjust their portfolios in response to the changing interest rate on government bonds. The Fed calibrates its buying and selling to try to achieve its target interest rate in the federal funds market.

The Discount Rate

The February 2005 announcement by the FOMC also included an increase in the discount rate. The discount rate is the interest rate from another market—in this case a market established by the Fed itself.

We have said that if a bank is short on reserves, it can borrow. One source of loans is the federal funds market. Another source of loans is the Fed itself. Member banks have the privilege of borrowing from the Fed, and the rate at which a bank can borrow is called the **discount rate**. The Fed directly controls this interest rate. The Federal Reserve's policies on such loans are set out in "Regulation A" of the Fed's Board of Governors: "A Federal Reserve Bank [that is, a Regional Fed] may extend primary credit on a very short-term basis, usually overnight, as a backup source of funding to a depository institution that is in generally sound financial condition in the judgment of the Reserve Bank. Such primary credit ordinarily is extended with minimal administrative burden on the borrower." "Regulation A (12 C.F.R. 201 as amended effective December 9, 2009)," Federal Reserve, accessed July 20, 2011, http://www.frbdiscountwindow.org/regulationa.cfm?hdrID=14&dtlID=77. Once a bank has established the right to borrow at the Fed's "discount window," the execution of such a loan is straightforward. The bank simply makes a toll-free call and provides a few pieces of basic information.

To see how this tool works, suppose the discount rate were very high, much higher than the interest the bank can earn by making a loan. Then the bank would find it prohibitively expensive to borrow from the Fed. If the bank were unsure that it could meet the needs of depositors, it would respond by holding reserves in excess of the reserve requirement. That is, with a very high discount rate, the bank would lend out a smaller fraction of its deposits. By contrast, if the Fed were to set the discount rate very low, the bank would make more loans and hold fewer reserves, safe in the knowledge that it could always borrow from the Fed if necessary.

From this reasoning, we can see that as the discount rate is increased, banks hold more excess reserves and lend less. This shows up in Figure 25.6.4 as a shift inward in the supply of credit. Thus the Fed *can increase interest rates by increasing the discount rate*.

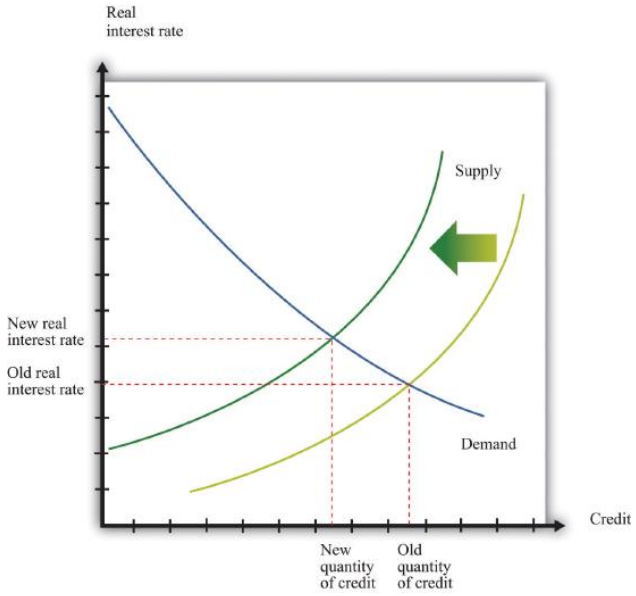


Figure 25.6.4: An Increase in the Discount Rate. An increase in the discount rate reduces the supply of credit and therefore increases the real interest rate.

Reserve Requirements

Reserve requirements are outlined in Section 19 (A) of the Federal Reserve Act:

(A) Each depository institution shall maintain reserves against its transaction accounts as the Board may prescribe by regulation solely for the purpose of implementing monetary policy—

1. in the ratio of 3 per centum for that portion of its total transaction accounts of $25,000,000 or less, subject to subparagraph (C); and
2. in the ratio of 12 per centum, or in such other ratio as the Board may prescribe not greater than 14 per centum and not less than 8 per centum, for that portion of its total transaction accounts in excess of $25,000,000, subject to subparagraph (C) [which stipulate that the reserve requirements could be changed].

Suppose the Fed were to increase the reserve requirement from 10 percent to 20 percent. In the previous example, all else being the same, a bank with deposits of $1,000 would be required to have at least $200 on deposit, rather than the $100 that was required originally. To fulfill this larger reserve requirement, the bank would be allowed to lend only $800 at most. Banks therefore respond to an increase in the reserve requirement by holding a larger fraction of deposits on reserve and lending out a smaller fraction of their deposits. This reduces the supply of credit in the economy since a smaller fraction of saving is actually being lent.

As shown in Figure 25.6.5, the supply of credit shifts inward, and the interest rate increases. This picture is exactly the same as Figure 25.6.4. When we think about the credit market, the increase in the discount rate and the increase in the reserve requirement have the same effect. Thus we learn that the Fed *can increase interest rates by increasing the reserve requirement*. Often, increases in the reserve requirement are coupled with other measures, such as open-market operations, to increase interest rates. A decrease in the reserve requirement works in a symmetric fashion, though in the opposite direction.

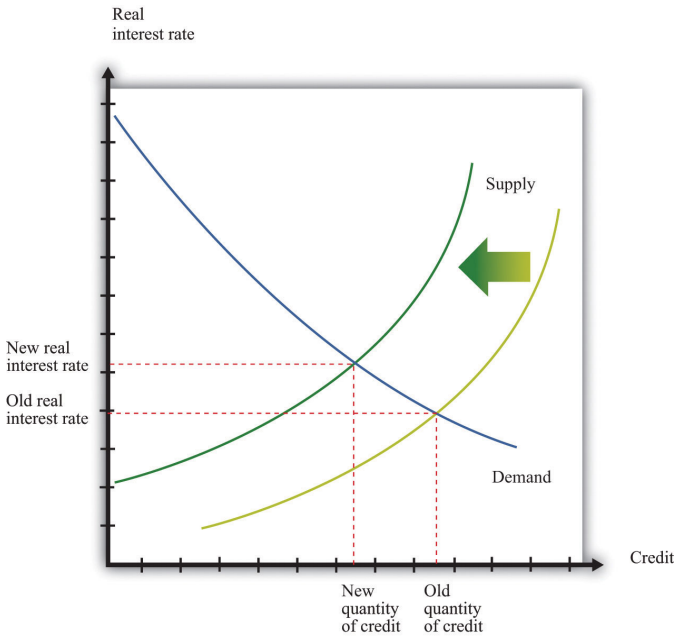


Figure 25.6.5: An Increase in Reserve Requirements. An increase in reserve requirements reduces the supply of credit and therefore increases the real interest rate.

Key Takeaways

1. Banks act as intermediaries, taking the deposits of households and making loans to firms and households who wish to borrow. Banks also borrow from other banks and from the Fed.
2. The main tools of the Fed are as follows: (a) open-market operations, (b) lending at the discount rate to member banks, and (c) setting the reserve requirements on member banks.

Exercises

1. Can a bank borrow from the Fed?
2. What are reserve requirements?
3. In an open market sale, does the money supply increase or decrease?

This page titled 25.6: The Tools of the Fed is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

25.7: The Fed in Action 25.7: The Fed in ActionEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What monetary policy did the Fed pursue during the Great Depression?
2. Why is stabilization of the economy through monetary policy so difficult?

We finish this chapter by going back to the actual actions of the Fed and focusing on two periods. First, we consider the Great Depression from a monetary perspective. Chapter 22 discusses that period in more detail and pays more attention to fiscal policy. Then we consider the period leading up to the February 2005 announcement.

The Great Depression Revisited

The Fed was in fact not very active during the Great Depression (some commentators might even say that this section should be titled "The Fed Inaction"). Yet monetary events were still critical.

A key short-term interest rate at that time was the so-called commercial paper rate. This rate decreased from about 6 percent in 1929 to a low of 0.8 percent by 1935. At first glance, therefore, it seems as if the monetary authority was implementing cuts in interest rates that could stimulate the economy. On closer examination, however, the picture is not so simple. During the Great Depression the inflation rate was negative—prices were decreasing on average. From the Fisher equation, a negative inflation rate means that the nominal interest rate *understates* the cost of borrowing. Decreasing prices mean that the nominal interest rate is smaller than the real interest rate. Even though nominal interest rates were decreasing in the early 1930s, the inflation rate was decreasing faster. As a result, the real interest rate increased. It became more expensive for households and firms to borrow, so spending decreased.

When prices decrease, the obligations of borrowers increase in real terms. People at the time did not typically anticipate these decreasing prices, so there was *unanticipated deflation*. Unanticipated deflation redistributes wealth from borrowers to lenders. Many firms, banks, and households were left with large (real) debts during the Great Depression. These led to bankruptcies and contributed to the contraction in economic activity.

Thus along with the high real interest rates came a series of bank failures. In addition, banks tended to hold more in excess reserves during this period, and thus loans, relative to deposits, decreased. These banking problems meant that the financial markets became less effective at connecting the savings of individual households with the investment plans of firms. It is perhaps not surprising that investment and spending on consumer durable goods decreased so much during the Great Depression.

In retrospect, the monetary authority could have been much more aggressive in dealing with the high real interest rates. They could have conducted open-market operations, buying bonds and decreasing interest rates. At the same time, this would have provided additional funds (sometimes called *liquidity*) to the banking system. Yet the Fed did not do so. Many observers now think that the severity of the Depression can be blamed in large part on these failures of the Fed. If so, this is good news, for it tells us that we are much more likely to be able to avert similar economic catastrophes in the future.

Monetary Policy from 1999 to 2005

Here is a brief summary of the target federal funds rate over the period from June 1999 to May 2005. Remember that these are nominal interest rates.

- Starting in June 1999, the target federal funds rate increased from 4.75 percent to 6.5 percent by January 2001.
- Starting in February 2001, the target federal funds rate decreased from 6.5 percent to a low of 1 percent by July 2004.
- In August 2004 the target federal funds rate was increased to 1.25 percent and was increased steadily to a level of 2.75 percent by May 2005.

We have already examined these targets, together with the actual federal funds rates, in Figure 25.7.3.

The time of tighter monetary policy, from June 1999 to January 2001, was a period of inflation concern. In the first part of 1999, the inflation rate averaged about 2 percent, and the unemployment rate was decreasing, reaching 4 percent in May 1999. Even though inflation was low, the Federal Open Market Committee (FOMC) statement from June 1999 called for an increase in the target federal funds rate, pointing to *potential inflation* as a rationale for increasing the target rate: "The Committee, nonetheless,

recognizes that in the current dynamic environment it must be especially alert to the emergence, or potential emergence, of inflationary forces that could undermine economic growth." Federal Open Market Committee, "Press Release," Federal Reserve, June 30, 1999, accessed August 8, 2011, http://www.federalreserve.gov/boarddocs/press/general/1999/19990630/default.htm. The Fed's tightening had the effect of reducing durable spending and thus bringing gross domestic product (GDP) down closer to potential output. As a consequence, there was less pressure on prices.

This policy continued through January 2001. By that point, the United States was very close to recession. (According to the National Bureau of Economic Research Business Cycle dating group, a recession began in March 2001.) From December 2000 to January 2001, the unemployment rate jumped from 3.7 percent to 4.7 percent. The Fed responded by allowing the federal funds rate to decrease steadily, starting in February 2001. This policy led to a federal funds rate of 1 percent by July 2003, a level that was maintained for a year. Historically, this was a very low rate. Over the year, inflation averaged about 2.3 percent, so the real federal funds rate was actually negative.

A turnaround in Fed policy occurred in August 2004. Inflation had started to increase somewhat in early 2004, and the unemployment rate had decreased to 5.3 percent in May 2004. So in August 2004, the Fed started a gradual increase of the target federal funds rate. Look back at Figure 25.7.3. Recall that part of the monetary transmission mechanism is the link between the nominal federal funds rate, which is very short term, and much longer-term rates. Figure 25.7.3 shows the federal funds rate along with the 1-year and 10-year Treasury bond yields. The loosening of monetary policy in February 2001 is evident from the decrease in the federal funds rate and the 1-year Treasury rate.

But the long-term Treasury rate seems not to follow the short-term rates that closely. In fact, it seems that the long-term rates started to decrease before the reductions in the federal funds rate began, and then the long-term rates did not decrease nearly as much over the February 2001–August 2004 period. After that time, although the federal funds rate was increased, the long-term rate did not respond much at all.

This reminds us of one the biggest challenges of monetary policy. Although the Fed is able to closely target the federal funds rate, it has much less ability to control longer-term rates. Someone making a loan for a long period of time will try to anticipate economic events over the course of the entire loan period. As a consequence, the loan rate may reflect anticipated events (such as the Fed's loosening of monetary policy in February 2001) and may also not respond as much to rate changes that are seen as temporary.

Why Do Central Bankers Get Paid So Much?

We have made monetary policy look easy. The effects of the actions of the monetary authority are summarized by Figure 25.7.1. Given a choice of a target inflation rate and a target level of economic activity, the Fed (and other central banks) ought to know exactly what to do to reach these goals. So why are central bankers so vital to the functioning of the macroeconomy?

What Is the State of the Economy?

In 25.3.4 Section "Closing the Circle: From Inflation to Interest Rates", we described the Taylor rule as relating the target federal funds rate to the state of the economy, specifically the inflation rate and the output gap. As a matter of theory, this is straightforward to describe. The practice is rather harder.

First, it is a significant challenge simply to know the current state of the economy. In the United States, part of the preparation for FOMC meetings is an attempt to figure out the current output gap and other variables. The Board of Governors of the Federal Reserve has a large staff of professional economists, as do the various regional Federal Reserve banks. These economists spend much of their time helping the members of the FOMC understand the current state of the economy.

One particular problem is that the level of real GDP itself is calculated only on a quarterly basis. Potential GDP, meanwhile, is a theoretical construct that requires some guesses about "full employment." It is not directly measured. So if the Fed learns that real GDP is growing rapidly, it has to judge whether this is because potential GDP is growing rapidly or because actual GDP is above potential.

Since the Fed does not meet to determine policy each day and the Fed's policies themselves take time to work through the economy, it is not even enough to know the current state of the economy. The FOMC must also forecast the state of the economy for the near future. One talent of the previous Fed chairman, Alan Greenspan, was apparently his use of relatively unorthodox sources to get a sense of the state of the economy.

What Are the Effects of Monetary Policy?

Even if there were no uncertainty about the current state of the economy—that is, the inflation rate and the output gap—monetary policy is still difficult for other reasons. First, as we emphasized earlier, the Fed does not have direct control over the long-term real interest rates that matter for durable goods spending. The Fed can influence a short-term nominal rate, which in turn influences the long-term real rates. But the exact link from one interest rate to the other is not known by the Fed and may change over time. The Fed may fail to achieve the long-term rate that it is aiming for.

Second, the Fed does not have perfect knowledge of the monetary transmission mechanism. Consider again the links between real interest rates and output, as shown in Figure 25.7.9. In reality, the Fed does not know exactly what the relationship between interest rates and output looks like. Reality looks more like Figure 25.7.1. In this picture the Fed is aiming for a high level of output. However, it misses its target real interest rate and actually ends up setting a higher real rate than it wanted. In addition, real GDP is more sensitive to interest rates than it thought, so the high rate leads to a big reduction in GDP. Thus because the Fed fails to achieve its target interest rate and also misjudges the monetary transmission mechanism, it ends up with much lower real GDP than it wanted.

Finally, the Fed has imperfect knowledge of the link between economy activity and price adjustment. Recall that the price setting equation stipulates that inflation depends on the output gap and something called autonomous inflation. As we have seen, this last term captures several factors, including the influence of expectations about the future on current price-setting behavior. This presents a double challenge to the Fed. First, to evaluate the effects of its policy on prices, the Fed needs to know the expectations that underlie autonomous inflation. Second, the Fed must recognize that *its actions and statements influence these expectations*. This is why the individuals involved in the making of monetary policy are so careful both about what they do and about what they say about what they do.

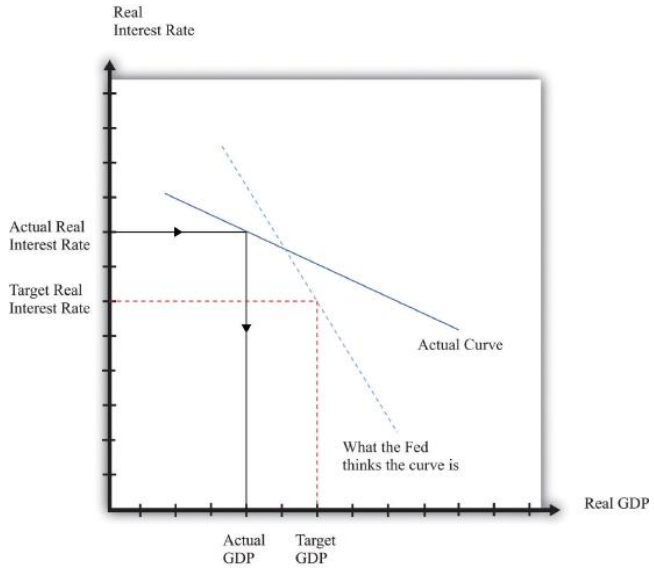


Figure 25.7.1: Controlling the Economy. The Fed's ability to control the economy depends on how knowledgeable it is about the state of the economy and on how accurately it can target interest rates.

What Should the Fed Do When Its Goals Are in Conflict?

We know that the goals of the Fed include price and output stability. Sometimes these goals conflict, and when they do, the task of central bankers becomes even more complicated.

The FOMC statement with which we opened this chapter stated that the "Committee perceives the upside and downside risks to the attainment of both sustainable growth and price stability for the next few quarters to be roughly equal." But what if instead it had said the "Committee perceives the risks of low output growth and high inflation for the next few quarters to be roughly equal"? What would the appropriate monetary policy be in this case? Should the Fed use its power to stabilize prices or to promote economic activity?

The tension is evident from the Taylor rule. Here is an example: the target real interest rate increases when inflation is high and decreases when the output gap is high:

$$real\ interest\ rate = -(1/2) \times (output\ gap) + (1/2) \times (inflation\ rate - 4\ percent)$$

$$\text{real interest rate} = (1/2) \times (\text{output gap}) + (1/2) \times (\text{inflation rate} - 4 \text{ percent}).$$

Remember that a positive output gap means that that the economy is in a recession: actual GDP is below potential. When the economy is in recession and inflation is not very high, the Taylor rule says that the Fed should reduce the real interest rate. And—from this same rule—the Fed should increase the real interest rate in the face of high inflation and a negative output gap. But what should the Fed do when inflation is high *and* there is a recession? High inflation argues for increasing real interest rates, but a positive output gap argues for a cut in rates.

The Fed—and, indeed, monetary authorities throughout the world—faced exactly this conflict in the mid-1970s when oil prices increased substantially as a result of actions by the Organization of Petroleum Exporting Countries. Researchers who have examined data over the past three decades have found that an increase in oil prices is typically met with an increase in the federal funds rate. The following discussion elaborates on the Fed's response to oil price increases: Federal Reserve Bank of Cleveland, accessed July 20, 2011, www.clevelandfed.org/Research/inflation/Readingroom/Viewpoint/2005/oil-prices-economy04-05.cfm. A speech by then Fed Governor Ben Bernanke in 2004 provides more details: "Remarks by Governor Ben S. Bernanke at the Distinguished Lecture Series, Darton College, Albany, Georgia," Federal Reserve, October 21, 2004, accessed July 20, 2011, http://www.federalreserve.gov/boardDocs/speeches/2004/20041021/default.htm. Thus, when faced with conflicting goals stemming from an oil price increase, the Fed seems to have put more weight on the goal of price stability.

When Things Go Badly Wrong

Everything that we have talked about in this section helps to explain why central bankers must be skilled and knowledgeable individuals with a good grasp of both economics and the workings of financial markets. Still, we have essentially been describing the job of a technocrat. Central bankers really earn their salaries in abnormal rather than normal times.

Starting in 2007 and stretching well into 2008, the United States and other countries began to experience financial crises that were similar in some ways to those experienced in the Great Depression. The financial crisis of 2008 is discussed in Chapter 19 and Chapter 30. The crisis seemed to begin innocently enough, with a decrease in housing prices that left some people unable or unwilling to cover their mortgage payments. But because of the way financial markets work, it became very hard for lenders to work out which of their assets were "nonperforming"—that is, unlikely to be repaid. As a result, financial markets froze up.

Part of the Fed's response was an aggressive use of the tools that we have described in this chapter. For example, the Fed reduced the federal funds rate down to 0.25 percent. At that point, the Fed had just about reached the limit of what was possible with monetary stimulus. The problem is that nominal interest rates cannot go below zero because cash has a nominal interest rate of zero. If you keep a dollar bill from this year to next year, it is worth $1 next year. Therefore it would always be better just to keep cash rather than invest in an asset with a negative nominal return. The Fed had hit what is known as the **zero lower bound**.

Even though it was at the zero lower bound, the Fed still had other options. In normal circumstances, it operates in the economy by buying and selling short-term government debt, one of the many assets in the economy. But these were highly abnormal circumstances, and it is possible for the Fed to buy and sell other assets as well. This is what the Fed did. During the crisis, the Fed started purchasing many other assets, such as commercial paper. In other words, instead of just lending to banks, the Fed started lending directly to firms in the economy. Central banks in some other countries, such as the United Kingdom, pursued similar policies.Explaining what happened in 2008 involves understanding the actions of the Fed, but it requires many of our other tools as well. For that reason, we take up this crisis in more detail in Chapter 30.

Key Takeaways

1. Despite the large reduction in aggregate economy activity and deflation during the Great Depression, the Fed did not pursue a very aggressive policy. The effectiveness of the Fed was hampered by the unwillingness of households to deposit funds in banks and the unwillingness of banks to make loans.
2. The conduct of monetary policy is made difficult by uncertainty over the current state of the economy and the inexact nature of the effects of interest rates on real GDP and prices.

Exercises

1. In what ways was the Fed not very aggressive during the Great Depression?
2. How could the goals of the Fed be in conflict?
3. Does the Fed know the current state of the economy when it makes decisions?

This page titled 25.7: The Fed in Action is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

25.8: End-of-Chapter Material 25.8: End-of-Chapter MaterialEdit Update titlecancel

In Conclusion

A driving analogy is sometimes used to illustrate the problems of the Fed. In the best of all worlds, we would drive a car in perfect weather along straight, wide, dry roads. We would look out crystal clear windows with complete knowledge of exactly where we are on the road and what driving conditions are like up ahead. Then, with complete control over the car, we could adjust speed and direction to reach our destination.

This is not the right picture for monetary policy. Instead, the windshield is very dirty, obscuring current conditions and making predictions almost impossible. Although the driver is well trained, the connection between the tools of the car and its direction and speed is haphazard.

Suppose the driver sees a steep downhill in the distance that requires some slowing down. Putting on the brakes will eventually slow the car down, but the delay is hard to predict. Making matters worse, by the time the car slows, the road may be going uphill again.

More precisely, the first challenge for the Fed is determining the current state of the economy. The Fed must rely on economic data to determine the current state of the economy. This is not easy; data often arrive with lags and with measurement error. Furthermore, the data often provide conflicting signals about the current state of the economy.

The second challenge for the Fed is that the transmission mechanism is not cast in stone. Reducing real interest rates by, say, one percentage point does not create the same response in spending at all times. Instead, the links in the monetary transmission mechanism change over time and depend on numerous other variables in the economy. Understanding these links remains a key area of research in economics and is also a challenge for those responsible for the conduct of monetary policy.

Key Links

- Board of Governors purposes and functions: http://www.federalreserve.gov/aboutthefed/default.htm
- Federal Reserve Act: http://www.federalreserve.gov/generalinfo/fract/default.htm
- Board of Governors, Federal Open Market Committee (FOMC), monetary policy tools: http://www.federalreserve.gov/monetarypolicy/fomc.htm
- European Central Bank: http://www.ecb.int/home/html/index.en.html
- History of money
  - Public Broadcasting System: www.pbs.org/newshour/on2/money/history.html
  - Federal Reserve Bank of Minneapolis: www.minneapolisfed.org/community_education/teacher/history.cfm

exercises

1. Have you ever noticed that banks are often housed in big imposing buildings? Why do you think this is the case?
2. Consider a Taylor rule given by

$$real\ interest\ rate = -(1/2) \times (output\ gap) + (1/2) \times (inflation\ rate - 4\ percent).$$

   1. Describe this rule in words. What is the target inflation rate in this rule?
   2. If the inflation rate is 6 percent and the GDP gap is −2 percent, what should the real interest rate be? What nominal interest rate should the Fed set?

3. (Advanced) Draw a version of Figure 25.8.4 where you show how to relate the target interest rate to the output gap. Explain in words what it means to move along the curve. What shifts the curve you have drawn?
4. What would happen if the Fed set the discount rate below the rate of return on government bonds?
5. Do open-market operations have to be in the form of the Fed buying and selling government debt? Could an open-market operation occur with the Fed buying the stock of a company?
6. Explain why an increase in interest rates reduces the demand for durable goods.
7. Suppose the relationship between investment and interest rates is $investment = 100 - 4 \times real\ interest\ rate$ and suppose the multiplier is 2. If the interest rate decreases by one percentage point, what happens to real GDP (assuming no change in the price level)?

8. Give two reasons why it is difficult to conduct monetary policy.
9. Suppose the central bank in country A is more worried about inflation than the output gap, but the opposite is true in country B. What differences in the Taylor rule would you expect to see in the two countries? Must it be the case that country A has a lower target inflation rate than country B?
10. Explain why a positive output gap does not necessarily lead to decreasing prices.

Economics Detective

1. Find the most recent announcement of the Federal Open Market Committee (FOMC). How does it differ from the one from February 2, 2005? Who is currently on the FOMC?
2. Use the site http://www.hsh.com/calc-payment.html to calculate how your monthly payment would change as you vary the interest rate charged on a car loan for a $30,000 car. This will give you a sense of how actions of the Fed would affect your monthly payments on a loan.
3. Find the names of five other central banks in the world economy. Find some information about their history (when were they established, for example), their design (are they independent?), and their operating procedures.
4. Find the web page for the Board of Governors of the Federal Reserve System and read about the tools of monetary policy. Based on your reading, (a) how often does the FOMC meet, and (b) how is its membership determined?
5. If you live in the United States, find the web page for the regional Fed closest to you. Try to find its most recent report on local economic conditions. Do you agree with this assessment of the local economy? What can you learn about the president of the regional Fed? What about the director of research, who is the staff member most likely to give advice to the president of the regional Fed about monetary policy?
6. Using your web research skills, find a discussion of Fed policy during times of high oil prices. How did the Fed resolve the tensions between increasing rates to combat inflation and decreasing rates to deal with unemployment? Try to find data on (real) oil prices and the federal funds rate. Did these two economic variables move together during periods of high oil prices?
7. In March 2008, the Fed opened the discount window to add liquidity into the financial system. Find the policy statements associated with this action and describe exactly what the Fed did.
8. Get data on the US economy to see how well the Taylor rule,

$$real\ interest\ rate = -(1/2) \times (output\ gap) + (1/2) \times (inflation\ rate - 4\ percent),$$

   fits the facts for the past five years.
9. Find an occasion when the Fed has changed reserve requirements. Did it also make other policy adjustments at the same time?

This page titled 25.8: End-of-Chapter Material is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

26: Inflations Big and Small

26.1: Rising Prices
26.2: The Quantity Theory of Money
26.3: Facts about Inflation and Money Growth
26.4: The Causes of Inflation
26.5: The Costs of Inflation
26.6: Policy Remedies
26.7: End-of-Chapter Material

This page titled 26: Inflations Big and Small is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

26.1: Rising Prices 26.1: Rising PricesEdit Update titlecancel

Through the years, people have been willing to wear some absurd slogans on their clothing. But surely one of the worst was the "WIN" button, introduced by United States President Gerald Ford in a speech on October 8, 1974. President Ford's speech can be read and heard here: http://www.fordlibrarymuseum.gov/library/speeches/740121.asp. The button, shown in the following figure, was the symbol of a campaign against a perceived a social evil. And what was this great evil? It was inflation. "WIN" stood for "whip inflation now." President Ford asked citizens to wear WIN buttons as a sign that they were enlisted in the battle against inflation.



Figure 26.1.1: Button: Whip Inflation Now

Wearing buttons might not have been the first bit of advice economists would have given to a leader interested in battling inflation. But this episode makes it evident that President Ford and his advisors viewed inflation as a major social problem. The president even invoked wartime imagery, concluding his speech by saying the following: President Ford's speech can be read and heard here: http://www.fordlibrarymuseum.gov/library/speeches/740121.asp.

Only two of my predecessors have come in person to call upon Congress for a declaration of war, and I shall not do that. But I say to you with all sincerity that our inflation, our public enemy number one, will, unless whipped, destroy our country, our homes, our liberties, our property, and finally our national pride, as surely as any well-armed wartime enemy.

I concede there will be no sudden Pearl Harbor to shock us into unity and to sacrifice, but I think we have had enough early warnings. The time to intercept is right now. The time to intercept is almost gone.

My friends and former colleagues, will you enlist now? My friends and fellow Americans, will you enlist now? Together with discipline and determination, we will win.

When President Ford initiated this campaign, the US inflation rate was about 12 percent. In other words, a shirt that cost $10.00 in 1973 cost about $11.20 in 1974. This was the highest inflation rate that the United States had experienced since World War II. Inflation running at this rate is, at the very least, a significant inconvenience.

Still, compared to the experience of many countries, this level of inflation is negligible. Between World War I and World War II, Germany, Hungary, Austria, and Poland experienced massive rates of inflation. In one month in 1923, the annual inflation rate in Germany was 6,829 percent. This number is very difficult to fathom; it is astronomical compared to the inflation that President Ford was facing. At this rate of inflation, *prices were doubling every three to four days*.

Such rapid price increases forced people to change their behavior in extraordinary ways. The instant workers received their pay, they would rush out and spend it, for even a delay of a few hours could mean that your wages would buy fewer goods and services. Even ordering in a café became a game to beat inflation: "The price increases began to be dizzying. Menus in cafes could not be revised quickly enough. A student at Freiburg University ordered a cup of coffee at a cafe. The price on the menu was 5,000 Marks. He had two cups. When the bill came, it was for 14,000 Marks. 'If you want to save money," he was told, "and you want two cups of coffee, you should order them both at the same time.'" This comes from a PBS page with many excerpts about the German hyperinflation: "The German Hyperinflation, 1923," Commanding Heights: The Battle for the World Economy, PBS, accessed

September 20, 2011, http://www.pbs.org/wgbh/commandingheights/shared/minitext/ess_germanhyperinflation.html. And these are not just stories from long ago. In the past 25 years, there have been large inflations in Yugoslavia, Israel, Argentina, Brazil, Mexico, Ukraine, and Zimbabwe, for example.

What is the cause of inflation?

Road Map

In this chapter, we study the causes and consequences of inflation. Times of rapid inflation are especially helpful for understanding inflation in general. When inflation is the dominant feature of an economy, it is very easy to isolate the main forces at work. We will see, moreover, that the most interesting periods to study are the beginning and end of large inflations, for such times provide a particular insight into the connection between fiscal policy and monetary policy.

We first study the relationship between the inflation rate and changes in the amount of money circulating in an economy and explain that, in the long run, there is a close connection between the inflation rate and the growth rate of the money supply. We look at some data both for the United States and for other countries and examine some examples of hyperinflation. Then we explore the underlying cause of hyperinflations, which turn out to be connected to the tax and spending choices that governments make, and we conclude by discussing government policy to control inflation.

This page titled 26.1: Rising Prices is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

26.2: The Quantity Theory of Money 26.2: The Quantity Theory of MoneyEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions.

1. What is the quantity theory of money?
2. What is the classical dichotomy?
3. According to the quantity theory, what determines the inflation rate in the long run?

We begin by presenting a framework to highlight the link between money growth and inflation over long periods of time.The framework complements our discussion of inflation in the short run, contained in Chapter 25. The **quantity theory of money** is a relationship among money, output, and prices that is used to study inflation. It is based on an accounting identity that can be traced back to the **circular flow of income**. Among other things, the circular flow tells us that

$$nominal\ spending = nominal\ gross\ domestic\ product\ (GDP).$$

The "nominal spending" in this expression is carried out using money. While money consists of many different assets, you can—as a metaphor—think of money as consisting entirely of dollar bills. Nominal spending in the economy would then take the form of these dollar bills going from person to person. If there are not very many dollar bills relative to total nominal spending, then each bill must be involved in a large number of transactions.

The **velocity of money** is a measure of how rapidly (on average) these dollar bills change hands in the economy. It is calculated by dividing nominal spending by the money supply, which is the total stock of money in the economy:

If the velocity is high, then for each dollar, the economy produces a large amount of nominal GDP.

Using the fact that nominal GDP equals $real\ GDP \times the\ price\ level$, we see that

And if we multiply both sides of this equation by the money supply, we get the **quantity equation**, which is one of the most famous expressions in economics:

$$money\ supply \times velocity\ of\ money = price\ level \times real\ GDP.$$

Let us see how these equations work by looking at 2005. In that year, nominal GDP was about $13 trillion in the United States. The amount of money circulating in the economy was about $6.5 trillion. In Chapter 24, we discussed the fact that there is no simple single definition of money. This figure refers to a number called "M2," which includes currency and also deposits in banks that are readily accessible for spending. If this money took the form of 6.5 trillion dollar bills changing hands for each transaction that we count in GDP, then, on average, each bill must have changed hands twice during the year $(13/6.5 = 2)$. So the velocity of money was 2 in 2005.

Toolkit: Section 31.27

You can review the circular flow of income in the toolkit.

The Classical Dichotomy

So far, we have just written a definition. There are two steps that take us from this definition to a theory of inflation. First we use the quantity equation to give us a theory of the price level. Then we examine the *growth rate* of the price level, which is the inflation rate.

In macroeconomics we are always careful to distinguish between nominal and real variables:

- **Nominal variables** are defined and measured in terms of money. Examples include nominal GDP, the nominal wage, the dollar price of a carton of milk, the price level, and so forth. (Most nominal variables are measured in monetary units, but some are just numbers. For example, the nominal interest rate tells you how many dollars you will obtain next year for each dollar you invest in an asset this year. It is thus measured as "dollars per dollar," so it is a number.)

- All variables not defined or measured in terms of money are **real variables**. They include all the variables that we divide by a price index in order to correct for the effects of inflation, such as real GDP, real consumption, the capital stock, the real wage, and so forth. For the sake of intuition, you can think of these variables as being measured in terms of units of (base year) GDP (so when we talk about real consumption, for example, you can think about the actual consumption of a bundle of goods and services by a household). Real variables also include the supply of labor (measured in hours) and many variables that have no specific units but are just numbers, such as the velocity of money or the capital-to-output ratio of an economy.

Prior to the Great Depression, the dominant view in economics was an economic theory called the **classical dichotomy**. Although this term sounds imposing, the idea is not. According to the classical dichotomy, real variables are determined independently of nominal variables. In other words, if you take the long list of variables used by macroeconomists and write them in two columns—real variables on the left and nominal variables on the right—then you can figure out all the real variables without needing to know any of the nominal variables.

Following the Great Depression, economists turned instead to the **aggregate expenditure model** to better understand the fluctuations of the aggregate economy. In that framework, the classical dichotomy does not hold. Economists still believe the classical dichotomy is important, but today economists think that the *classical dichotomy only applies in the long run*.

The classical dichotomy can be seen from the following thought experiment. Start with a situation in which the economy is in equilibrium, meaning that supply and demand are in balance in all the different markets in the economy. The classical dichotomy tells us that this equilibrium determines *relative prices* (the price of one good in terms of another), not absolute prices. We can understand this result by thinking about the markets for labor, goods, and credit.

Figure 26.2.1 presents the labor market equilibrium. On the vertical axis is the real wage because households and firms make their labor supply and demand decisions based on real, not nominal, wages. Households want to know how much additional consumption they can get by working more, whereas firms want to know the cost of hiring more labor in terms of output. In both cases, it is the real wage that determines economic choices.

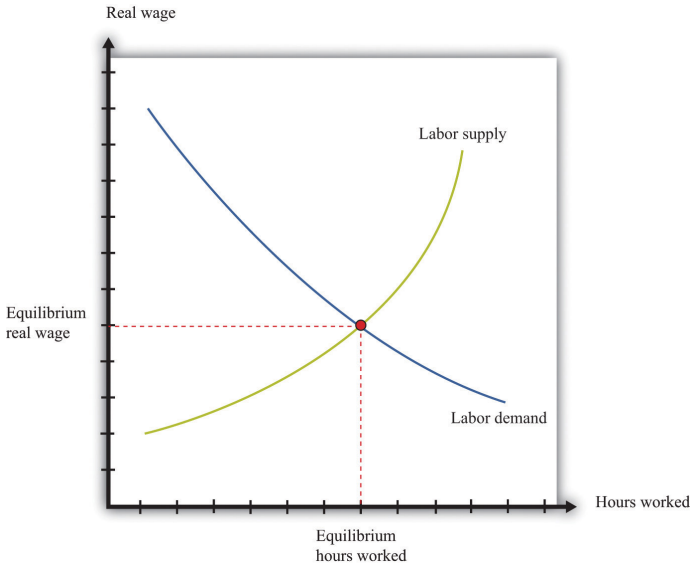


Figure 26.2.1: Labor Market Equilibrium

Now think about the markets for goods and services. The demand for any good or service depends on the real income of households and the real price of the good or service. We can calculate real prices by **correcting for inflation**: that is, by dividing each nominal price by the aggregate price level. Household demand decisions depend on real variables, such as real income and relative prices.If you have studied the principles of microeconomics, remember that the budget constraint of a household depends on income divided by the price of one good and on the price of one good in terms of another. If there are multiple goods, the budget constraint can be determined by dividing income by the price level and by dividing all prices by the same price level. The same is true for the supply decisions of firms. We have already argued that labor demand depends on only the real wage. Hence the supply of output also depends on the real, not the nominal, wage. More generally, if the firm uses other inputs in the production process, what matters to the firm's decision is the price of these inputs relative to the price of its output, or—more generally—relative to the overall price level.If you have studied the principles of microeconomics, the condition that price equals marginal cost is used to characterize the output decision of a firm. What matters then is the price of the input, relative to the price of output.

What about credit markets? The supply and demand for credit depends on the real interest rate. This means that those supplying credit think about the return they receive on making loans in real terms: although the loan may be stated in terms of money, the supply of credit actually depends on the real return. The same is true for borrowers: a loan contract may stipulate a nominal interest rate, but the real interest rate determines the cost of borrowing in terms of goods. The supply of and demand for credit is illustrated in Figure 26.2.2.

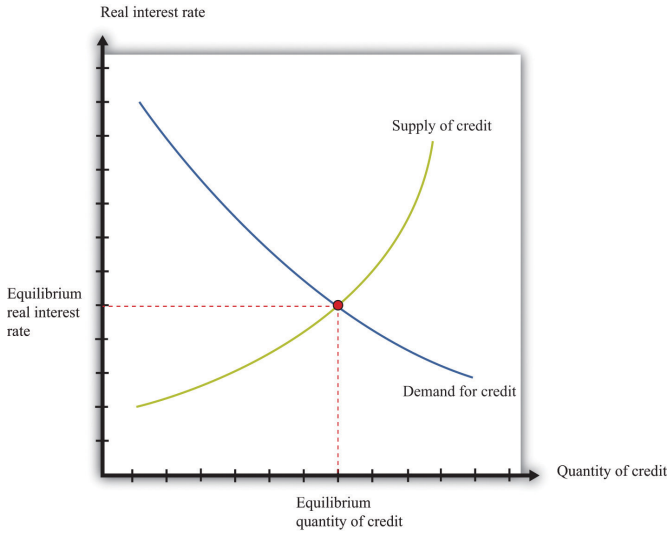


Figure 26.2.2: Credit Market Equilibrium. The credit market equilibrium occurs at a quantity of credit extended (loans) and a real interest rate where the quantity supplied is equal to the quantity demanded.

Toolkit: Section 31.3, Section 31.24, and Section 31.8

You can review the labor market and the credit market, together with the underlying demand and supply curves, in the toolkit. You can also review how to correct for inflation.

The classical dichotomy has a key implication that we can study through a **comparative statics** exercise. Recall that in a comparative statics exercise we examine how the equilibrium prices and output change when something else, outside of the market, changes. Here we ask: what happens to real GDP and the long-run price level when the money supply changes? To find the answer, we begin with the quantity equation:

$$money\ supply \times velocity\ of\ money = price\ level \times real\ GDP.$$

Previously we discussed this equation as an *identity*—something that must be true by the definition of the variables. Now we turn it into a theory. To do so, we make the *assumption* that the velocity of money is fixed. This means that any increase in the money supply must increase the left-hand side of the quantity equation. When the left-hand side of the quantity equation increases, then, for any given level of output, the price level is higher (equivalently, for any given value of the price level, the level of real GDP is higher).

What then changes when we change the money supply: output, prices, or both? Based on the classical dichotomy, we know the answer. Real variables, such as real GDP and the velocity of money, stay constant. A change in a nominal variable—the money supply—leads to changes in other nominal variables, but real variables do not change. The fact that changes in the money supply have no long-run effect on real variables is called the **long-run neutrality of money**.

Toolkit: Section 31.16

You can find more details on how to conduct comparative static exercises in the toolkit.

We continue to assume that the velocity of money is a constant.In fact, the velocity of money might also grow over time as a result of developments in the financial sector. Saying that the velocity of money is constant is the same as saying that its growth rate is zero. Using this fact and rearranging the equation, we discover that the long-run inflation rate depends on the difference between how rapidly the money supply grows and how rapidly output grows:

$$inflation\ rate = growth\ rate\ of\ money\ supply - growth\ rate\ of\ output$$

$$\text{inflation rate} = \text{growth rate of money supply} - \text{growth rate of output.}$$

The long-run growth rate of output does not depend on the growth rate of the money supply or the inflation rate. We know this because long-run output growth depends on the accumulation of capital, labor, and technology. From our discussion of labor and credit markets, equilibrium in these markets is described by real variables. Equilibrium in the labor market depends on the real wage and not on any nominal variables. Likewise, equilibrium in the credit market tells us that the level of investment does not depend on nominal variables. Since the capital stock in any period is just the accumulation of past investment, we know that the stock of capital is also independent of nominal variables.

Therefore there is a direct link between the money supply growth rate and the inflation rate. The classical dichotomy teaches us that changes in the money supply do not affect the velocity of money or the level of output. It follows that any changes in the growth rate of the money supply will show up one-for-one as changes in the inflation rate. We say more about monetary policy later, but notice that there are immediate implications for the conduct of monetary policy:

- In a growing economy, there are more transactions taking place, so there is typically a need for more money to facilitate those transactions. Thus some growth of the money supply is probably desirable to match the increased income.
- If the monetary authorities want a stable price level—zero inflation—in the long run, then they should try to set the growth rate of the money supply equal to the (long-run) growth rate of output.
- If the monetary authorities want a low level of inflation in the long run, then they should aim to have the money supply grow just a little bit faster than the growth rate of output.

Keep in mind that this is just a theory. The quantity equation holds as an identity. But the assumption of constant velocity and the statement that long-run output growth is independent of money growth are assertions based on a body of theory. We now look at how well this theory fits the facts.

Key Takeaways

1. The quantity theory of money states that the supply of money times the velocity of money equals nominal GDP.
2. According to the classical dichotomy, real variables, such as real GDP, consumption, investment, the real wage, and the real interest rate, are determined independently of nominal variables, such as the money supply.
3. Using the quantity equation along with the classical dichotomy, in the long run the inflation rate equals the rate of money growth minus the growth rate of output.

Exercises

1. Is the real wage a nominal variable? What about the money supply?
2. If velocity of money decreases by 2 percent and the money supply does not grow, can you say what will happen to nominal GDP growth? Can you say what will happen to inflation?

This page titled 26.2: The Quantity Theory of Money is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

26.3: Facts about Inflation and Money Growth 26.3: Facts about Inflation and Money GrowthEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What does it mean to say that “inflation is always and everywhere a monetary phenomenon”?
2. What do we know about inflation and money growth in the United States?
3. What happened during past and recent hyperinflations?

According to the quantity equation, the inflation rate and the rate of money growth are closely linked. As the famous economist Milton Friedman said, “Inflation is always and everywhere a monetary phenomenon.”This quote comes from Milton Friedman and Anna Schwartz, *A Monetary History of the United States, 1867–1960* (Princeton, NJ: Princeton University Press, 1963). By this he meant that inflation could always ultimately be traced to “excessive” money growth. Keep in mind that we are talking about the long run here. Over shorter periods of time, changes in the money supply affect the level of real economic activity and have correspondingly less effect on the inflation rate.

Inflation and Money Growth in the United States

Figure 26.3.1 and Figure 26.3.2 show the relationship between inflation and money growth for the United States. For this discussion, money growth is measured as *M1*. The rate of money growth is on the horizontal axis, and the annual inflation rate is on the vertical axis.

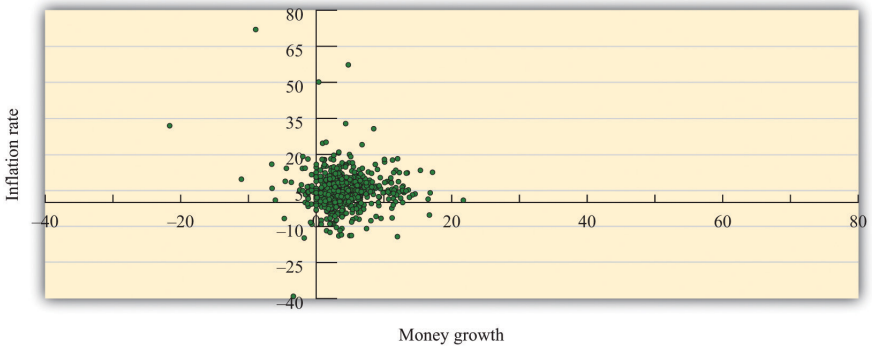


Figure 26.3.1: Inflation and Money Growth in the Short Run

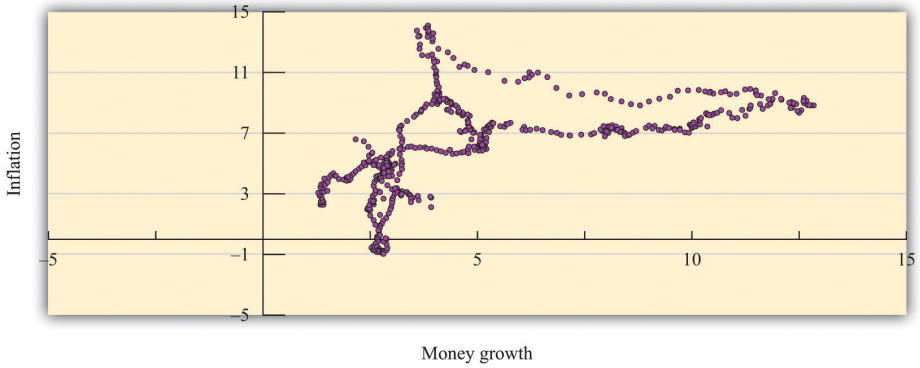


Figure 26.3.2: Inflation and Money Growth in the Long Run

The two figures differ in the time horizon used to compute the growth rates. In Figure 26.3.1, month-to-month changes in money and prices are used to calculate annual growth rates. If you listen to a radio report or read the newspaper about inflation, typically you will first be told about the monthly Consumer Price Index (CPI) and then be given an annual inflation rate. The annual growth rate is the amount by which the variable would increase if the monthly growth rate persisted for a year. The conversion is simply to take the monthly percentage change and convert it into an annual percentage change by multiplying by 12. So if the CPI increased from 112 to 118 over the past month, then the change *for the month* would be calculated as follows:

If prices increased at this rate each month at this same rate, then prices would increase by 12 × 5.36 percent = 64.32 percent over the year. The data for Figure 26.3.1 start in January 1959 and end in December 2010. So the first observation is the annual percentage change between January and February 1959.

Figure 26.3.2 examines annual growth rates based on observing the money supply and the price level at five-year intervals. The first observation is the annual growth rate for the period starting in January 1959 and ending in January 1964. The annual growth rates for a five-year period are computed for each month starting in January 1964. Here, instead of multiplying a monthly growth rate by 12 to get an annual rate, we divide a five-year rate by 5 to get an annual rate. The point of examining growth rates over longer periods of time goes back to the idea that we are investigating the relationship between prices and the money supply over long periods of time.

Comparing these two figures, you can see that the relationship between money growth and inflation is much tighter when we examine five-year periods, as in Figure 26.3.2, rather than the monthly changes in Figure 26.3.1. This is consistent with the view that the relationship between money growth and inflation is a long-term relationship, not a short-term relationship.

In the monthly data, the link between money growth and inflation is relatively weak. The correlation, a measure of how closely two variables move together, is only 0.20 in the monthly data. In contrast, for the annual growth rates computed by looking over a five-year period, the correlation is about 0.65, indicating that money growth and inflation move more closely together over longer periods of time.

Toolkit: Section 31.23

You can review the meaning and measurement of correlation in the toolkit.

Money Growth and Inflation in Other Countries

In the United States, money growth and inflation rates are relatively moderate. Looking back at Figure 26.3.1, we see that the highest inflation rate in the past half-century was about 15 percent, in 1980. Some other countries have had a very different experience.

Figure 26.3.3 shows data on money growth and inflation from 110 countries.See George McCandless Jr. and Warren Weber, “Some Monetary Facts,” *Federal Reserve Bank of Minneapolis Quarterly Review* 19, no. 3 (Summer 1995): 2–11. The article provides a complete description of the data and the countries. On the vertical axis of the figure is the inflation rate, measured as the annual rate of change of the CPI. On the horizontal axis is the rate of growth of the money supply. So a point in the figure represents a single country and shows that country’s combination of inflation and money growth. The sample period used is 1960–1990, meaning that each point is an average over a three-decade period.

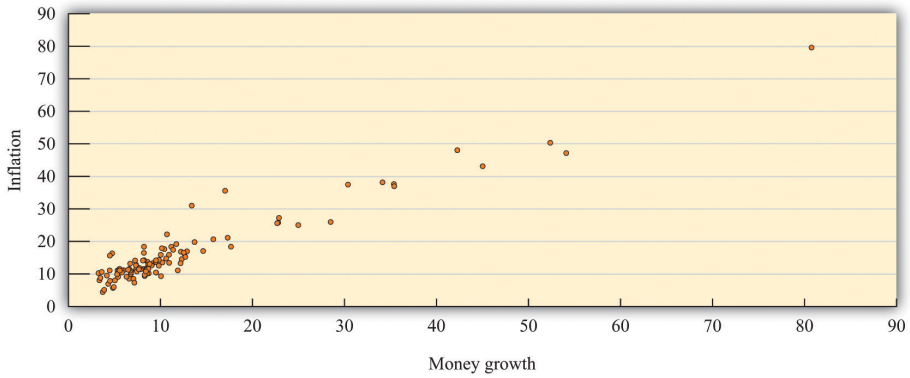


Figure 26.3.3: Inflation and Money Growth in Different Countries

Figure 26.3.3 clearly indicates that countries with high money growth are the countries that experience high inflation. If you were to draw a line through the points that came as close as possible to them, that line would have a positive slope. McCandless and Weber conclude as follows: “In the long run, there is a high (almost unity) correlation between the rate of growth of the money supply and the inflation rate. This holds across three definitions of money and across the full sample of countries and two subsamples.” George McCandless Jr. and Warren Weber, “Some Monetary Facts,” *Federal Reserve Bank of Minneapolis Quarterly Review* 19, no. 3 (Summer 1995): 2–11.

Big Inflations

Most of the countries in Figure 26.3.3 have inflation and money growth that are less than 20 percent. There are some outliers, however. For example, there is one country with inflation and money growth at 80 percent annually over the sample. This country is Argentina; we return to it later. There have been episodes in history where the rates of inflation were so large that they are difficult to comprehend.

Germany, 1922–24

Table 26.3.1 "Prices in Germany" contains data for Germany in the early 1920s.The data come from Thomas Sargent, “The Ends of Four Big Inflations,” in *Inflation: Causes and Effects*, ed. Robert Hall (Cambridge, MA: National Bureau of Economic Research, 1982). The data in this case show the levels of wholesale prices because reliable consumer price indices were not available. The second column is a measure of prices for each month, from January 1922 to June 1924. The third column computes the annual inflation rate by multiplying the monthly inflation rate by 12. The final column indicates the amount of time in days it would take for prices to double at the annual inflation rate indicated in the third column. (When the number in the last column is negative, it tells you how long it would take the price level to halve.)

Table 26.3.1: Prices in Germany

Month and Year	Price Level	Annual Growth Rate (%)	Doubling Time in Days
January 1922	3,670	60.3	419
February 1922	4,100	133.0	190
March 1922	5,430	337.1	75
April 1922	6,360	189.7	133
May 1922	6,460	18.7	1351
June 1922	7,030	101.5	249
July 1922	10,160	441.9	57
August 1922	19,200	763.7	33
September 1922	28,700	482.4	52
October 1922	56,600	814.9	31
November 1922	115,100	851.8	30
December 1922	147,480	297.5	85
January 1923	278,500	762.9	33
February 1923	588,500	897.8	28
March 1923	488,800	−222.7	−113.6
April 1923	521,200	77.0	328
May 1923	817,000	539.4	47
June 1923	1,938,500	1036.8	24
July 1923	7,478,700	1620.2	16
August 1923	94,404,100	3042.6	8
September 1923	2,394,889,300	3880.2	6
October 1923	709,480,000,000	6829.4	4
November 1923	72,570,000,000,000	5553.3	5
December 1923	126,160,000,000,000	663.6	38

Month and Year	Price Level	Annual Growth Rate (%)	Doubling Time in Days
January 1924	117,320,000,000,000	−87.2	−290
February 1924	116,170,000,000,000	−11.8	−2140
March 1924	120,670,000,000,000	45.6	555
April 1924	124,050,000,000,000	33.2	763
May 1924	122,460,000,000,000	−15.5	−1634
June 1924	115,900,000,000,000	−66.1	−383

From the table, you can get a vivid sense of the pace of prices simply by counting the number of digits used to describe the price level. At the height of the inflation in October 1923, the annual inflation rate was over 6,800 percent. It is hard to make sense of a number like this, which is why we include the fourth column: at this inflation rate, prices double every 3 to 4 days. Rapid inflation of this kind is called **hyperinflation**.

Where does hyperinflation come from? The quantity theory tells us that the rapid price increases must be related to growth in the money supply, a reduction in output growth, or rapid growth in the velocity of money. Drawing on the quote from Milton Friedman, it is natural to first examine the growth rate of the money supply. Figure 26.3.4 shows the money growth and inflation rates for Germany during this period. The graph clearly shows that as prices were exploding in Germany, so too was the money supply. In 1922, prices increased 93 percent, and the money stock grew at 52 percent. In the following year, the average inflation rate was up to 433 percent, and the money supply grew at almost 300 percent.These are calculated as January to January growth rates.

In October 1923, when the inflation rate peaked at over 6,800 percent, the money supply grew at nearly 6,000 percent on an annual basis. According to economist Thomas Sargent, 99 percent of the outstanding bank notes had been put in circulation during the previous month. At that point, both prices and the money supply were doubling in a matter of days. Thus the escalating prices were matched by enormous increases in the money supply.

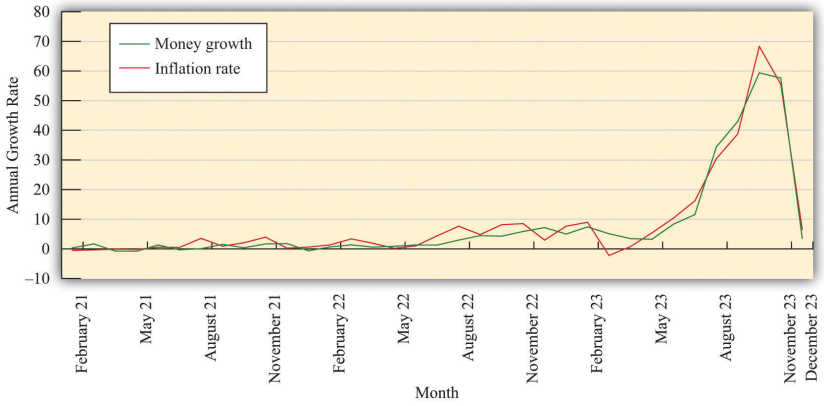


Figure 26.3.4: Money Growth and Inflation in Germany

At first glance, the German data seem to confirm the idea that large inflation rates are driven by large money growth rates. On closer examination, though, we notice that the inflation rates were greater than the growth rate of the money supply. Yet we said earlier that

$$inflation\ rate = growth\ rate\ of\ money\ supply + growth\ rate\ of\ velocity - growth\ rate\ of\ output.$$

It follows that the velocity of money must have been increasing or output must have been decreasing.

It is plausible, indeed likely, that the velocity of money will increase during a period of very high inflation. If you know that the cash in your pocket will lose its value from one hour to the next, then you want to get rid of it quickly. During the German hyperinflation, anyone with cash wanted to exchange it as quickly as possible for goods and services. Thus money changed hands more and more rapidly: in other words, the velocity of money increased.

Money had ceased to perform one of its key functions. It was no longer a store of value. Even if people were still using money as a medium of exchange, they could no longer rely on money to keep its value. A monetary system is a fragile institution: its success depends on everyone believing in it. See Chapter 24 for more discussion. People are willing to accept money because they think that others will, in turn, be willing to accept it from them. During a hyperinflation, this system breaks down. People are reluctant to accept money because they know that others will not want to accept it from them.

Rapid inflation is also disruptive to the general functioning of the economy. People have to devote much more time and energy to managing their cash. People insist on being paid more frequently and abandon work to shop as soon as they are paid. Furthermore, as discussed later, inflation acts as a tax on work. So higher inflation means a higher tax and thus a reduction in employment and output. Overall, output does tend to decrease during hyperinflation, increasing the inflation rate still further. For Germany, real output decreased by 46 percent in 1923 during the height of the hyperinflation. In contrast, 1924 was a good year for the economy, with real output growing at 35 percent.

So while rapid money growth sets hyperinflation in motion, hyperinflation then becomes self-fueling, powered by increases in the velocity of money and—to a minor extent—decreases in the growth rate of output. In the end, the system can collapse completely, with people no longer being willing to accept money at all. In Germany, this is what eventually happened. There are many anecdotes surrounding the German hyperinflation: children using piles of money as building blocks, households using money as wallpaper, and so forth. Figure 26.3.5 shows money being used in a furnace to heat a home.



Figure 26.3.5: The Use of Money in a Hyperinflatio

In December 1923, the hyperinflation came to an end. Look again at Table 26.3.1 "Prices in Germany". Prices in that month had increased to around a billion times greater than they had been two years previously. But from then the price level stayed roughly steady. In fact, it decreased for the next two months, then fluctuated somewhat. The price level in June 1924 was lower than it was at the start of the year. There is thus a new mystery to solve: what happened to bring the inflation to an end? We return to this question shortly.

Zimbabwe

We discussed the example of Germany in some detail because it is one of the most dramatic hyperinflations ever. But hyperinflations are not simply the stuff of economic history. Indeed, from around 2003 to 2009, the African country of Zimbabwe was embroiled in severe inflation. In 2008, prices were doubling on an almost daily basis. Banknotes were issued in denominations of 100,000,000,000,000 Zimbabwe dollars. “Zimbabwe Hyperinflation ‘Will Set World Record within Six Weeks,’” *The Telegraph*, November 13, 2008, accessed August 22, 2011, http://www.telegraph.co.uk/news/worldnews/africaandindianocean/zimbabwe/3453540/Zimbabwe-hyperinflation-will-set-world-record-within-six-weeks.html, accessed August 22, 2011; “A Worthless Currency,” *The Economist*, July 17, 2008, accessed August 22, 2011, http://www.economist.com/node/11751346?story_id=E1_TTSVTPQG, accessed August 22, 2011.

Table 26.3.2 "The Start of the Hyperinflation in Zimbabwe" presents some basic economic facts about Zimbabwe as it entered the hyperinflation; the data come from an International Monetary Fund country report (http://www.imf.org/external/pubs/ft/scr/2005/cr05359.pdf). Looking at these numbers, one is immediately struck by the severity of the decline in economic activity: real gross domestic product (GDP) decreased every year since 2000, including an 11 percent

decline in 2003. At the same time, the country experienced rapid inflation, reaching nearly 600 percent in 2003. As indicated by the third row of the table, the money supply (measured as M1) grew rapidly in 2003 and 2004, fueling the inflation.

Table 26.3.2: The Start of the Hyperinflation in Zimbabwe

Variable	2000	2001	2002	2003	2004
real GDP growth (% change, market prices)	−7.3	−2.7	−4.4	−10.9	−3.5
consumer prices (% change)	55.2	112.1	198.9	598.7	132.7
money supply (billions)	52.6	128.5	348.5	2,059.3	6,867.0

Stories from Zimbabwe resemble the experiences from the 1920s in Germany. The British Broadcasting Company presented some interviews about life during this period of rampant inflation.

THE STUDENT When I go to withdraw my money, I have to wait around 30 minutes because there are so many people waiting.

It's so difficult.

Maybe you want 10 million but they only give you 2.8, because there is not enough at the bank.

THE LECTURER Children in Harare play in uncollected rubbish. Hyperinflation has meant an end to rubbish collections. It's a very strange environment.

There are a lot of pay rises, but they are meaningless.

They are always eroded the minute they give us the pay rise.

Also, considering we have so much to pay—we have parents in the countryside, and we have families—it doesn't work.

People are willing to lend money, but they are not willing to lend it for nothing. It's usually at a rate of 90 or 100 percent.

Sometimes these are your relatives or people you work with, taking advantage of this.

People are cannibalizing each other.

THE MOTHER Because my income hasn't risen as much as the prices in the shops, we have had to adjust quite a bit.

The things that we buy—the groceries at home, the things we get for our two children—we have to buy immediately, as soon as we get the money.

We know that if we wait a bit, the prices are going to go up again. If we wait another week, we will not be able to afford anything.

People are taking the money out in suitcases or carrier bags. "Zimbabwe: Living with Hyperinflation," BBC News, January 31, 2006, accessed July 21, 2011, http://news.bbc.co.uk/2/hi/africa/4665854.stm.

Zimbabwe's citizens increasingly turned to other currencies to conduct transactions, even though the Zimbabwe dollar was officially the only legal tender in the country. The Zimbabwe hyperinflation eventually ended in January 2009, when the Finance Minister officially permitted citizens to use other currencies in places of the Zimbabwe dollar. "Zimbabwe Abandons Its Currency," BBC News, January 29, 2009, accessed August 22, 2011, http://news.bbc.co.uk/2/hi/7859033.stm.

Key Takeaways

1. The quote by Milton Friedman that "inflation is always and everywhere a monetary phenomenon" points out the connection between money growth and price growth (inflation). From this perspective, the source of inflation is money growth.
2. Over long periods of time, inflation and money growth are closely linked in the United States.
3. The hyperinflations in many countries, such as Germany and Zimbabwe, were times of rapid growth in prices stemming from rapid expansions of the money supply and subsequently fueled by increases in the velocity of money.

Exercises

1. What happens to the velocity of money during a hyperinflation?
2. What is the difference between a monthly inflation rate and an annual inflation rate?

This page titled 26.3: Facts about Inflation and Money Growth is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

26.4: The Causes of Inflation 26.4: The Causes of InflationEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What is the inflation tax?
2. How is inflation caused by the central bank's commitment problem?
3. What happens if there are multiple regions (states or countries) independently choosing how much money to print?

We have argued so far that inflation is caused by excessive money growth, which in turn leads to increases in the velocity of money. But we have also documented that rapid inflations are damaging to the functioning of an economy. There is therefore a deeper question to be asked: why on earth do monetary authorities pursue policies that lead to such disastrous outcomes?

The Inflation Tax

Suppose your country is at war. Wars are expensive. Not only are there soldiers to be paid and kept supplied, but your valuable aircraft and tanks are liable to be destroyed by the enemy while you are in turn throwing costly ammunition and missiles at them. How do governments pay for all these expenses? One thing that the government can do is to tax the population to pay these bills. It may not be feasible to collect enough tax revenue in the time of a war, however. Many governments instead borrow during times of large expenses. This allows the government to spread the tax burdens over time.

So far, taxation and borrowing are the only two possibilities that we have considered. But there is a third possibility: a government can simply print the money it needs. There is a government budget constraint that says the government budget constraint is discussed in Chapter 29.

$$deficit = change\ in\ government\ debt + change\ in\ money\ supply.$$

The left side of this equation is the deficit of the government. The deficit is the difference between government outlays and government receipts. The right side of this equation describes how the government finances its deficit. This equation says that the government can finance its deficit by issuing either new government bonds or new money.

Toolkit: Section 31.33

You can review the details of the government budget constraint in the toolkit.

There is a puzzle here. Money is just a piece of paper with writing on it. The government can print it at will. Yet the government can take these pieces of paper and exchange them for goods and services of real value. It can pay soldiers, or nurses, or construction workers who are building roads. It can print money, hand it over to Airbus or Boeing, and get a new airplane. So who is really paying in this case?

We already know everything we need to know to figure out the answer. When the government prints more money, prices will eventually increase. This comes directly from the quantity equation once we remember that real variables are independent of the money supply in the long run. In the long run, the extra money will just result in higher prices and no additional output. *And increased prices mean that existing money becomes less valuable*. If the price level increases by 10 percent, existing dollar bills are worth 10 percent less than they were; they will buy (roughly) 10 percent less in terms of goods and services. Inflation is exactly like a tax on the money that people currently hold in their wallets and pocketbooks. Indeed, we say that there is an **inflation tax** when the government prints money to finance its deficit.

Examine the government budget constraint again. If we write out the deficit in full, the equation says

$$government\ purchases + transfers - tax\ receipts = change\ in\ government\ debt + change\ in\ money\ supply.$$

Suppose that government purchases increase, say due to a war, by $100 billion. This equation tells us that, to finance this expense, the government could

- increase taxes now by $100 billion,
- increase taxes now by less than $100 billion and sell some government debt,
- increase taxes now by less than $100 billion and print some money.

In some sense, these are all versions of the same thing: to finance the spending of $100 billion, the government will have to increase taxes. Those taxes may be paid now, they may be paid later (when the government repays the debt), or they may be paid through the inflation tax. The government must decide how to best increase taxes to finance the extra spending, and the inflation tax is one option available to the government.

Commitment

It is hard to imagine that a government acting in the interests of its citizens would choose to bring about hyperinflation. Why do governments apply such misguided policies? The leading explanations all fall under the heading of a "weak" central bank. A weak central bank is unable to pursue its normal goal of price stability and instead becomes a tool of other interests, such as the fiscal authorities.

A government entity, such as a central bank or the treasury, suffers from a **commitment problem** when it is not able to make credible promises to pursue certain actions. Suppose a central bank wishes to pursue a strategy of stabilizing prices. If the economy is in a deep recession, the central bank might instead come under pressure to reduce interest rates. Reductions in interest rates require the central bank to increase the money supply and ultimately create inflation, yet if it could commit to a policy, the central bank might prefer to focus on inflation and ignore the recession. Let us see how these types of commitment problems work through some examples.

Increasing Output

The level of potential output in an economy is not necessarily the *ideal* level of output. Even when the economy is at potential, there is some unemployment and some spare capacity. The monetary authority therefore might have a target level of output that is above potential output. Suppose (for simplicity) that its target level of inflation is zero. To understand what will happen, we use our model of **price adjustment**:

$$inflation\ rate = autonomous\ inflation - inflation\ sensitivity \times output\ gap.$$

Toolkit: Section 31.31

You can review the details of price adjustment in the toolkit.

To begin with, suppose that everyone in the economy believes that there will be zero inflation, so autonomous inflation is zero. Were output equal to potential output (so the output gap is zero), then actual inflation would also be zero. This situation is summarized in Figure 26.4.1. However, if the Fed follows a **Taylor rule**, it will react to the fact that output is below its target by reducing real interest rates with the aim of increasing spending and output. The price adjustment equation then tells us that there will be positive inflation. This outcome is also shown in Figure 26.4.1 as the combination of the target level of output and a positive inflation rate.

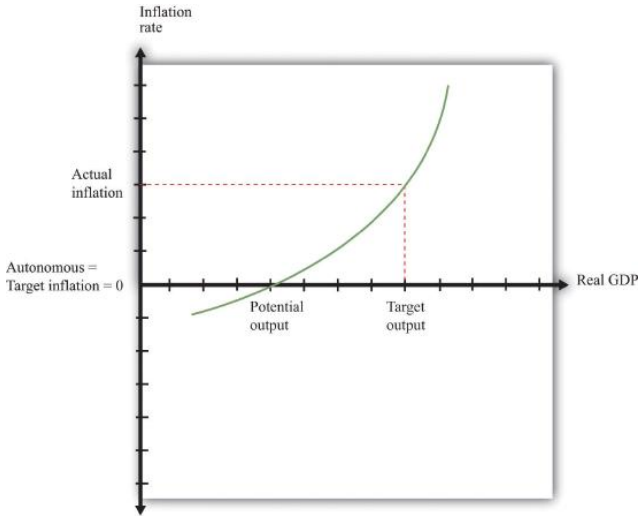


Figure 26.4.1: The Gains to Inflation. If target inflation = autonomous inflation = 0, but target output is above potential output, then the Fed will reduce the real interest rate and create more output to meet its target output. This will create inflation.

This is not the end of the story. Everyone in the economy is predicting zero inflation, yet the Fed is using its monetary policy to increase output and create positive inflation. Over time, people will notice that their expectations are wrong and will start to expect positive inflation instead. This results in an increase in autonomous inflation and a shift in the relationship between inflation and output.

At that point the Fed will have an incentive to create still more inflation to pursue its goal of output above potential. But additional inflation is costly to the Fed because it is now moving away from its target of zero inflation. Eventually inflation will be so high that the Fed no longer wants to create more inflation to increase output. The economy will end up with a positive inflation rate, where expectations of inflation are equal to actual inflation and no one is fooled. In the end, the Fed incurs an inflation rate above its target, yet it does not succeed in creating output above potential.

The final outcome involves costly inflation, but output remains at potential. Given that it cannot actually keep output above potential, the Fed would prefer zero inflation, yet it lacks the ability to *commit* to a zero-inflation policy. If the inflation rate is zero, the Fed has an incentive to create positive inflation.

The Politics of Fiscal Policy

The government budget constraint tells us that are three ways to fund spending: taxes today, the inflation tax, or debt (which means taxes at some future date). A government that has the best interests of its citizens at heart will decide on the best mix of these three. Optimists may believe that this is what governments try to do. Cynics might hold a very different view. Suppose—just suppose—that the leader of a government is more concerned with reelection than with sound economic policy and believes that her chances of reelection will be increased if she pledges not to increase taxes now or in the future. If this promise is credible, then the government budget constraint tells us that any increases in spending must be financed by money growth.

The monetary authority again has no power to commit to avoid inflation. The fiscal side of the government has set the level of spending and decided, based on the wishes of the political leaders, to have low taxes. Faced with this fiscal package, the monetary authority has no choice: it must print money to finance the government budget constraint. This story relies on the belief that individuals in the economy do not understand that the government, by its fiscal actions, is causing inflation and thus imposing a kind of tax.

A more extreme example arises when the government's expenditures are so great that it simply cannot finance them with current taxation. This can occur in poorer economies where the tax base is low and the mechanisms for collecting taxes are often imperfect. Moreover, a government can finance its deficit through borrowing only if the public is willing to purchase government bonds. If a government is in fiscal trouble—if its tax and spending policies appear to the public to be unsustainable—it will have great difficulty persuading the public and the international investment community to buy government debt. Investors will demand a very high interest rate (including a **risk premium**) to cover the possibility that the government may default on its debt. Interest rates on debt will increase.

At this point a government may find that the only option available to it is to finance its deficit through the printing of money. After all, no government wants to be in the position of being unable to pay its soldiers. The leaders of a country in such a position will decide to run the printing presses instead. The end result is inflation and, if the process gets completely out of control, hyperinflation. But from the government's point of view, at least it buys it some time. Thus although moderate inflations are caused by poor monetary policy, hyperinflations are almost always originally caused by unsustainable fiscal policies.

Regional Monetary Policy

Figure 26.4.2 shows the price level in Argentina from 1988 to 2005. Argentina experienced hyperinflation in the early 1990s. Prices were then stable for about a decade and then increased again in the early years of the 21st century. In Argentina, different regional governments have significant power over the decisions of the central government. (It is as if a state government in the United States could appeal for funds directly from Washington.) These transfers from the central government in turn must be funded either from tax revenues or by printing money. If a region is sufficiently powerful relative to the central government, then it is as if the regional government has the power to print currency.

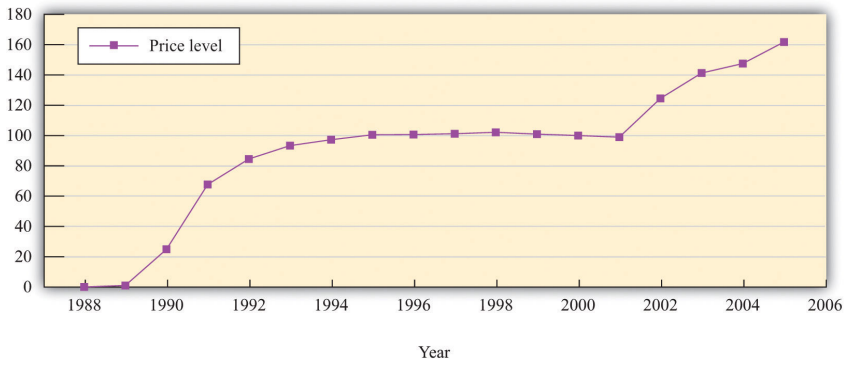


Figure 26.4.2: The Price Level in Argentina. Source: International Monetary Fund World Economic Outlook database

(http://www.imf.org/external/pubs/ft/weo/2010/01/index.htm).

A battle between regional governments can give rise to hyperinflation. Russell Cooper and Hubert Kempf, "Dollarization and the Conquest of Hyperinflation in Divided Societies," *Federal Reserve Bank of Minneapolis Quarterly Review*, Summer 2001, accessed July 21, 2011, http://www.minneapolisfed.org/research/QR/QR2531.pdf. To simplify the issue, suppose that each region in Argentina has its own printing press. Each region can then independently undertake monetary policy by printing Argentine pesos and using those pesos to fund projects within their regions. The inflation tax is very tempting in these circumstances: a regional government can in effect tax people in other regions to help pay for its own projects. Why? Because printing money results in an inflation tax on *everyone who has pesos*. If money is printed in one region, some of the inflation tax will be paid by people in other regions who have pesos.

To be concrete, imagine you are a politician in Buenos Aires who wants to raise 100 million pesos for a project in that city. You could levy an income tax on citizens of your area. Alternatively, you could print 100 million pesos. If you impose the income tax, your own citizens must pay it all. If you impose the inflation tax, people living in other regions of Argentina pay some of the tax. Your constituents get the benefit, but others bear a large part of the costs. Acting in the interests of your constituents, you print the pesos. Of course, this story is true not only for you but also for the leaders in all regions. In the end, there is excessive money growth in the economy as a whole and high inflation. The monetary authorities are weak because no single authority controls the overall money supply.

The situation we have described is sometimes called a **prisoners' dilemma**. In a prisoners' dilemma, the actions of one person imposes costs on others, and the behavior that is best for each individual decision maker (in this case, all the regional monetary authorities) is not best for the country as a whole.

Toolkit: Section 31.18

If you are interested in more detail on the prisoners' dilemma game, you can review it in the toolkit.

Reducing Distortions from Taxes

Suppose a government faces a large expense today and can tax labor income to pay for it. One option is to increase income taxes today by a lot to finance this expense. This would cause a reduction in labor supply and thus in employment and real gross domestic product (GDP). This *distortion* in labor supply is an economic cost of the tax. Alternatively, the government could increase taxes a little bit today and a little bit in future years. This spreads out the tax over many years and leads to less distortion. The government budget constraint tells us that the government can spread out taxes by borrowing new and levying taxes later to pay off the debt.

If the government had access to a nondistortionary tax instead of income taxes, it would be better to use that tax instead. For a tax not to be distortionary, it must be the case that economic decisions (how much to buy and sell, how much labor to supply, etc.) do not change when the tax changes. At first glance, it seems that the inflation tax might fit the bill. Remember the inflation tax makes people's *existing* stocks of money worth less in real terms. People have already decided how much money to hold. So if the government levies an inflation tax, it is not distortionary; people have already made their decisions on how much money they want to own.

But there is a danger here. Our argument rests on the idea that the decisions about which assets to hold have already been made by households. The inflation tax might be nondistortionary the first time that the government tried it. But people would rapidly come to anticipate that the government would be likely to use it again. At that point they *would* start changing their decisions about how much money to hold, and the tax would be distortionary after all.

Key Takeaways

1. A government that prints money to finance its deficit is using an inflation tax. Individuals who hold nominal assets such as currency pay the tax.
2. One source of inflation is a commitment problem of a central bank wishing to use inflation to boost output.
3. When there are multiple regions (states or countries) each with the power to print money, then inflation will tend to be higher than it would be if only a single central bank controlled the money supply.

Exercises

1. When you pay the inflation tax, do you have to fill out a form? If not, then how is the inflation tax collected?
2. If a government has a deficit of $400 billion and sells $200 billion in debt, how much must it increase the money supply so that the government budget constraint holds?

This page titled 26.4: The Causes of Inflation is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

26.5: The Costs of Inflation 26.5: The Costs of InflationEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What are the costs of an excessive inflation tax?
2. When does inflation cause a redistribution?

At the beginning of this chapter, we highlighted President Ford's campaign to "Whip Inflation Now." It is clear from that episode that even relatively moderate inflation is perceived as a bad thing. It is even more self-evident that massive inflations, such as those in Germany or Zimbabwe, are highly disruptive. It is hardly surprising that the stated primary objective of most central banks is price stability. All that said, we have not yet really explained exactly why inflation is costly.

An Excessive Inflation Tax

Inflation, used as one tax among many, may be an efficient way of raising some of a government's revenues. The effects of the inflation tax are like the effects of any tax: people respond by substituting away from the activity being taxed. When the government taxes cigarettes, people smoke less. When the government taxes income, people work less. When the government taxes the money people hold, people hold less money. These changes in behavior are the distortions caused by taxation. People substitute away from holding money in two ways: (1) during moderate inflations, people allocate more of their time to transactions; and (2) during high inflations, people may cease using money altogether.

During high inflations, the real value of money decreases quickly. So if you work and get paid in money, you had better go shopping quickly to make purchases. During hyperinflations, people may literally spend more time trying to get rid of their money than they do earning it in the first place. The same distortion applies, although less dramatically, in times of low to moderate inflation. People respond to inflation by carrying less cash, on average. To do so, they must spend more time standing in line in the bank and at automatic teller machines.

Imagine that ice cream were to be used as money. In a very cold climate, ice cream is just fine as a store of value. In a very hot climate, by contrast, ice cream is a bad store of value. You would probably want to get paid every day, and as soon as you received your ice cream, you would run to the store to buy other goods and services before your money melted. You and everyone else would spend much more time shopping and less time working. Melting ice cream, in this world, is like inflation.

There is a good reason why we do not use ice cream as a medium of exchange. Because it is such a bad store of value, people would quickly abandon it in terms of some other way of trading. During hyperinflations, this is exactly what we see: people substitute away from money completely and instead resort to barter trades. Often, some other commodity, such as cigarettes, starts being generally accepted as an alternative to money. But substitution away from money is costly to the economy. Money facilitates trade. It is generally easier to trade when everyone uses money rather than goods in exchange. When people respond to high inflation by eliminating money from trades, we are observing a distortion from the inflation tax.

Uncertainty and Real Interest Rates

It is the *real* interest rate that ultimately matters for saving and investment decisions. Yet loans are almost invariably quoted in nominal terms: a loan contract gives the borrower some money with a requirement to pay back that money plus interest in the future. The real and the nominal interest rates are linked by the **Fisher equation**:

$$real\ interest\ rate \approx nominal\ interest\ rate - inflation\ rate.$$

To calculate the real interest rate you subtract the inflation rate from the nominal interest rate. So, for example, if the annual interest rate on a car loan is 12 percent and the current inflation rate is 4 percent, then the real interest rate on the car loan is 8 percent.

Toolkit: Section 31.25

You can review the derivation and uses of the Fisher equation in the toolkit.

The Fisher equation glosses over an important point, however. Suppose you are thinking of taking out a loan this year, allowing you to borrow money now for repayment next year. The inflation rate that matters for this loan is the inflation between this year

and next. *At the time you sign the contract, you do not know what the inflation rate will be.* You must make your decision about the loan without knowing for sure what the real interest rate will be. You have to make a guess:

$$expected\ real\ interest\ rate \approx nominal\ interest\ rate - expected\ inflation\ rate.$$

Thus when a loan contract is signed, it is based on expectations of what will happen to prices in the future. If a borrower and lender would like to agree on a loan at a 4 percent real interest rate, but both expect 2 percent inflation, then they will agree on a 6 percent nominal interest rate.

What happens if the inflation rate turns out to be different from what the borrower and lender expected? Suppose the actual inflation rate turns out to be 4 percent. This means that the actual real interest rate, from the Fisher equation, is only 2 percent. This is good news for the borrower: he gets a loan at a lower rate than he expected. But it is bad news for the lender: she is repaid at a lower rate than she expected. The opposite is true if the inflation rate is lower than expected. Suppose the actual inflation rate is only 1 percent. Then the real interest rate is higher than anticipated—5 percent instead of 4 percent—which benefits the lender but is costly to the borrower.

Any divergence between actual and expected inflation therefore leads to a *redistribution*, either from the borrower to the lender or from the lender to the borrower. When inflation is higher than expected, the borrower is better off, and the lender is worse off. The opposite effects occur if inflation is lower than expected: the borrower loses, and the lender wins.

The possibility that the inflation rate will turn out to be unexpectedly high or unexpectedly low means that there is uncertainty whenever people sign loan contracts. A fixed nominal interest rate on a loan exposes both the borrower and the lender to the risk of inflation uncertainty. Uncertainty can prevent beneficial trades from taking place. Imagine that you were thinking of buying a used car, but you had to decide to buy without knowing whether the price was going to be $1,500 or $2,000. You might well decide not to buy in the face of this uncertainty. Similarly, people might sometimes decide not to sign loan contracts that would actually be beneficial to them.

The borrower and the lender could always change the form of their contract. Contracts do not *have* to specify nominal interest rates, and not all of them do. Some loans have interest rates that change with the actual inflation rate. In this way, borrowers and lenders can protect themselves from unexpected inflation. However, such contracts are unusual in practice and are most often seen in countries experiencing high and uncertain inflation. What should we conclude from the fact that loan contracts are rarely protected against inflation? Presumably one of two things is true: either such contracts are expensive to write or the benefit of these contracts is actually small.

Unexpected inflation can also have redistributive effects with other types of contracts. Labor contracts are an example. Although the worker and the firm ultimately care about real wages, most labor contracts are written in terms of nominal wages. That is, most labor arrangements are not indexed and thus leave the parties open to the effects of unanticipated inflation. So, for example, if inflation is higher than anticipated, then the real wage earned by the worker is lower than expected, which is a benefit to the firm.

Economies do respond to inflation, partly through the way in which people write contracts. In countries with high and volatile inflation, labor and other contracts generally provide some form of protection against inflation through indexation. For example, if you agree to a job that pays you $10 an hour this year, the nominal wage rate next year will change depending on inflation. If, for example, inflation was 20 percent this year, then under an indexed contract your nominal wage would automatically increase by 20 percent to $12. Under full indexation, the real wage you are paid is constant.

Key Takeaways

1. Inflation can distort choices, such as the holding of money. A small amount of inflation, so that it is one tax among many, makes economic sense, but high inflation leads to significant distortion in the economy.
2. Expected inflation is reflected in the terms of loan agreements. Unexpected inflation leads to a lower real interest rate and thus a redistribution from the lender to the borrower.

Exercises

1. If the inflation rate is lower than expected, who gains and who losses?
2. What costs of inflation are highlighted in our discussion of Zimbabwe in 26.3 Section "Zimbabwe"?

This page titled 26.5: The Costs of Inflation is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

26.6: Policy Remedies 26.6: Policy RemediesEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What actions can governments take to prevent excessive inflation?
2. How can hyperinflations be ended?
3. How do governments overcome the commitment problem?

We have already explained that money is a fragile social institution: money has value only because people believe it has value. Hyperinflations illustrate this fragility. Large inflations are impressive but, fortunately, are also relatively rare. In other words, most of the time monetary authorities are somehow able to maintain confidence in the system. To understand how they do so, we begin by looking at how hyperinflations come to an end.

Ending Big Inflations

As noted in 26.4 Section "The Inflation Tax", the rapid inflation in Germany ended abruptly. Although October 1923 was the month with the highest inflation rate, prices actually decreased in early 1924.

How did the hyperinflation end? The answer has to do with the conduct of fiscal policy. On October 15, 1923, a decree created a new currency from the old one. A key element of the decree was limits imposed on the money creation process by the central bank, particularly the provision of credit to the government. According to economist Thomas Sargent, who has studied how hyperinflations end, “This limitation on the amount of credit that could be extended to the government was announced at a time when the government was financing virtually 100 percent of its expenditures by means of note issue.” Thomas Sargent, “The Ends of Four Big Inflations,” in *Inflation: Causes and Effects*, ed. Robert Hall (Cambridge, MA: National Bureau of Economic Research, 1982), 83. Prior to October 1923, government spending was financed by printing money. After the decree, the printing presses were effectively turned off. As a consequence, the government’s budget went into surplus starting in January 1924. The hyperinflation was over once the printing presses were quiet.

Other countries that experienced hyperinflation around this time had similar stories: there was an abrupt end to hyperinflation after a regime change in which fiscal imbalances were restored. In Austria, for example, the inflation ended when the government established an independent central bank and adopted a fiscal policy that did not require financing by the central bank. The reforms in these countries had two effects: (1) the fiscal reforms limited the budget deficits, and (2) the monetary restrictions implied that deficits would not be financed by the printing of money.

A natural question is: what took them so long? Given the damage caused by these periods of hyperinflation, why did the countries not adopt these policies earlier? Part of the explanation may lie in political affiliations of the governments in these countries. Or, perhaps, these governments simply did not appreciate the rather complex links between fiscal and monetary policy.

Delegating Monetary Power to Another Country

Sometimes countries take even more drastic measures to shield monetary policy from political pressures. One is to effectively eliminate the monetary authority and delegate monetary policy to another country. Some small countries do this by simply using another country’s currency. Panama, El Salvador, and Ecuador, for example, have used the US dollar as their currency. Zimbabwe effectively did the same in 2009.

Argentina in the 1990s is an interesting example of a country that went almost—but not quite—that far. Figure 26.6.2 shows the price level in Argentina from 1988 to 2005. There are evidently three distinct periods: very high inflation, zero inflation, and then moderate inflation. From 1988 to 1993, there was substantial inflation. The annual inflation rate was about 343 percent in 1988 and was over 2,300 percent in 1990. But by 1993 it was only 10 percent, and from 1994 to 2001 it was effectively zero. Then, starting in 2002, there was a resurgence of inflation. What happened?

As we explained earlier, Argentina suffered from hyperinflation in the late 1980s as a consequence of a weak monetary authority. In 1991, Argentina adopted a novel monetary system called a **currency board**. Every single peso in circulation was “backed” by a US dollar held by the Central Bank of Argentina. If desired, people had the right to take their pesos to the Central Bank of Argentina and swap them for dollars. Thus Argentina both adopted a **fixed exchange rate** between the peso and the dollar (1 peso

equals $1) and also made that exchange rate credible by always having enough dollars on hand to exchange for the pesos in circulation. For all intents and purposes, Argentina had switched to using US dollars.

Argentina therefore avoided inflation by ceding control of monetary policy to the United States. Since the central bank in the United States controls the quantity of dollars and Argentina linked pesos to dollars, then, everything else the same, the Fed could change the amount of pesos in Argentina, whereas the Central Bank of Argentina could not. The Central Bank of Argentina could resist pressures to inflate by arguing that it did not control the money supply.

Many observers thought at the time that Argentina's currency board would ensure price stability in Argentina. They thought that there would no longer be pressure on the monetary authority from the fiscal side of the economy. This proved to be incorrect. Taking advantage of its healthy economy in the early 1990s, Argentina adopted expansionary fiscal policies. A combination of factors then triggered recession in the country. Unemployment increased to 18 percent. It was not possible to expand fiscal policy much further, and Argentina had given up its control over monetary problem. In the late 1990s and early 2000s, the recession became so severe that the political pressure on the monetary authority was insurmountable. Argentina abandoned its currency board. One result was a resurgence of inflation.

Another variation on the delegation of monetary policy is that adopted by many countries in Europe. They decided to abandon their currencies and their monetary autonomy in favor of a new currency called the euro. Monetary policy is run by the European Central Bank, which is highly independent. Independent central banks are better able to resist political pressure, so countries that had previously had weak central banks saw a significant advantage in adopting the euro.

Abandoning one's currency in favor of a new currency, as occurred throughout Europe, seems like a particularly powerful way for a country to commit to a new monetary regime. It is worth remembering, though, that no monetary system is cast is stone. Just as Argentina's currency board collapsed despite its apparent credibility, so too could a country decide to abandon the euro and reestablish its own currency. Indeed, following fiscal problems in several countries in Europe (most notably Greece, Portugal, and Ireland), there has been some speculation that some countries might eventually choose to do just that.

Independent Monetary Authorities

Hyperinflations arise when the central bank is weak and unable to resist the pressures put on it by others—notably politicians—to use monetary policy for other purposes. Monetary authorities must be able to "just say no." This suggests that monetary authorities will be able to do a better job if they are independent of other branches of government.

Economists have studied the relationship between measures of the independence of a country's central bank and the inflation rate in that country. Economists Alberto Alesina and Lawrence Summers examined both political and economic independence of the monetary authority. By political independence, they meant the process of appointing the leadership of the central bank and the role of government officials in the conduct of monetary policy. By economic independence, they meant the extent to which the monetary authority is under pressure to finance the government's budget deficit.

Figure 26.6.1 displays data from their research. The horizontal axis shows annual inflation, and the vertical axis is their index of central bank independence, with higher numbers indicating a more independent central bank. The data are averaged over the period 1955–1988. Each point in the figure refers to a particular country. Switzerland and Germany both receive very "high" central bank independence ratings of 4 and have relatively low average inflation. Spain, in contrast, has the second lowest measure of central bank independence and has the highest inflation rate in the study.

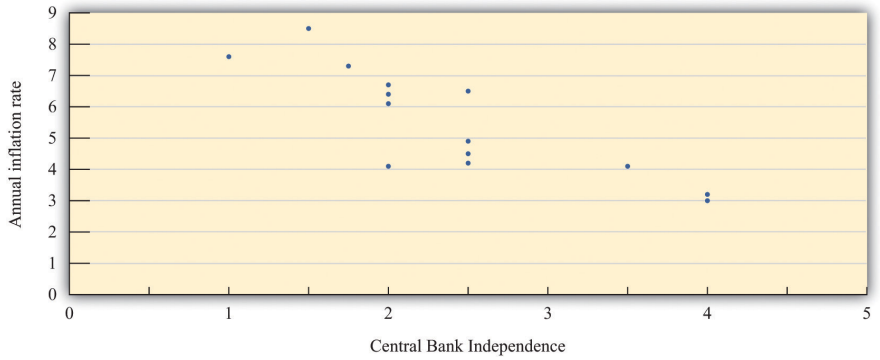


Figure 26.6.1: Central Bank Independence and Inflation

https://socialsci.libretexts.org/@go/page/50922

Since the work of Alesina and Summers (and other economists), more and more countries have become convinced of the virtues of having an independent central bank. For example, when the Labour Party came to power in Britain during the 1990s, one of their first acts was to make the Bank of England more independent. This was particularly striking because the Labour Party is a center-left political party, yet independent central banks tend to be conservative, focusing primarily on inflation and not worrying so much about employment and output.

Events in Argentina also attest to the value of an independent central bank. In 2003, the Congress in Argentina passed an act stating,

The Argentine Central Bank is a National State self-governed institution, whose primary and fundamental mission is to preserve the value of the Argentine currency.

When formulating and implementing the monetary and financial policy, it is not subject to the orders, guidelines or instructions of the National Executive branch of government."BCRA Law," Banco Central de la República Argentina, accessed September 20, 2011, http://www.bcra.gov.ar/institucional/in010000_i.asp.

There are two key elements in this act. First, the stated goal of the Central Bank of Argentina is to preserve the value of the currency. There is no mention of pursuing full employment, just a version of price stability. Second, the central bank is to be independent of the executive branch of the government.

Inflation Targeting

Under a policy regime called **inflation targeting**, some central banks use their tools to set the inflation rate as close as possible to a target. Just as we know that a monetary authority cannot literally control interest rates, we know it cannot literally set the inflation rate either. Rather, it can use the policy tools at its disposal to influence the economy in an attempt to reach the target inflation rate.

In its simplest form, the target is some publically announced inflation rate—say, 3 percent. If the monetary authority thinks the inflation rate is likely to be higher than 3 percent for the year, it adopts contractionary monetary policy to reduce the inflation rate. If it thinks that the inflation rate is likely to be lower than the target, it adopts an expansionary policy. Inflation in this world is relatively predictable.

What should the target be? If, as one might believe from all the discussion in the press and elsewhere, inflation is a pernicious problem, then perhaps the inflation target should be zero. Yet most central banks following this policy adopt targets with positive inflation rates, based on the belief that a little bit of inflation may be useful in the economy. One argument often heard is that deflation (negative inflation) is problematic. From a historical perspective, a prolonged period of deflation in the Unites States occurred during the Great Depression and coincided with a negative output growth. More recently, Japan experienced both slow real gross domestic product (real GDP) growth and some periods of deflation during the 1990s.

Many policymakers have apparently concluded that deflation is to be avoided because it could underlie a depression. An alternative possibility is that deflation is correlated with periods of low economic activity, but it is unclear whether it is the cause or the consequence of a sluggish economy. Whatever the connection between deflation and depression, the prevailing wisdom of the Fed (and other central banks) is to avoid deflation. Given that the central banks cannot always hit their targets precisely, aiming for zero inflation makes deflation more likely than when central banks adopt a target with positive inflation.

In addition, a little inflation may make it easier for relative prices and wages to adjust in an economy. If the demand for beef decreases and the demand for pork increases, then the price of beef should decrease, and the price of pork should increase. Such adjustment is straightforward. Similar logic says that if the demand for accountants decreases and the demand for systems analysts increases, then the wages of accountants should decrease, and the wages of systems analysts should increase. This may be more problematic. People typically respond very negatively if there is an attempt to cut their wages. It may be easier for employers to let inflation do the job of reducing the real wage instead. (This is an argument that makes some economists uncomfortable since it implies irrationality on the part of workers. Still, the psychological resistance to nominal wage cuts appears to be strong.)

A second issue is whether the inflation target should be allowed to vary. Instead of announcing a 3 percent target for all times, the monetary authority might decide that the target rate should depend on the state of the economy. For example, they could have a higher target rate in recession and a lower target rate in booms. This way monetary policy could still be used to help keep the economy at potential output.

Finally, there is the question of "punishment" for missing a target. If the purpose of inflation targeting is to support a particular (moderate) inflation rate, then a central banker missing the target ought to be fined or even terminated, just like a manager of a

store who persistently misses sales targets. Presumably, if a central bank has goals to achieve, it should also have incentives to meet those goals. Central bankers are often called to testify in front of bodies, such as the US Congress, who monitor the progress of the economy relative to particular targets.

Australia is an example of a country that follows an inflation target rule. According to the charter of the Reserve Bank of Australia (http://www.rba.gov.au/monetary-policy/about.html), the goal of Australian monetary policy is to maintain inflation between 2 and 3 percent annually, on average. The central bank does also recognize of the role of monetary policy for stabilization purposes. Thus even though it has a target range for inflation, it also examines the state of the economy when setting monetary policy. Moreover, the phrase *on average* means that the central bank has some leeway in the conduct of policy: they can allow inflation to increase above 3 percent for a short while, provided that they eventually take actions to bring the inflation rate back down.

Key Takeaways

1. Governments can take a variety of actions to prevent excessive inflation. These include the delegation of monetary policy to another central bank, the creation of an independent monetary authority, and constraining monetary policy to focus solely on inflation.
2. Historically hyperinflations ended when the government restored fiscal balance by eliminating deficit spending.
3. A government can overcome a commitment problem by delegating the conduct of monetary policy to a more conservative central bank. This can be achieved through a currency board or through joining a monetary union.

Exercises

1. What is the difference between a fixed exchange rate and a currency board?
2. What is an independent central bank?

This page titled 26.6: Policy Remedies is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

26.7: End-of-Chapter Material 26.7: End-of-Chapter MaterialEdit Update titlecancel

In Conclusion

We have studied some severe and extreme cases of inflation to reveal the sources of rapid price increases. The quantity equation and the data both clearly indicate that inflation is linked to money supply growth. During periods of rapid inflation, the money supply is growing as well. The velocity of money also increases in times of rapid inflation, reflecting the collapse of confidence in the monetary system.

Money supply growth, in turn, comes about because money creation can help finance government budget deficits. Instead of using taxes to finance government spending, governments just print money. This increases prices and thus acts like a tax on those holding money and other nominal assets. Like all taxes, the inflation tax is distortionary. Used in moderation, there is an argument for using this tax along with others. But for some countries in some time periods, the use of the inflation tax has been excessive and there have been very costly hyperinflations.

The way to avoid excessive inflation is to create fiscal balance and monetary discipline. The big inflations between World War I and World War II ended when fiscal balance was restored. Monetary discipline comes in many forms. It requires an independent central bank, immune from political pressures. It may also require a central bank focused on an inflation target, paying less attention to other macroeconomic issues.

Key Links

- PBS history of German hyperinflation: http://www.pbs.org/wgbh/commandingheights/shared/minitext/ess_germanhyperinflation.html
- Federal Reserve releases on the measures of the money stock: http://www.federalreserve.gov/releases/h6
- International Monetary Fund country report on Zimbabwe: http://www.imf.org/external/pubs/ft/scr/2011/cr11135.pdf
- Australia Monetary Policy: http://www.rba.gov.au/monetary-policy/about.html
- Jim Bullard, Federal Reserve Bank president, Seven Faces of “the Peril”: http://research.stlouisfed.org/publications/review/10/09/Bullard.pdf
- Gary Stern, Federal Reserve Bank president, on deflation: http://www.minneapolisfed.org/research/pub_display.cfm?id=3354
- Federal Reserve Bank of Minneapolis publication on deflation: http://www.minneapolisfed.org/publications_papers/pub_display.cfm?id=3350

exercises

1. What is the difference between the quantity equation and the quantity theory of money?
2. According to the classical dichotomy, what happens to the real money supply if the nominal money supply grows at 10 percent?
3. If you were to draw a line through the points in Figure 26.7.3, it would not pass through the origin. Can you explain why? (Hint: examine the equation for the quantity equation, expressed in growth rates.)
4. Looking at Figure 26.7.2, you might be fooled into thinking that the inflation between 2002 and 2003 was almost as bad as that between 1990 and 1991. Why would this reasoning be a mistake?
5. The chapter contains two perspectives on Germany. The first is the hyperinflation in Germany during the 1920s, and the second is current Germany with low inflation and an independent central bank. How would you describe the differences in economic achievement (inflation, output growth, and unemployment) between these two versions of Germany? What were the institutional differences between these two versions of Germany?
6. Looking at Table 26.3.2 "The Start of the Hyperinflation in Zimbabwe", what happened to the velocity of money in Zimbabwe during the hyperinflation?
7. If the central bank takes the view that producing at potential GDP is efficient, then does it face a commitment problem?
8. In 2010, the state of California faced severe budgetary problems. If the state could print dollars, how would that relieve its budget problems? Who would pay the inflation tax?

Economics Detective

1. In the United States, how many central banks are there?

2. In note 5, we mention a measure of the money supply called “M2.” There are other measures of the money supply. For example, “M1” refers to currency and other assets that are immediately available for spending purposes. Find the most recent measure of the stocks of M1 and M2 for the United States.
3. Calculate the velocity of money for a country other than the United States.
4. The chapter did not present data on other recent periods of high inflation in countries such as Argentina, Brazil, Israel, and others. Search the Internet to find data on the inflation experiences of these countries. Create a graph of the growth rates of inflation and money in one of these countries.
5. It might be that countries have high money growth and thus high inflation because these are the goals of their monetary authority. See whether you can find a monetary authority with a stated goal of high inflation. If not, then think about why countries experience inflation if that is not the objective of the monetary authority?
6. What countries are dollarized in the world economy? Try to find out how dollarization influenced the inflation rate in that country.
7. Try to find a statement of the objectives of the Central Bank of Argentina. Part of independence is the way in which the decision makers at the central bank are appointed. How are these appointments made in Argentina?
8. Go to the web page for the Bank of Australia to learn about inflation targeting. What is their inflation target? How is it determined? What happens if they miss the target? Compare this to the objective and policy decisions of the Fed in the United States. What other central banks follow an inflation-targeting rule?
9. Is monetary policy in the United States guided by an inflation target? Does the European Central Bank use an inflation target?

Spreadsheet Exercise

1. Create a version of Table 26.3.1 "Prices in Germany" using a spreadsheet. Examine quarterly data for the United States or another country for the years 2007 to 2009. For prices, use the GDP implicit price deflator. Use the spreadsheet to calculate the inflation rate. Then put in a measure of the money supply and real GDP. Use the spreadsheet to calculate the velocity of money. Is the velocity of money approximately constant?

This page titled 26.7: End-of-Chapter Material is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

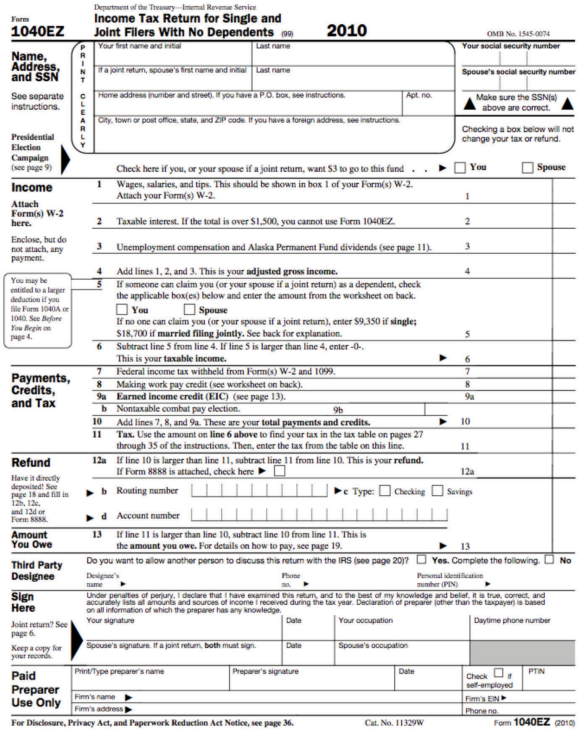
27: Income Taxes

27.1: Tax Day
27.2: Basic Concepts of Taxation
27.3: The Kennedy Tax Cut of 1964
27.4: Income Taxes and Saving
27.5: The Reagan Tax Cut
27.6: End-of-Chapter Material

This page titled 27: Income Taxes is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

27.1: Tax Day 27.1: Tax DayEdit Update titlecancel

Every year, in the middle of April, US citizens and residents are required to file an income tax form. The following figure shows the 1040EZ tax form, which is the simplest of all these tax forms. For the majority of us, this is one of the most direct pieces of contact that we have with the government. Based on the declarations we file, we are required to pay taxes on the income we have earned over the year. These tax revenues are used to finance a wide variety of government purchases of goods and services and transfers to households and firms. Of course, income taxes are not unique to the United States; most other countries require their residents to complete a similar kind of form.



Department of the Treasury—Internal Revenue Service

Form **1040EZ** **Income Tax Return for Single and Joint Filers With No Dependents** (99) **2010** OMB No. 1545-0074

Name, Address, and SSN — See separate instructions. PRINT / CLEARLY

Your first name and initial | Last name | Your social security number

If a joint return, spouse's first name and initial | Last name | Spouse's social security number

Home address (number and street). If you have a P.O. box, see instructions. | Apt. no. | Make sure the SSN(s) above are correct.

City, town or post office, state, and ZIP code. If you have a foreign address, see instructions. | Checking a box below will not change your tax or refund.

Presidential Election Campaign (see page 9) — Check here if you, or your spouse if a joint return, want $3 to go to this fund . . ▶ ☐ You ☐ Spouse

Income — Attach Form(s) W-2 here. Enclose, but do not attach, any payment.

1	Wages, salaries, and tips. This should be shown in box 1 of your Form(s) W-2. Attach your Form(s) W-2.	1
2	Taxable interest. If the total is over $1,500, you cannot use Form 1040EZ.	2
3	Unemployment compensation and Alaska Permanent Fund dividends (see page 11).	3
4	Add lines 1, 2, and 3. This is your **adjusted gross income.**	4
5	If someone can claim you (or your spouse if a joint return) as a dependent, check the applicable box(es) below and enter the amount from the worksheet on back. ☐ You ☐ Spouse. If no one can claim you (or your spouse if a joint return), enter $9,350 if **single;** $18,700 if **married filing jointly.** See back for explanation.	5
6	Subtract line 5 from line 4. If line 5 is larger than line 4, enter -0-. This is your **taxable income.** ▶	6

You may be entitled to a larger deduction if you file Form 1040A or 1040. See *Before You Begin* on page 4.

Payments, Credits, and Tax

7	Federal income tax withheld from Form(s) W-2 and 1099.	7
8	Making work pay credit (see worksheet on back).	8
9a	**Earned income credit (EIC)** (see page 13).	9a
b	Nontaxable combat pay election. 9b	
10	Add lines 7, 8, and 9a. These are your **total payments and credits.** ▶	10
11	**Tax.** Use the amount on **line 6 above** to find your tax in the tax table on pages 27 through 35 of the instructions. Then, enter the tax from the table on this line.	11

Refund — Have it directly deposited! See page 18 and fill in 12b, 12c, and 12d or Form 8888.

12a	If line 10 is larger than line 11, subtract line 11 from line 10. This is your **refund.** If Form 8888 is attached, check here ▶ ☐	12a
▶ b	Routing number ▶ c Type: ☐ Checking ☐ Savings	
▶ d	Account number	

Amount You Owe

13	If line 11 is larger than line 10, subtract line 10 from line 11. This is the **amount you owe.** For details on how to pay, see page 19. ▶	13

Third Party Designee — Do you want to allow another person to discuss this return with the IRS (see page 20)? ☐ **Yes.** Complete the following. ☐ **No**

Designee's name ▶ | Phone no. ▶ | Personal identification number (PIN) ▶

Sign Here — Joint return? See page 6. Keep a copy for your records.

Under penalties of perjury, I declare that I have examined this return, and to the best of my knowledge and belief, it is true, correct, and accurately lists all amounts and sources of income I received during the tax year. Declaration of preparer (other than the taxpayer) is based on all information of which the preparer has any knowledge.

Your signature | Date | Your occupation | Daytime phone number

Spouse's signature. If a joint return, **both** must sign. | Date | Spouse's occupation

Paid Preparer Use Only

Print/Type preparer's name | Preparer's signature | Date | Check ☐ if self-employed | PTIN

Firm's name ▶ | Firm's EIN ▶

Firm's address ▶ | Phone no.

For Disclosure, Privacy Act, and Paperwork Reduction Act Notice, see page 36. Cat. No. 11329W Form **1040EZ** (2010)

Figure 27.1.1: Easy Tax Form

From the perspective of a household or a firm, the tax form is a statement of financial responsibility. From the viewpoint of the government, the 1040 tax form is an instrument of fiscal policy. The 1040 form is based on the US tax code, and changes in that code can have profound effects on the economy—both in the short run and in the long run.

In this chapter, we study the various ways in which income taxes affect the economy. An understanding of taxes is critical for policymakers who devise tax policies and for voters who elect them. Tax policies are often controversial, in large part because they affect the economy in several different ways. For example, in the 2004 and 2008 US presidential campaigns, one of the most contentious economic policy issues was an income tax cut that President George W. Bush had initiated in his first term and that the Republican Party wished to make permanent. That issue returned to the forefront of political discussion in 2010, when these tax cuts were renewed.

Politicians have argued about such matters since the country was founded. Should the government ensure it has enough tax revenue to balance its budget? How should we raise the revenues to pay for our government programs? What is the appropriate tax on the income received by individuals and corporations? Fiscal policy questions like these are debated in the United States and other countries throughout the world. They are tough questions for politicians and economists alike.

Politicians focus largely on who wins and loses—which groups will bear the burden of taxes and receive the benefits of government spending and transfers? They do so for political reasons and because one goal of a tax system is to redistribute income. Economists emphasize something rather different. Economists know that taxes are necessary to finance government expenditures. At the same time, they know that taxes can have the negative effect of distorting people's decisions and lead to inefficiency. Hence

economists focus on designing a tax system that achieves its goals of raising revenue and redistributing income, without distorting the decisions of individuals and firms too much.

In addition, macroeconomists have observed that taxes significantly affect overall economic performance, as measured by variables such as real gross domestic product (real GDP) growth or the unemployment rate. The government can use changes in taxes as a means of influencing aggregate spending in the economy. In the United States, the federal government has often changed income taxes to affect overall economic performance. In this chapter, we examine two examples: the tax policies of the Kennedy administration of 1960–63 and the Reagan administration of 1980–88.

Our discussion of the Kennedy tax cut experience highlights the way in which variations in income taxes are used to help stabilize the macroeconomy. We use the Reagan tax cuts of the early 1980s to explore the growth implications of income taxes, which are often called “supply-side effects.”

Road Map

Our approach to understanding the effects of income taxes on the economy is summarized in Figure 27.1.2:

- Taxes affect consumption and hence aggregate expenditure and output.
- Taxes affect saving and hence the capital stock and output.
- Taxes affect labor supply and hence output.

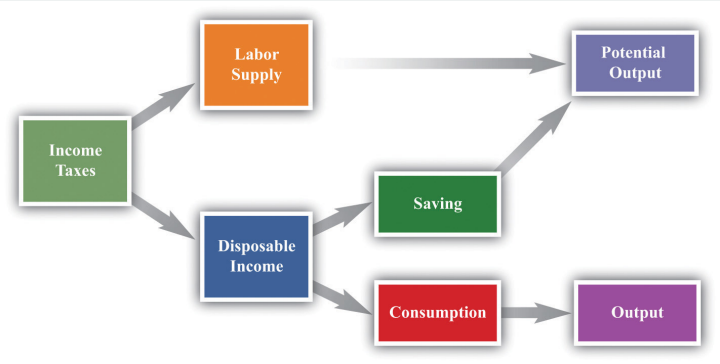


Figure 27.1.2: Macroeconomic Effects of Tax Policy

Any change in the income tax regime affects both the spending and the supply sides of the economy. Our reason for thinking separately about the Kennedy and Reagan tax experiments is to isolate the spending effects and the supply effects. Once you understand these different channels, you will be equipped to evaluate other tax policies, such as those adopted later by President George W. Bush. Finally, the figure reveals that the choice between consumption and saving and the choice between work and leisure are at the heart of our analysis.

This page titled 27.1: Tax Day is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

27.2: Basic Concepts of Taxation 27.2: Basic Concepts of TaxationEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What is the difference between a marginal and an average tax rate?
2. How does the tax system redistribute income?

Before delving into the details of President Kennedy's tax policy, we review the basics of personal income taxation. This review is not only helpful for your study of economics but also may be useful when you have to fill out your own income tax form. Even a quick glance at the 1040EZ form in Figure 27.2.1 suggests that taxes are a very complex topic. Indeed, the US federal tax code governing income taxes alone runs to thousands of pages. The taxes that you pay depend on your adjusted gross income (line 4), which is the income you receive from a variety of sources (the main components noted on the return are wages, interest income, and unemployment compensation). But there is also a "standard deduction" and an "exemption" (line 5)—for a single person in 2010, these totaled $9,350. For the EZ form, your taxable income is given as the following:

$$taxable\ income = adjusted\ gross\ income - (deduction + exemption).$$

If your financial situation is very simple, you can file this EZ form. However, if you receive income from other sources (such as dividends on stocks), or if you wish to "itemize" your deductions (for payments of interest on home mortgages, dependent children, property taxes, and so forth), you have to file a more complicated form, often with several other forms containing supplementary information. Thus the calculation of adjusted gross income and deductions can be quite complex. For all individuals, however, the basic relationship still holds:

$$taxable\ income = adjusted\ gross\ income - (deductions\ and\ exemptions).$$

Once you know your taxable income, there are then different tax rates for different income levels.Even this is not quite the whole story. There are various tax credits for which some individuals are eligible, and there is also something called the alternative minimum tax, which must be calculated.

Marginal and Average Tax Rates

From the perspective of macroeconomics, this complexity is daunting, particularly when we remember that the details of the tax system vary from country to country and year to year. The income tax is evidently not a simple thing that can be incorporated in a straightforward way into our frameworks. We cannot hope to incorporate all these features of the tax code into our theory without getting completely bogged down in the details. If we are going to make sense of how taxes affect consumption behavior, we must leave out most of these complicating elements. The challenge for economists is to decide which features of the tax system are critical for our analysis and which are peripheral and can be safely ignored.

One noteworthy feature of the income tax system is that not everyone pays the same amount of tax. Table 27.2.1 "Revised 2010 Tax Rate Schedules" shows the income tax schedule for the year 2010 for a single taxpayer.There are other schedules for members of a household filing jointly. These and related tables are available from "Forms and Publications," Internal Revenue Service, accessed September 20, 2011, http://www.irs.gov/formspubs/index.html. It indicates how much tax a must be paid for a given level of taxable income.

Table 27.2.1: Revised 2010 Tax Rate Schedules

If Taxable Income		The Tax Is Then		
Is Over (in US$)	But Not Over (in US$)	This Amount (in US$)	Plus This (%)	Of the Excess Over (in US$)
0	8,375	0	10	0
8,375	34,000	837.50	15	8,375
34,000	82.400	4,681.25	25	34,000

If Taxable Income		The Tax Is Then		
82.400	171,850	16,781.25	28	82.400
171,850	373,650	41,827.25	33	171,850
373,650	—	108,421.25	35	373,650

To use this table, you must first find your taxable income. Suppose it is $20,000. Your tax is then determined from the second row of the table. You would owe $837.50 + 0.15 \times (20,000 - 8,375)$, which is $2,581.25.

Figure 27.2.1 shows the relationship between taxes and income implicit in the tax schedule summarized in Table 27.2.1 "Revised 2010 Tax Rate Schedules". This figure shows the amount of tax you must pay given your adjusted gross income (upper panel) and your taxable income (lower panel). We see two key facts:

1. As an individual's income increases, he or she pays more in tax (the line slopes upward).
2. As an individual's income increases, he or she pays a larger fraction of *additional* income in tax (the line becomes steeper at higher levels of income).

This leads us to two ways to think about the tax schedule a household faces.

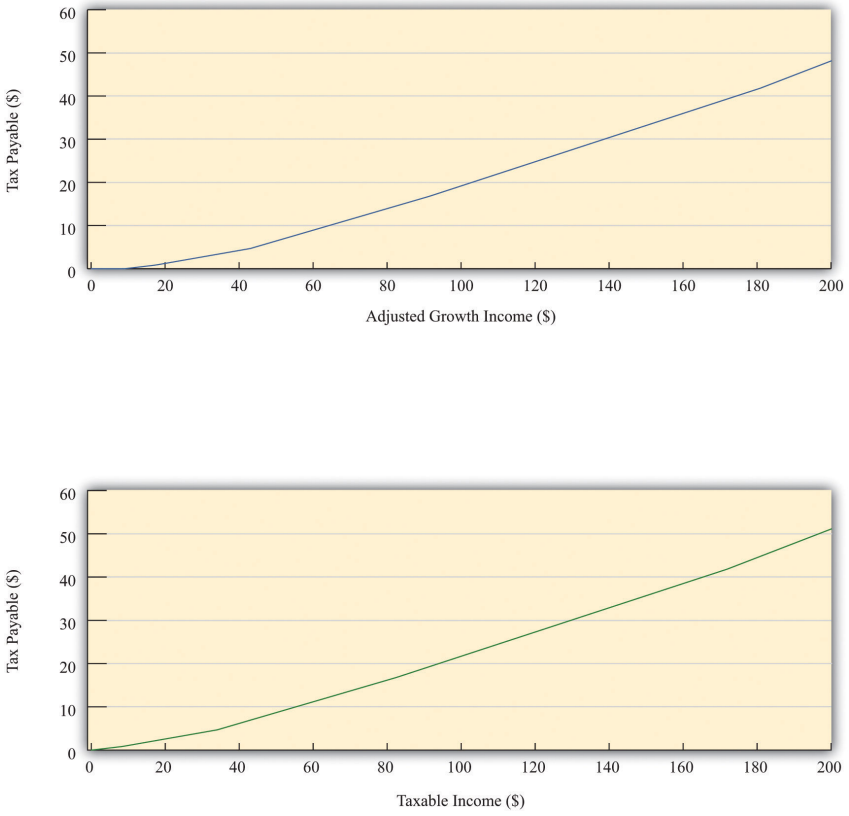


Figure 27.2.1

The figure shows the amount of tax owed by a single individual in the United States who takes the "standard deduction." The upper panel has adjusted gross income on the horizontal axis, whereas the lower panel has taxable income on the horizontal axis.

As shown in Table 27.2.1 "Revised 2010 Tax Rate Schedules", there were six different tax rates in effect in 2010, ranging from 10 percent for low-income individuals to 35 percent for high-income individuals. The tax rates in the fourth column are the **marginal tax rates** since they represent the tax rate paid on marginal (that is, additional) income. Thus higher income households pay higher marginal tax rates. The marginal tax rate can be seen graphically as the slope of the line in Figure 27.2.1.

We are often interested in knowing what fraction of an individual's income goes to taxes. This is called the **average tax rate**. Returning to the example we calculated earlier, if you have an income of $20,000 and thus pay taxes of $2,581.25, your average tax rate is equal to , or 12.9 percent. The marginal tax rate of 15 percent is greater than the average tax rate of 12.9 percent. There is a

difference between the tax you pay on average and the tax rate charged on the last dollar of income.The average tax rate can also be given a graphical interpretation. It is the slope of a line from the origin to the point on the graph.

Leaving aside the details of exemptions and deductions, the essence of the income tax code is captured in the table and figures we have just presented. Even these, however, are quite complicated. We want to build income taxes into our framework of the economy, so it would be nice if we could decide on a simpler way to represent the tax code. The art of economics lies in deciding how to take something complicated, like the US income tax code, and represent it in as simple a way as possible *while still retaining the features that matter to the problem under discussion.*

Looking at Figure 27.2.1, we can see that the relationship between taxes paid and taxable income looks approximately like a straight line. It is not exactly a straight line because it becomes steeper as marginal tax rates increase. For our purposes in this chapter, however, it is a reasonable simplification to represent this relationship as a line—that is, to suppose that the marginal tax rate is constant.

In addition, we ignore the standard deduction and exemption. That is, we suppose that people start paying taxes on their very first dollar of income. Thus we suppose that

$$taxes\ paid = tax\ rate \times income.$$

Representing the tax schedule this way is fine if we want to examine the economy as a whole and are not particularly concerned with the way in which taxes affect different households. We use this simplified model of the tax system at various times in this chapter.

Effects of Changes in Tax Rates

We can use this simple model of the tax system to see how a change in the income tax rate affects both individuals and the economy as a whole. Suppose there is a cut in the tax rate. Since $taxes\ paid = tax\ rate \times income$, the immediate impact is to reduce the amount of taxes households pay: for a given income, a reduction in the tax rate reduces taxes paid. This means that **disposable income**, which is the income left over after paying taxes and receiving transfers, increases.

What do households do with the increase in disposable income? A likely answer is that a typical household spends some of this extra income and saves the remainder. If all households follow this pattern, then the increased spending by each household translates into larger consumption in the aggregate economy. At this point, the power of the **circular flow of income** will take over, and the level of income and output in the economy will increase even further.

> **Toolkit: Section 31.27**
>
> You can review the circular flow of income in the toolkit.

As the economy expands, the amount of taxes paid starts to increase. In other words, one consequence of a tax cut is that the tax base (income) expands. The ultimate effect of a tax cut on the overall amount of taxes paid depends on both this expansion of the tax base (income) and the reduction of the tax rate.

Taxes and Income Distribution

The effects of a tax cut are not the same for everyone. Changes in the tax code affect the distribution of income. If we want to understand such effects, however, it is a mistake to use our simple model of the tax system. We must instead examine how marginal tax rates are different at different levels of income. Suppose that marginal tax rates increase with income, which means that average tax rates increase with income. Higher income households then pay a larger fraction of their income as taxes to the government. As a result, the distribution of income after taxes is more equal than the distribution of income before taxes.

Imagine that we take two individuals with different levels of income and calculate their tax payments and after-tax income. Suppose that the first individual earns $20,000 per year and the other earns $200,000. Table 27.2.2 "The Redistributive Effects of Taxation (in US$)" shows the amount of tax each pays and their income after taxes, based on the tax schedule from Table 27.2.1 "Revised 2010 Tax Rate Schedules". Notice from the table that the marginal tax of the high-income household is 33 percent, compared with the 15 percent marginal tax of the low-income household. The total tax paid by the high-income individual is $51,116.75, which is almost 20 times the tax paid by the low-income household. Whereas the pre-tax income of the richer household was 10 times greater than that of the poorer household, its after-tax income is 8.5 times greater.

Table 27.2.2: The Redistributive Effects of Taxation (in US$)

Income	Tax Paid	Income after Taxes
20,000	2,581.25	17,418.75
200,000	51,116.75	148,883.25

This example shows that the tax code redistributes income from high-income to low-income households. What is more, the redistribution does not necessarily stop here. We have not said anything about what the government does with the tax revenues it receives. If the government transfers all those revenues to low-income households, then the combined redistributive effect of taxes and transfers is even stronger.

When we talk about the effects of taxes on labor supply and disposable income, keep in mind that the size of these effects is different for households at different levels of income. These varying effects matter for the politics of tax cuts because lawmakers pay close attention to which income groups are affected by tax policy.

Key Takeaways

1. The marginal tax rate is the rate paid on an additional dollar of income, and the average tax rate is the ratio of taxes paid to income.
2. When the marginal tax rate is increasing in income, then the tax system redistributes from richer households to poorer households. In this case, after-tax income is more equal than income before taxes are paid.

Exercises

1. Use Table 27.2.1 "Revised 2010 Tax Rate Schedules" to calculate the tax you would pay if your income were $30,000.
2. If taxes paid equal the tax rate times income, what happens to the average tax rate when the marginal tax rate changes?

This page titled 27.2: Basic Concepts of Taxation is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

27.3: The Kennedy Tax Cut of 1964 27.3: The Kennedy Tax Cut of 1964Edit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What was the state of the economy prior to the Kennedy tax cut of 1964?
2. What framework did economists at that time use to predict the effects of this tax cut?
3. What was the response of the economy to this tax cut?

Now that we have some basic idea of how income taxes work, we turn to the Kennedy tax cut of 1964. We begin with some background information; we then develop the economic tools needed to analyze the effects of the tax policy on household consumption and thus on real gross domestic product (real GDP).

The Scenario

In his inaugural presidential address, President Kennedy famously said, “My fellow Americans, ask not what your country can do for you; ask what you can do for your country.” The Kennedy administration recruited top individuals in all fields (“the best and the brightest”) to come to Washington in this new spirit of commitment to public service.See David Halberstam, *The Best and the Brightest* (New York: Ballantine Books, 1972).

Every president has a group of economists, known as the Council of Economic Advisors (CEA; http://www.whitehouse.gov/cea), that provides advice on economics and economic policy. The list of members and staff of the 1961 CEA reads today like a “who’s who” of economics. James Tobin and Robert Solow were prominent members of the economics team; both went on to win Nobel Prizes in Economics. The chairman of the CEA was Walter Heller, an economist known for a wide variety of contributions on the conduct of macroeconomic policy.

The economists in the Kennedy administration observed that there had been three recessions in the two Eisenhower administrations (1952–1960): one from 1953 to 1954 after the Korean War, one from 1957 to 1958, and one in 1960. You can see these in Figure 27.3.1. The CEA members and staff thought that more aggressive fiscal and monetary policies could be used to keep the economy more stable and prevent such recessions. Their goal of moderating fluctuations in the economy was based on the framework of the basic **aggregate expenditure model**, which had been developed in the aftermath of the Great Depression, augmented by some developments in economic thinking from the 1940s and 1950s. Based on that analysis, they believed that fiscal and monetary policies could be used to control aggregate spending and hence real GDP.

Toolkit: Section 31.30

You can review the aggregate expenditure model in the toolkit.

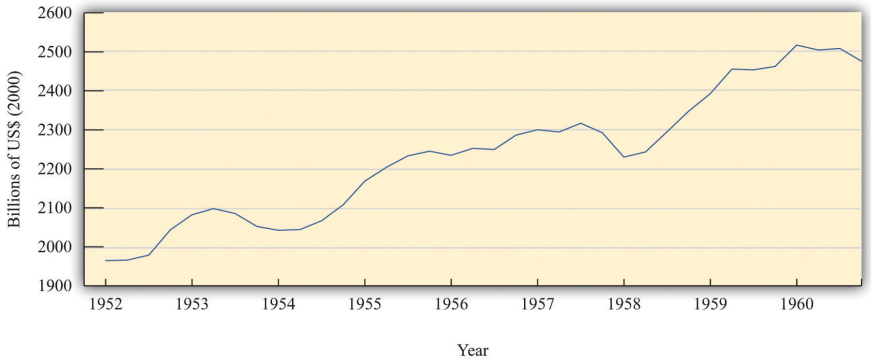


Figure 27.3.1: *Real GDP in the 1950s*. The chart shows real GDP in the United States between 1952 and 1960, measured in billions of year 2000 dollars.

Source: Bureau of Economic Analysis.

This group of economists had, on one hand, a clearly defined goal of stabilizing the macroeconomy and, on the other hand, a set of policy *instruments*—economic variables such as taxes, government spending, and interest rates—that were under the control of policymakers. They also had a framework of analysis (the aggregate expenditure model) that explained how these instruments could be used to achieve their goals. Finally, they had a president who was willing to listen and take their advice. Never before had economists had such tools and wielded such influence.

The opportunity to test their ideas arose toward the middle of the Kennedy presidency. In the middle of 1962, it was apparent to the Kennedy administration economists that the economy was beginning to sputter. The growth rate of real GDP was 7.1 percent in 1959 but decreased to 2.5 percent and 2.3 percent in 1960 and 1961, respectively.*Economic Report of the President* (Washington, DC: GPO, 2005):table B-2, GPO Access, accessed September 20, 2011, www.gpoaccess.gov/eop/2005/2005_erp.pdf. Their response was to initiate a tax cut.

As is usually the case when a major fiscal policy action is under consideration, there was a lengthy time lag between the initiation of the policy and its implementation. Even though the tax cut was proposed in 1962, President Kennedy never saw it put into effect. He was assassinated in November 1963; the tax cut for individual households and corporations was not enacted until early 1964. For households, tax withholding rates decreased from 18 percent to 14 percent, leading to an estimated tax reduction of about $6.7 billion. Taxes on corporations were also decreased; the reduction in taxes for 1964 was expected to be about $1.7 billion. By 1965, the economists expected that taxes would be lower by $11 billion. In 1965, nominal GDP was about $719 billion, so these changes were about 1.5 percent of nominal GDP.

For many observers of the macroeconomy, this was a watershed event. The *Economic Report of the President* proclaimed 1965 the "Year of the Tax Cut." In retrospect, these years were the heyday of Keynesian macroeconomics: for the first time, the government was using tax policy in an attempt to fine-tune the economy.

Figure 27.3.2 shows what happened to average and marginal tax rates. Marginal tax rates were very high at the time—much greater than in the present day. At high levels of income, more than 90 cents of every additional dollar had to be paid to the government in taxes. Consequently, average tax rates were also high: an individual with taxable income of $100,000 (a very high level of income back then) had to pay about two thirds of that amount to the government. The Kennedy tax cuts reduced these tax rates. Even after the tax cut, the marginal and average tax rates both increased with income. In other words, the tax system still redistributed income across households. But when we compare 1963 and 1964, we see that the marginal tax rate did not increase as rapidly under the new tax policy. Therefore, this channel of redistribution was weaker under the new tax policy.

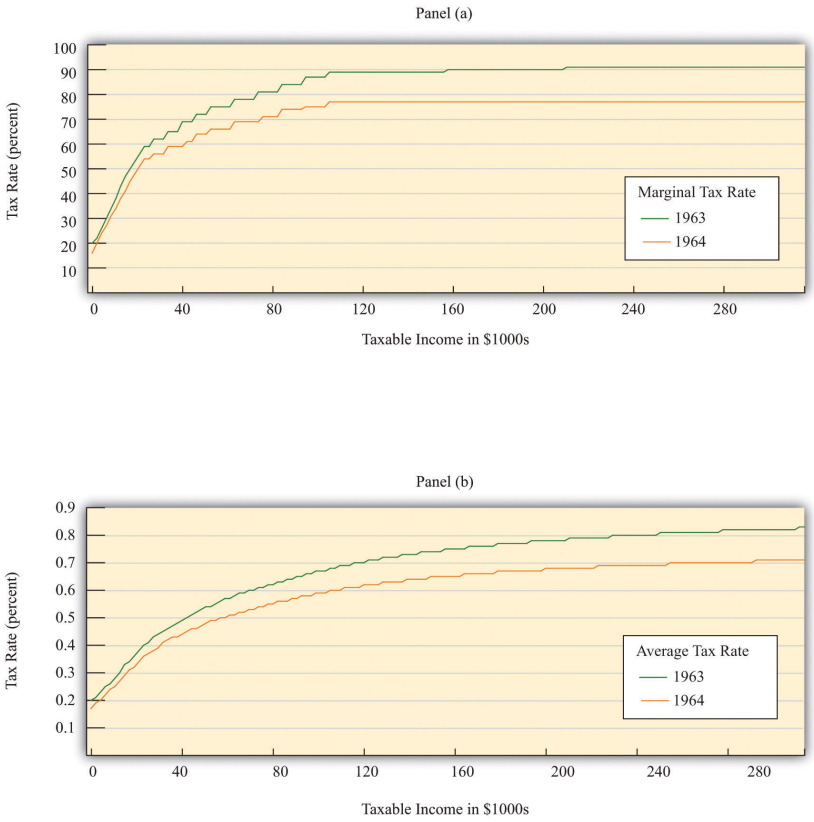


Figure 27.3.2: Tax Policy during the Kennedy Administration. The charts show the impact of the Kennedy tax cut. Part (a) highlights how the marginal tax rates for households changed from 1963 to 1964, and part (b) shows the impact on average tax rates.

Source: Department of the Treasury, IRS 1987, "Tax Rates and Tables for Prior Years" Rev 9-87

For their policy to be successful, Kennedy's advisors had to ask and then answer a series of questions. How big a tax cut should they recommend? How long should it last? What would be the effect on government revenues? What would be the effect on real GDP and consumption? Economists working in government today confront exactly the same questions when contemplating changes in tax policy. Questions such as these epitomize economics and economists at work.

Looking back at this experiment with almost half a century of hindsight, we can ask additional questions. How well did these policies work in terms of achieving their goal of economic stabilization? What actually happened to consumption and output? Was the tax policy successful?

The Kennedy economists needed a quantitative model of economic behavior: a formalization of the links between their policy tools (tax rates) and the outcomes that they cared about, such as consumption and output. Using the aggregate expenditure model, they wanted to know how big a change in real GDP they could expect from a given change in the tax rate. To use the model to study income taxes, we need to add some theory about how spending responds to changes in taxes. Accordingly, we study the effects of income taxes on household consumption and then discuss how changes in consumption lead to changes in output.

Although we are using a historical episode to help us understand the effect of taxes on the economy, this chapter is not intended as a lesson in economic history. Variations of this same model are still used today to analyze current economic policies. Indeed, in response to the economic crisis of 2008, many countries around the world cut taxes in an attempt to stimulate their economies. By studying the experience of the early 1960s, we gain insight into a critical part of macroeconomics: the linkage between consumption and output.

Having said that, economics has advanced significantly since the 1960s, and the state-of-the-art analysis for that time seems oversimplified today. Modern economists think that the policy advisers in the 1960s neglected some key aspects of the economy. Their insights were not wrong, but they were incomplete. Our understanding of the economy has evolved since Tobin, Solow, and Heller designed the nation's tax policy.

Household Consumption

We begin by studying the relationship between consumption and income. We first develop some ideas about how households make consumption decisions, and, on the basis of those ideas, we make some predictions about what we expect to happen when there is a cut in taxes. We then examine the evidence from the Kennedy tax cut.

Income, Consumption, and Saving

In microeconomics, we study how a consumer allocates incomes across a wide variety of products. Microeconomists interested in studying, say, the market for ice cream examine how households choose between ice cream and other products that are close substitutes, such as frozen yogurt, and between ice cream and other products that are complements, such as hot fudge sauce. When studying microeconomics, however, we focus on choices for goods made at a particular point in time.

Macroeconomics has a different emphasis. It emphasizes the choice between consumption and saving. Instead of thinking about the consumption of ice cream today versus frozen yogurt today, we study the choice between consumption today and consumption in the future. To highlight this decision, macroeconomists downplay the choices among different goods and services. Of course, in reality, households decide both how much to spend and how much to save, and what products to purchase. But it is convenient to treat these decisions separately.

The same basic ideas of household decision making apply in either case. Households distribute their income across goods to ensure that no redistribution of that spending would make them better off. This is true whether we are talking about ice cream and frozen yogurt, or about consumption and saving. Households allocate their income between consumption and savings in a way that makes them as well off as possible. They do not spend all their income this year because they want to save some for consumption in the future.

Suppose a household in the United States had taxable income of $20,000 in 2010. Some of this income goes to the payment of taxes to federal and state governments. (From our earlier discussion, the average federal tax rate is 13.25 percent.) The rest is either spent on goods and services or saved. The income that is spent on goods and services today is spread over the many products that a household buys. The income that is saved will likewise be used in the future to purchase different goods and services.

Personal Income and Disposable Income

The most basic measure of aggregate economic activity is real GDP, which is the total amount of final goods and services produced in our economy over a period of time, such as a year. The rules of national income accounting mean that real GDP measures three different things at once: the production or output of the economy, the spending in the economy; and the income generated in the economy. We use real GDP as our most general measure of income.

We work in this chapter with two further concepts of income from the national accounts: **personal income** and **disposable income**. Some of the income generated in the economy is retained by firms to finance new investment, so it does not go to households. Personal income refers to that portion of GDP that finds its way directly into the hands of households. (At the level of an individual household, it corresponds closely to adjusted gross income on the tax form.) Disposable income is what remains after we subtract from personal income the taxes paid by households to the government and add to personal income the transfers (such as welfare payments) received by households from the government. For a household, disposable income measures its available resources after taxes have been paid and transfers received.

Consumption Smoothing

Our starting point for understanding consumption choices is the household budget constraint for a typical household. The household receives income from working and other sources and pays taxes to the government. The remainder is the household's disposable income. The household budget constraint reminds us that, ultimately, you must either spend the income you receive or save it; there are no other choices. That is,

$$disposable\ income = consumption + saving.$$

A theory of consumption is a theory of how households decide to divide their income between consumption and saving. Saving is a way to convert current income into future consumption. A theory of consumption is equivalently a theory of saving. A fundamental idea about household behavior is that people do not wish their consumption to vary a lot from month to month or year to year. This principle is so important that economists give it a special name: **consumption smoothing**. Households use saving and borrowing to smooth out fluctuations in their income and keep their consumption relatively smooth. People will tend to save when their income

is high and will dissave when their income is low. (*Dissave* is the word economists use to mean either running down one's existing wealth or borrowing against future earnings.)

> **Toolkit: Section 31.32**
>
> You can review the consumption-saving decision in the toolkit.

Perfect consumption smoothing means that the household consumes exactly the same amount in each period of time (for example, a month or a year). If a construction worker earns $10,000 per month working from May to October but nothing for the rest of the year, we do not expect that he will spend $10,000 per month in the summer and then starve in the winter. It is much more likely that he will save half of his income in the summer and spend those savings in the winter, so that he spends about $5,000 per month throughout the year.

The logic of consumption smoothing is the same as the argument for why households buy many different goods rather than one single good. Households typically take their income and spend it on a wide variety of products. Furthermore, when income increases, the household will spread this extra income across the spectrum of goods it consumes; not all of it is spent on one good. If you obtain more income, you do not spend all this extra income on ice cream, for example. You buy more of many different goods.

The Consumption Function

One way to represent consumption smoothing is by means of a **consumption function**. This is an equation that relates current consumption to current disposable income. It allows us to go from an abstract idea about consumption behavior—consumption smoothing—to a specific formulation of consumption that we can use in a model of the aggregate economy.

We suppose the consumption function can be represented by the following equation:

$$consumption = autonomous\ consumption + marginal\ propensity\ to\ consume \times disposable\ income.$$

- We make three assumptions:
  1. Autonomous consumption is positive. Households consume something even if their income is zero. If the household has accumulated a lot of wealth in the past or if the household expects its future income to be larger, autonomous consumption will be larger. It captures both the past and the future.
  2. We assume that the marginal propensity to consume is positive. The marginal propensity to consume captures the present; it tells us how changes in current income lead to changes in current consumption. Consumption increases as current income increases; the larger the marginal propensity to consume, the more sensitive current spending is to current disposable income. By contrast, the smaller the marginal propensity to consume, the stronger is the consumption-smoothing effect.
  3. We also assume that the marginal propensity to consume is less than one. This says that not *all* additional income is consumed. When the household receives more income, it consumes some and saves some. The marginal propensity to save is the amount of additional income that is saved; it equals (1 – marginal propensity to consume).

Table 27.3.1 "Consumption, Income, and Saving" contains an example of a consumption function where autonomous consumption equals 10,000 and the **marginal propensity to consume** equals 0.8. If the household earns no income at all (disposable income = $0), it still spends $10,000 on consumption. In this case, savings equal −$10,000. This means the household is either drawing on existing wealth (accumulated savings from the past) or borrowing against income expected in the future. The marginal propensity to consume tells us how the household divides *additional* income between consumption and saving. In our example, the household spends 80 percent of any additional income and saves 20 percent.

Table 27.3.1: Consumption, Income, and Saving

Disposable Income ($)	Consumption ($)	Saving ($)
0	10,000	−10,000
10,000	18,000	−8,000
20,000	26,000	−6,000
30,000	34,000	−4,000

Disposable Income ($)	Consumption ($)	Saving ($)
40,000	42,000	−2,000
50,000	50,000	0
60,000	58,000	2,000
70,000	66,000	4,000
80,000	74,000	6,000
90,000	82,000	8,000
100,000	90,000	10,000

For example, when income is equal to $20,000, consumption can be calculated as follows:

$$consumption = \$10,000 + 0.8 \times \$20,000 = \$10,000 + 0.8 \times \$20,000 = \$26,000.$$

The household is still dissaving but now only by $6,000. Table 27.3.1 "Consumption, Income, and Saving" also shows that when income equals $50,000, consumption and income are equal, so savings are exactly zero. At income levels above $50,000, the household has positive savings.

Figure 27.3.3 shows the relationship between consumption and income graphically. We also graph the savings function in Figure 27.3.3. The savings function has a negative intercept because when income is zero, the household will dissave. The savings function has a positive slope because the **marginal propensity to save** is positive.

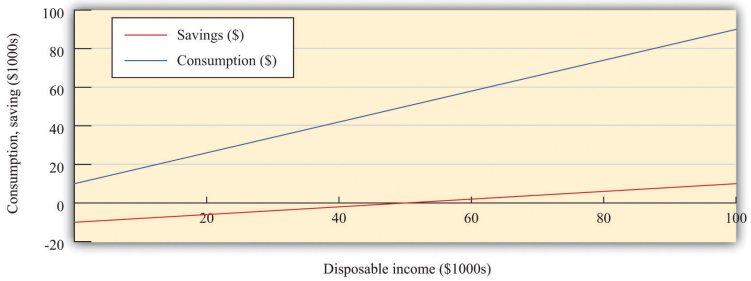


Figure 27.3.3: Consumption, Saving, and Income. The graph shows the relationship between consumption and disposable income, where autonomous consumption is $10,000 and the marginal propensity to consume is 0.8. When disposable income is below $50,000, savings are negative, whereas at income levels above $50,000, savings are positive.

As well as the marginal propensity to consume and the marginal propensity to save, we can examine the **average propensity to consume**, which measures how much income goes to consumption on average. It is calculated as follows:

When disposable income increases, consumption increases but by a smaller amount. This means that when disposable income increases, people consume a smaller fraction of their income: the average propensity to consume decreases.In terms of mathematics, we are saying that, if we divide through the consumption function by disposable income, we get an increase in disposable income reduces the first term and the average propensity to consume. Meanwhile, the ratio of saving to disposable income is called the **savings rate**. For example,

The savings rate and the average propensity to consume together sum to 1. In other words, a decline in the average propensity to consume equivalently means that households are saving a larger fraction of their income.

Because the consumption and savings relationships are two sides of the same coin, economists wishing to find the actual values of autonomous consumption and the marginal propensity to consume can examine data on consumption, savings, or both. If the data were perfect, we would get the same answer either way. For the United States, both consumption and savings data are readily available, but in some countries the data on savings may be of higher quality than the consumption data, in which case economists use savings data to understand consumption behavior.

Some Warnings about the Consumption Function

The consumption function is useful because it captures two fundamental insights: households seek to smooth their consumption, but consumption nonetheless responds to current income. But the consumption function is really too simple.Refining our theory of consumption is a subject for Chapter 28.

First, it ignores the role of accumulated wealth. If you consider two households with the same level of current income but different amounts of accumulated wealth, the one with higher wealth will probably consume more. Second, the consumption function does not explicitly include the role of expectations. A household's consumption reflects not only income today and the accumulation of income in the form of wealth but also anticipated income. So, for example, if a government announces that it will increase income tax rates in two years, we expect that households will respond immediately to smooth out the effects of these future taxes. The only way the consumption function allows us to capture wealth or expectations of future income is through autonomous consumption. This is fine as far as it goes, but it means that we are taking too many aspects of consumption as given, rather than explaining them with our theory.

Another complication is that changes in income today are often *correlated* with changes in income in the future. If your income increases today, is this an indication that your income will also be higher in the future? To see why this matters, consider two extreme examples. First, suppose that you receive a one-time inheritance of $10 million. What will you do with this income? According to the consumption smoothing argument, you will save some of this income to increase your consumption in the future. Roughly speaking, if you thought you had 10 years left to live, you might increase your consumption by about $1 million per year. In this case your marginal propensity to consume would be only 0.1.

Now suppose that instead of a $10 million windfall, you learn you will receive $1 million each year for the next 10 years. In this case, your income is already spread out over your lifetime. So, in this second case, you will again want to smooth your consumption. But since the increase in income will be maintained for your lifetime, you can increase your consumption by an amount equal to the increase in your income. Your marginal propensity to consume will be 1.0.

The difference between these two situations is that in the first case the income increase is temporary, and in the second it is permanent. The logic of consumption smoothing implies that the marginal propensity to consume is near 1 for permanent changes in income but much smaller for temporary changes in income.

The Effects of a Change in Income Taxes

We can now figure out the effects of a cut in taxes on consumption and saving. A reduction in taxes will increase disposable income. From the consumption function, this results in an increase in consumption equal to the marginal propensity to consume times the increase in disposable income. The average propensity to consume decreases. To summarize, if taxes are cut in the economy, we expect to see the following:

- An increase in disposable income
- An increase in consumption that is smaller than the increase in disposable income (that is, a marginal propensity to consume less than 1)
- A decline in the average propensity to consume

When natural scientists such as molecular biologists or particle physicists want to see how good their theories are, they conduct experiments. Economists and other social scientists have much less ability to carry out experiments—certainly at the level of the macroeconomy. The Kennedy tax cut, however, is like a "natural" experiment in that there was a major policy change that we can think of as a change in an **exogenous variable**. It is not, in truth, completely exogenous. We already explained that the tax cut was enacted in response to the poor performance of the economy. We are not badly misled by thinking of it as an exogenous event, however. We can therefore use it to see how well our theory performs. Specifically, we can look to see whether disposable income and consumption do behave as we have predicted.

Empirical Evidence on Consumption

Before we turn to those specific questions, let us examine some data on consumption. Figure 27.3.4 shows the behavior of consumption and disposable income from 1962 to 2010. The measures of both income and consumption are in year 2005 dollars. This means that the nominal (money) levels of income and consumption for each of the years have been corrected for inflation, so that we can see how the real level of consumption relates to the real level of income.

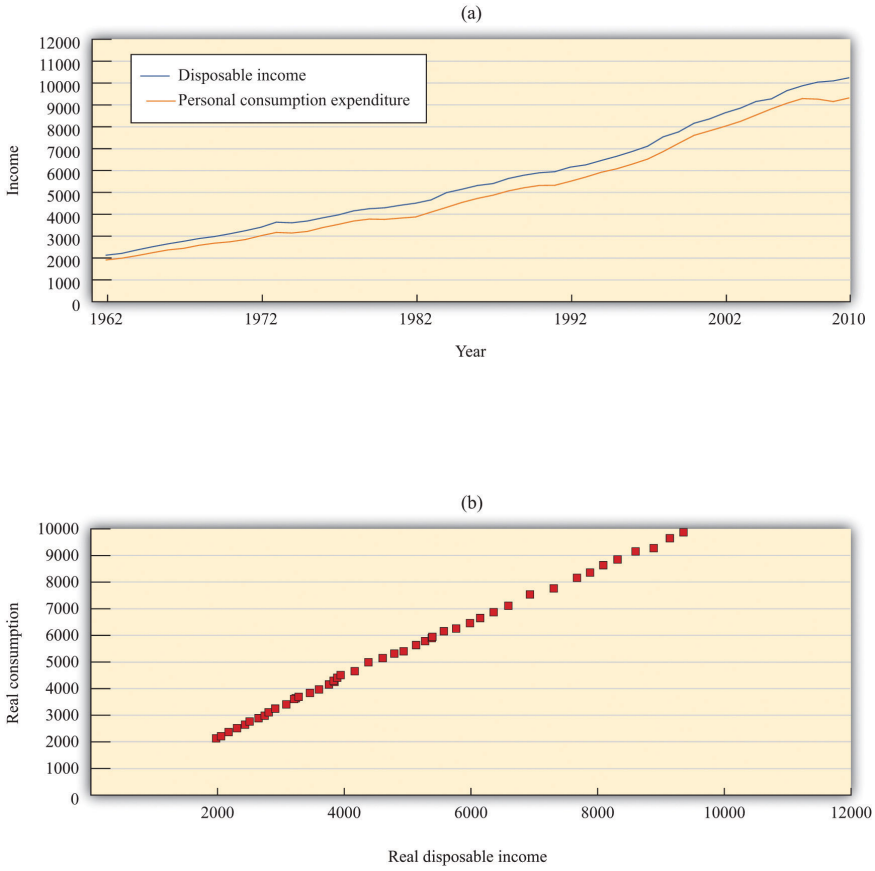


Figure 27.3.4: Consumption and Income. The charts show consumption and personal disposable income (in billions of year 2005 dollars) from 1962 to 2010. Consumption and disposable income grew substantially over this time (a) and there is a close relationship between consumption and income (b).

Source: Economic Report of the President (Washington, DC: GPO, 2011), table B-31, accessed September 20, 2011, www.gpoaccess.gov/eop/tables11.html.

Toolkit: Section 31.8

You can review how to correct for inflation in the toolkit.

The first thing we see in Figure 27.3.4 is that both consumption and disposable income grew substantially over the 1962–2010 period. This should come as no surprise. We know that the US economy grew over this period, so we would expect that disposable income and consumption would also grow. Figure 27.3.4 reveals that, as a consequence, there is a close relationship between consumption and income, and consumption expenditures are, on average, about 91 percent of disposable income. Although Figure 27.3.4 looks something like a consumption function, we should not take this relationship as strong evidence for our theory because it is primarily caused by the fact that both variables grew over time.

Consumption Response to the Kennedy Tax Cut

Now we return to the Kennedy tax cut. How well does our model perform in predicting the effects of the tax changes on consumption? Superficially, this seems like an easy question. We can examine the changes in consumption and income that arose after the tax changes and see whether these changes are consistent with the model.

There is a critical difference between our theory and reality, however. When we discussed the effects of a tax cut using our theory, we implicitly held everything else constant. We presumed that there was a change in taxes *and no change in any other variable*. For example, we assumed that government spending, investment spending, and net exports all did not change. In fact, other economic variables were changing at the same time that the new tax policy went into effect; these changes could also have affected consumption and disposable income. Looking at particular tax experiments is a messy business.

Taxes were cut in February 1964, and (real) disposable income increased by $430 billion, a much larger increase than in previous time periods. Consumption expenditures increased considerably during this period. Table 27.3.2 "Consumption and Income in the 1960s (Seasonally Adjusted, Annual Rates)" summarizes the behavior of GDP, disposable income, consumption, and the average propensity to consume over the 1960–68 period. Remember that these are real variables, measured in year 2000 dollars. The average propensity to consume is calculated as consumption divided by disposable income, and the marginal propensity to consume is calculated as the change in consumption divided by the change in disposable income.

Table 27.3.2 Consumption and Income in the 1960s (Seasonally Adjusted, Annual Rates)

Year	Real GDP ($)	Disposable Income ($)	Consumption ($)	APC	MPC
1960	2,501.8	1,759.7	1,597.4	0.91	—
1961	2,560.0	1,819.2	1,630.3	0.90	0.55
1962	2,715.2	1,908.2	1,711.1	0.90	0.91
1963	2,834.0	1,979.1	1,781.6	0.90	0.99
1964	2,998.6	2,122.8	1,888.4	0.89	0.74
1965	3,191.1	2,253.3	2,007.7	0.89	0.97
1966	3,399.1	2,371.9	2,121.8	0.89	0.96
1967	3,484.6	2,475.9	2,185.0	0.88	0.61
1968	3,652.7	2,588.0	2,310.5	0.89	1.11
APC, average propensity to consume; MPC, marginal propensity to consume.					

Source: *Economic Report of the President* (Washington, DC: GPO 2004), accessed September 20, 2011, www.gpoaccess.gov/eop.

Disposable income increased as did consumption, in accordance with the predictions of our theory. As the theory predicts, the average propensity to consume decreased for most of this period. Likewise, in line with the theory, the marginal propensity to consume was less than 1 (in all years except 1968). Thus the evidence from this period is broadly consistent with the predictions that we made on the basis of our model.

Aggregate Income, Aggregate Consumption, and Aggregate Saving

The 1964 tax cut was not designed to influence consumption in isolation but rather to have an impact on the overall economy via its effect on consumption. So far, we have argued that a change in taxes leads to a change in disposable income and hence a change in consumption. Now we complete the story, noting that a change in consumption will itself affect the level of real GDP and hence have further effects on the level of disposable income.

In the case of the Kennedy tax cut of 1964, the economists advising the administration at that time had a fairly specific idea of how changes in consumption would affect the overall economy. They argued that the $10 billion tax cut would lead to an increase in GDP of about $20 billion each year. How did they create this estimate? To answer this question, we need to embed our theory of consumption in the aggregate expenditure model.

We motivated our consumption function by thinking about the behavior of an individual household. We now presume that our household is in some sense average, or representative of the entire economy, so the consumption relationship holds at an economy-wide level. Different households might actually have different consumption functions, but when we add them together, we still expect to find an aggregate relationship similar to the one we have described.

The economists of the time used a framework that was closely based on the aggregate expenditure model. When prices are sticky, the level of GDP is determined in that model by the condition that planned spending and actual spending are equal. The model tells us that the level of real GDP depends on the level of autonomous spending and the multiplier,

$$real\ GDP = multiplier \times autonomous\ spending,$$

where the multiplier is calculated as Given the level of autonomous spending in the economy and given a value for the marginal propensity to spend, we can calculate the equilibrium level of real GDP.

The marginal propensity to spend is not the same thing as the marginal propensity to consume, although they are connected. The marginal propensity to spend tells us how much *total* spending changes when GDP changes. Total spending includes not only consumption but also investment, government purchases, and net exports, so if any of these are responsive to changes in GDP, then the marginal propensity to spend is affected. Likewise, autonomous spending is not the same as autonomous consumption. Autonomous spending is the sum of autonomous consumption, autonomous investment, autonomous government purchases, and autonomous net exports. Finally, the marginal propensity to consume measures how consumption responds to changes in disposable income, not GDP.

Toolkit: Section 31.30

You can review the aggregate expenditure model and the multiplier in the toolkit.

In our analysis here, we continue to focus on consumption and suppose that the other components of spending—government spending, investment, and net exports—are exogenous. That is, these variables are all unaffected by changes in income and so are all included in autonomous spending. In addition, we presume that the amount that the government spends is not affected by the amount that it receives in tax revenue.

To find out the effects on the economy of a change in income taxes, we take the equation for real GDP and write it in terms of changes:

$$change\ in\ real\ GDP = multiplier \times change\ in\ autonomous\ spending.$$

This equation tells us we need two pieces of information to work out the effect of a tax change:

1. The marginal propensity to spend because this allows us to calculate the multiplier
2. The effect of a tax change on autonomous spending

Let us think about the marginal propensity to spend first. We want to know the answer to the following question: if GDP changes by some amount (say, $100), what will happen to spending? There are three pieces to the answer.

1. A change in GDP leads to a change in personal income. Remember from the circular flow of income that GDP measures production, income, and expenditure in the economy. Firms receive income when they sell their products. Most of that income finds its way into the hands of households in the form of wage and salary payments or dividend payments. Firms hold onto some of the income that they generate, however, to replace worn-out capital goods and finance new investments. In the early 1960s, personal income was about 78 percent of GDP. So if GDP increased by $100, we would expect personal income to increase by about $78.
2. A change in personal income leads in turn to a change in disposable income. As we explained at length, personal income is taxed, so disposable income is less than personal income. Since we are considering the effects of a *change* in taxes, we need an estimate of the marginal tax rate facing consumers. We know from Figure 27.3.1 that this varied across individuals, but researchers have estimated that, for the economy as a whole, the marginal tax rate in 1964 was about 22 percent.Robert J. Barro and Chaipat Sasakahu provide estimates of the "average marginal tax rate." Barro and Sasakahu, "Measuring the Average Marginal Tax Rate from the Individual Income Tax" (NBER Working Paper No. 1060 [Reprint No. r0487], June 1984), http://www.nber.org/papers/w1060. To put it another way, households would keep about 78 percent (= 100 percent – 22 percent) of their personal income. So if personal income increased by $78, disposable income would increase by about $61 (= 0.78 × $78). (It is a meaningless coincidence that these two numbers are both 78 percent.)
3. Finally, a change in disposable income leads to a change in consumption. According to the 1964 *Economic Report of the President*, the CEA thought that the marginal propensity to consume was about 0.93. So if disposable income increased by $61, we would expect consumption to increase by about $57 (= 0.93 × $61).

Putting these three together, therefore, we see that an increase in GDP of $100 causes consumption to increase by $57. The marginal propensity to spend in this economy was equal to about 57 percent. It follows that the CEA thought that the multiplier was equal to about 2.3 because

Now let us think about the change in autonomous spending. We have said that taxes were cut by about $10 billion. We expect that most of this tax cut ended up in the hands of consumers. Based on the marginal propensity to consume of 0.93, we would therefore

expect there to be an increase of about $9.3 billion in autonomous consumption,

$$change\ in\ autonomous\ spending = \$9.3 billion.$$

Putting these two results together, we find that our prediction for the change in GDP as a result of the tax cut is

$$change\ in\ real\ GDP = multiplier \times change\ in\ autonomous\ spending = 2.3 \times \$9.3 billion = \$21.4 billion.$$

Our answer is not exactly equal to the $20 billion predicted by the CEA, but it is very close. As you might expect, the CEA was working with a more complicated model than the one we have explained here, and, as a result, they came up with a slightly smaller number for the multiplier.

A Word of Warning

All our analysis so far has ignored the fact that, through the price adjustment equation, increased real GDP causes the price level to rise. This increase in prices serves to choke off some of the effects of the increase in spending. In effect, we have ignored the supply side of the economy. It is not that the Kennedy-Johnson administration economists were naïve about the supply side, but they thought the demand side movements were much more relevant for short-run policymaking purposes. More recent economic experience has convinced economists that we ignore the supply side of the economy at our peril. Modern macroeconomists would be careful to augment this story with a discussion of price adjustment.

Toolkit: Section 31.31

You can review price adjustment in the toolkit.

Key Takeaways

1. Beginning in the early 1960s, growth of real GDP began to slow. This provided the basis for the tax cut of 1964.
2. The CEA economists used the aggregate expenditure model as the basis for their analysis of the effects of the tax cut.
3. In response to the tax cut, consumption and real GDP both increased. This fits with the prediction of the aggregate expenditure model.

checking you understanding

1. Is the marginal propensity to consume independent of whether an income increase is viewed as temporary or permanent?
2. If autonomous consumption is positive, is the level of savings at zero disposable income positive or negative?

This page titled 27.3: The Kennedy Tax Cut of 1964 is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

27.4: Income Taxes and Saving 27.4: Income Taxes and SavingEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. When income taxes are cut, what happens to private savings?
2. When income taxes are cut, what happens to national savings?

Look back at Figure 27.4.2. We explained that there are three channels through which income taxes affect the economy. In 27.3 Section "The Kennedy Tax Cut of 1964", we discussed the first of these in some depth: a cut in income taxes can stimulate consumption and increase aggregate spending. Figure 27.4.2 reveals that taxes can also affect **potential output**, both through their influence on saving (and hence capital accumulation) and through their effect on labor supply. We turn next to the savings channel.

Tax Cuts and Private Saving

We have already conducted most of the analysis we need to examine the effects of tax cuts on saving. We know that a tax cut increases disposable income. Our theory of consumption smoothing tells us that households will respond by increasing consumption and savings. Specifically, we predict that a dollar's worth of tax cuts will cause saving to increase by $(1 - marginalpropensitytoconsume)$.

It is tempting to conclude that tax cuts therefore will lead both to higher consumption, increasing output now, and to higher saving, increasing output in the future. Such an argument is not right because it looks only at saving by households. We also need to look at the effect of the tax cut on the government surplus or deficit.

Tax Cuts and National Saving

If the government is spending more than it receives in tax revenues, then it is running a deficit. Conversely, if it is spending less than it receives in tax revenues, it is running a surplus. National savings is the combined savings of the government and the private sector. If the government is running a deficit,

$$national\ savings = private\ savings - government\ deficit,$$

and if the government is running a surplus,

$$national\ savings = private\ savings + government\ surplus.$$

These are just two different ways of saying the same thing because, by definition, the government surplus equals minus the government deficit.

What happens if the government cuts taxes? If there are no associated changes in government spending, then tax cuts translate dollar for dollar into the government budget. One million dollars worth of tax cuts will increase the deficit (or decrease the surplus) by exactly $1 million. So even though a tax cut of a dollar increases private savings by $\$(1 - marginalpropensitytoconsume)$, it costs the government $1. The net effect (to begin with) is to *reduce* national savings by an amount equal to the marginal propensity to consume.

If the tax cut succeeds in increasing income, there is additional savings resulting from the multiplier process. Still, we expect the overall effect is a decrease in national savings. For example, consider the Kennedy tax cut again. Taxes were cut by $10 billion. The resulting change in income was roughly $20 billion. With the marginal propensity to save equal to 0.07, the offsetting increase in private savings would have been about $1.4 billion. Evidently, the result was a large decrease in national savings.

Here we see one of the biggest problems with tax cuts. They are attractive in the short run because they stimulate aggregate demand and increase output. They are also attractive politically, for obvious reasons. Unfortunately, they have the adverse long-run consequence of reducing national savings. When national savings decreases, the economy does not build up its capital stock so quickly, so future living standards are lower than they would otherwise be.

Key Takeaways

1. Since the marginal propensity to consume is less than 1, a tax cut will lead to a household to consume more and save more.
2. National savings, the sum of public and private savings, will generally decrease when there is a tax cut.

Exercises

1. If the marginal propensity to consume from a tax cut is zero, what will happen to national savings when taxes are cut?
2. If income taxes increase, what happens to the incentive of an entrepreneur to start a new business?

This page titled 27.4: Income Taxes and Saving is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

27.5: The Reagan Tax Cut 27.5: The Reagan Tax CutEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What was the state of the economy at the time of the Reagan tax cut?
2. What framework was used for analyzing the effects of this tax cut?
3. What were the effects of the tax cut?

When Ronald Reagan was elected US president in 1980, the US economy was not in very good shape. The 1970s had been a very difficult time for economies throughout the world. The oil-producing nations of the world, acting as a cartel, had increased oil prices substantially, and, as a result, energy costs had increased. These energy prices triggered a severe recession in the mid 1970s and a smaller recession in the late 1970s. Figure 27.5.1 shows the US real gross domestic product (GDP) for this period. As well as recessions, the United States was suffering from inflation that was very high by historical standards: prices were increasing by more than 10 percent a year.

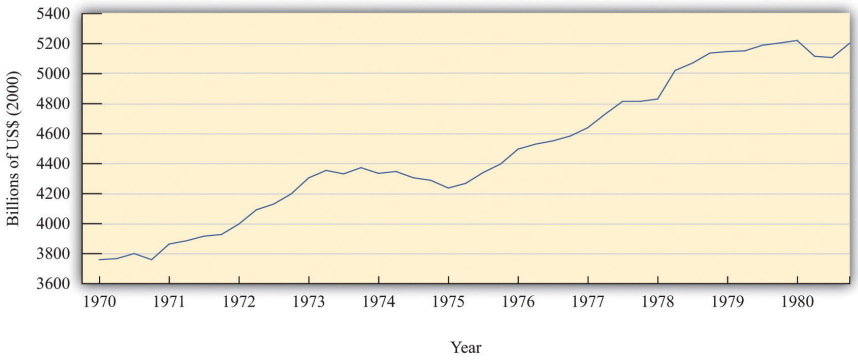


Figure 27.5.1: Real GDP in the 1970s. The figure shows real GDP in the 1970s. There was a protracted recession in the mid-1970s and a smaller recession toward the end of the decade.

Source: Bureau of Economic Analysis.

President Reagan and his economic advisors argued that high taxes were one of the causes of the relatively poor performance of the US economy. In particular, they claimed that taxes on income were deterring people from working as hard as they would otherwise. Unlike President Kennedy's advisors, who had argued that income tax cuts would increase real GDP by stimulating aggregate expenditure, Reagan's advisors said that tax cuts would increase potential output. Proponents of this economic view became known as supply siders because their focus was on the production of goods and services rather than the amount of spending on goods and services.

After his inauguration, President Reagan pushed hard for changes in the tax code, and Congress enacted the Economic Recovery Tax Act (ERTA) in 1981. This law reduced tax rates substantially: Figure 27.9 shows marginal and average tax rates for 1982, 1983, and 1984. The marginal tax rates are shown in part (a) in Figure 27.9: marginal rates decreased significantly for taxable income up to about $80,000.In contrast to Figure 27.5.1, no tax was payable until taxable income was $2,300. This is because the definition of taxable income at that time included the exemption. As a consequence, average tax rates also decreased significantly between 1982 and 1984 (part (b) in Figure 27.9).

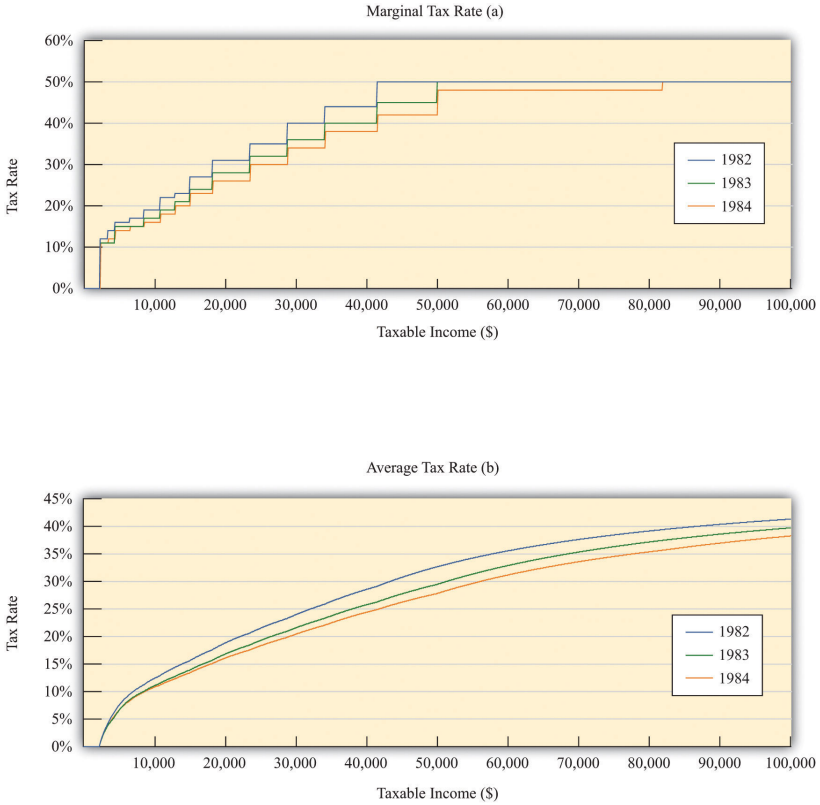


Figure 27.5.2: Marginal and Average Tax Rates, 1982 to 1984. The figure shows marginal (a) and average (b) tax rates from 1982 to 1984, the period of the Reagan tax cuts. Both marginal and average rates decreased substantially.

Source: Department of the Treasury, IRS 1987, "Tax Rates and Tables for Prior Years" Rev 9-87

The main mechanism that the supply siders proposed was that lower income taxes would increase the incentive to work. To analyze this claim, we need to investigate how the decision to supply labor depends on income taxes. As with our analysis of consumption, we look at labor supply by thinking about the behavior of a single household. We then suppose that the household can be taken as representative of the entire economy.

Labor Supply

Each individual faces a time constraint: there are only 24 hours in the day, which must be divided between working hours and leisure hours. An individual's **time budget constraint** says that, on a daily basis,

$$leisure\ hours + working\ hours = 24 hours.$$

The labor supply decision is equivalently the decision about how much leisure time to enjoy. This decision is based on the trade-off between enjoying leisure and working to purchase consumption goods. People like having leisure time, and they prefer more leisure to less. Leisure can be thought of as a "good," just like chocolate or blue jeans or cans of Coca-Cola. People sacrifice leisure, working instead, because the money they earn allows them to purchase goods and services.

To see this, we first rewrite the time budget constraint in money terms. The value of an hour of time is given by the nominal wage. Multiplying through the time budget constraint by the nominal wage gives us a budget constraint in dollars rather than hours:

$$(leisure\ hours \times nominal\ wage) + nominal\ wage\ income = 24 \times nominal\ wage.$$

The second term on the left-hand side is "nominal wage income" since that is equal to the number of hours worked times the hourly wage.

Because wage income is used to buy consumption goods, we replace it by total nominal spending on consumption, which equals the price level times the quantity of consumption goods purchased:

$$(leisure\ hours \times nominal\ wage) + (pricelevel \times consumption) = 24 \times nominal\ wage.$$

This is the budget constraint faced by an individual choosing between consuming leisure and consumption. Think of it as follows: it is as if the individual first sells all her labor at the going wage, yielding the income on the right-hand side. With this income, she then "buys" leisure and consumption goods. The price of an hour of leisure is just the wage rate, and the price of a unit of consumption goods is the price level. Finally, if we divide this equation through by the price level, we see that it is the **real wage** (the wage divided by the price level) that appears in the budget constraint:

$$leisure\ hours \times real\ wage + consumption = 24 \times real\ wage.$$

It is the real wage, not the nominal wage, that matters for the labor supply decision.

> **Toolkit: Section 31.3**
>
> You can review the labor market in the toolkit.

Changes in the Real Wage

What happens if there is an increase in the real wage? There are two effects:

1. There is a **substitution effect**. An increase in the real wage means that leisure has become relatively more expensive. You have to give up more consumption goods to get an hour of leisure time. If leisure becomes more expensive, we would expect the household to "buy" fewer hours of leisure and more consumption goods—that is, to *substitute* from leisure to consumption. This effect predicts that the quantity of labor supplied will increase.
2. There is an **income effect**. An increase in the real wage makes the individual richer—remember that we can think of income as equaling 24 × the real wage. In response to higher income, we expect to see the household increase its consumption of goods and services and also *increase* its consumption of leisure. This effect predicts that the quantity of labor supplied will decrease.

Putting these predictions together, we must conclude that we do not know what will happen to the quantity of labor supplied when the real wage increases. On the one hand, higher real wages make it attractive to work more since you can get more goods and services for each hour of time that you give up (the substitution effect). On the other hand, you can get the same amount of consumption goods with less effort, which makes it attractive to work less (the income effect). If the substitution effect is stronger, the labor supply curve has the standard shape: it slopes upward, as in Figure 27.5.3.

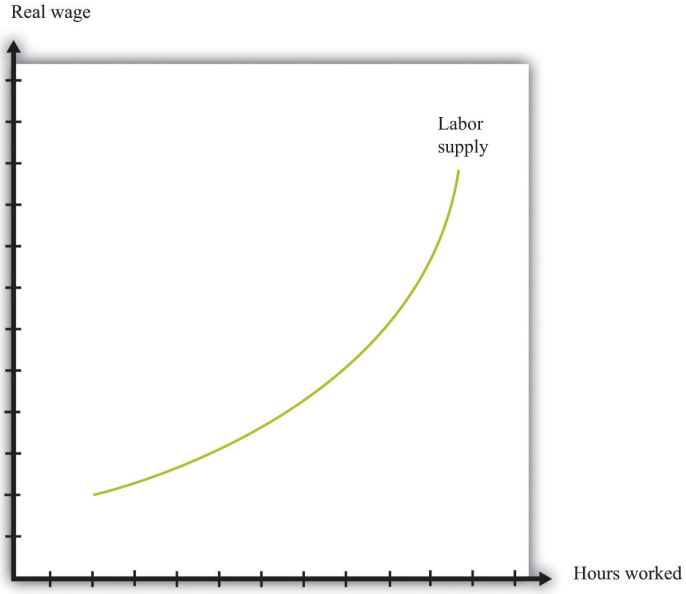


Figure 27.5.3: Labor Supply. The response of the quantity of labor supplied to the real wage depends on both an income effect and a substitution effect. When the substitution effect is larger than the income effect, the supply curve has the "normal" upward-sloping shape.

In the end, the shape of the labor supply curve is an empirical question; we can answer it only by going to the data. And as you might be able to guess, it turns out to be a difficult question to answer, once we start dealing with the complexities of different kinds of labor. The view of most economists who have studied labor supply is that higher real wages do lead to a greater quantity of

labor supplied, but the effect is not very strong. The income effect almost cancels out the substitution effect. This means that the labor supply curve slopes upward but is quite steep.

The Effect of the Reagan Tax Cuts on the Supply of Labor

Suppose an individual knows the nominal wage but also knows that she is going to be taxed on any income that she earns at the going income tax rate. The wage rate that matters for her decision is the after-tax real wage. Her real disposable income is

All our discussion of labor supply continues to hold in this case, except that we need to replace the real wage with the after-tax real wage since it is the after-tax wage that matters to the individual.

Figure 27.5.4 shows the effect of a cut in taxes. If the labor supply curve slopes upward, the tax cut leads to an increase in the quantity of labor supplied. And if labor supply increases, then potential output also increases. In other words, one effect of tax cuts is to induce people to work harder and produce more real GDP. To keep things simple, Figure 27.5.4 is drawn supposing that there is no change in the equilibrium real wage as a result of the tax cut. In fact, we would expect the real wage to decrease somewhat as well. Buyers of labor as well as sellers of labor would benefit from the tax cut. Indeed, it is this decrease in the real wage that induces firms to purchase the extra labor that individuals wish to supply. (If we included this in our picture, then the after-tax real wage would still increase but by less than shown in the figure.)

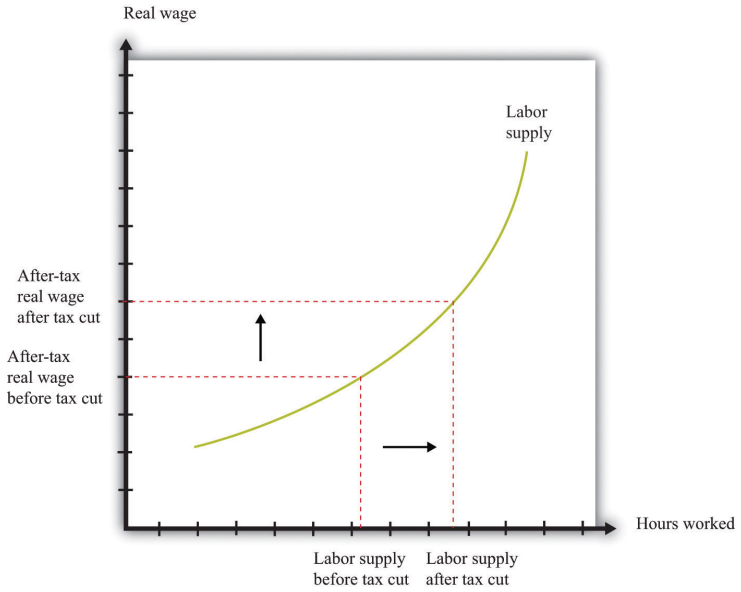


Figure 27.5.4: Labor Supply Response to Tax Cut. The wage that matters for labor supply decisions is the after-tax real wage. If income taxes are cut, and the real wage is unchanged, then households will supply more labor.

The Laffer Curve

Supply-side economics was controversial and generated a great deal of debate back in the 1980s and since. Yet the argument that we have just presented is not really controversial at all. Almost all economists agreed that as a matter of theory, cuts in taxes could lead to increases in the quantity of labor supplied. The disagreements concerned the magnitude of the effect.

Some proponents of supply-side economics made a much stronger claim. They said that the positive effects on labor supply could be so large that total tax revenues would increase, not decrease. They argued that even though the government would get less tax revenue on each dollar earned, people would work so much harder and generate so much more taxable income that the government would end up with more revenue than before.

This argument was encapsulated in the so-called Laffer curve. Economist Arthur Laffer asked what would happen if you graphed tax revenues as a function of the tax rate. Obviously (he observed) if the tax rate is zero, then tax revenues must be zero. And, Laffer argued, if the tax rate were 100 percent, so the government took every penny you earned, then no one would have an incentive to work at all, and the quantity of labor supplied would drop down to zero. Once again, income tax revenues would be zero. In between, tax revenues are positive.

Figure 27.5.5 shows an example of a Laffer curve. There is some tax rate that will lead to the maximum possible revenue for the government. This itself is not that interesting: the goal of the government is not to raise as much tax revenue as possible. But if the

tax rate lies to the right of that point, then—as the picture shows—a cut in taxes will increase tax revenues.

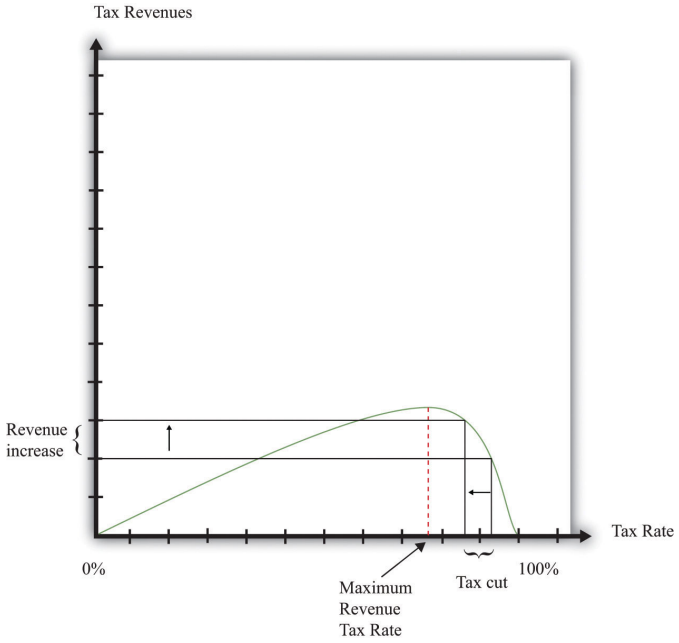


Figure 27.5.5: Laffer Curve. The Laffer curve says that it is possible for a reduction in the tax rate to lead to an increase in tax revenues. Although this is a theoretical possibility at very high tax rates, most economists view the Laffer curve as a theoretical curiosity with limited applicability to real economies.

Just as almost all economists agreed that there would be *some* supply-side effects of income tax cuts, almost all economists agreed that the Laffer curve argument was inapplicable to the US economy (or indeed any other economy). The evidence indicated that the effects of tax cuts on hours worked were likely to be relatively small. Almost no economists actually believed that the economy was on the wrong side of the Laffer curve, where tax cuts could pay for themselves.

Unfortunately, the Laffer curve argument was politically appealing, even though it was not supported by economic evidence. Buoyed by this argument, President Reagan oversaw both tax cuts and big increases in government spending. As a result, the US government ran large budget deficits. Following on from the ERTA, President Reagan and President George H. W. Bush after him were both forced to increase taxes to bring the budget back under control.The economic history of the United States in the 1980s was quite complex. Because this chapter concerns income taxes, we have considered only one of the policy changes of the Reagan administration. Other changes in tax policy were designed to promote savings. We have not discussed other aspects of President Reagan's fiscal policy (there were large increases in government purchases), the tight monetary policy pursued by the Federal Reserve, or the behavior of interest rates and exchange rates. All these are matters for other chapters.

Key Takeaways

1. Prior to the Reagan tax cut, the US economy was experiencing both low growth in real GDP and high inflation.
2. Reagan's economic advisors stressed the effects of taxes on the supply side of the economy, and in particular the incentive effects of taxes on labor supply and investment.
3. The Reagan tax cuts led to considerably higher deficits in the United States.

Exercises

1. What matters for labor supply decisions—the marginal tax rate or the average tax rate?
2. According to the Laffer curve, does a tax cut always increase tax revenues?

This page titled 27.5: The Reagan Tax Cut is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

27.6: End-of-Chapter Material 27.6: End-of-Chapter MaterialEdit Update titlecancel

In Conclusion

Our goal in this chapter was to understand the effects of tax changes on aggregate consumption and aggregate output. A tax cut puts more income in the hands of households, and thus consumption increases. The increase in consumption in turn leads to an expansion in the overall level of economic activity. The framework does a good job of describing and explaining actual economic outcomes during the Kennedy tax cut. We can thus have some faith that our basic framework is reasonably sound. Having said that, it is a very simple model that does have some deficiencies, most notably its neglect of the supply side of the economy.

Income tax cuts also decrease overall national saving. Income tax cuts increase household disposable income and lead to increased saving by households (as well as increased consumption). At the same time, however, income tax cuts mean that the government is saving less (or borrowing more). The net effect is to decrease national saving. The theory of economic growth tells us that reduced saving has the effect of decreasing future standards of living.

We then examined the Reagan tax cuts of the 1980s. These tax cuts were aimed at stimulating employment and output by encouraging people to work more. The belief that tax cuts lead to an increase in the quantity of labor supplied is consistent with basic microeconomic principles, but there is disagreement about the likely size of the effect.

Although we cast our discussion of the effects of taxes on spending using the tax cuts of the Kennedy and Reagan administrations, the lesson is more general. It is common for the United States and other countries to use variations in income tax rates as a tool of intervention. We highlighted several effects of such interventions. Income tax changes alter the level of household disposable income and thus influence consumption expenditures; they affect saving and capital accumulation; and they affect labor supply. This policy tool therefore gives the government considerable influence on the aggregate economy.

Indeed, when the crisis of 2008 hit the world's economies, many countries responded by implementing expansionary fiscal policies, including cuts in taxes. Australia, the United Kingdom, Singapore, Austria, and Brazil are just a few of the countries who cut taxes in response to the crisis.

We used the Kennedy tax cut to illustrate demand-side effects and the Reagan tax cut to illustrate supply-side effects because those were the channels emphasized by the economic advisors at the time. Just about every change in the income tax code, however, has effects on consumption, saving, and labor supply. Every change in the code has short-run effects and long-run effects, and, as we have seen, these effects can be contradictory. Thus whenever you hear or read about proposed changes in taxes, you should try to remember that all these different stories will be in operation. The politicians and pundits who are supporting or opposing the change will typically talk about only one of them, depending on the spin they wish to convey. The analysis of this chapter should help you always see the bigger picture.

Finally, remember that tax changes will typically have major effects on the distribution of income. There are winners and losers from every change in the tax code. This, above all, is why changes in the tax code are an endless source of political debate.

Key Links

- Council of Economic Advisors: http://www.whitehouse.gov/administration/eop/cea
- Christina Romer and David Romer narrative of tax changes in the United States: http://elsa.berkeley.edu/~dromer/papers/nadraft609.pdf
- Internal Revenue Service
  - History: www.irs.gov/irs/article/0,,id=149200,00.html
  - Forms: http://www.irs.gov/formspubs/index.html?portlet=3
  - Withholding Tax Calculator: http://www.irs.gov/individuals/article/0,,id=96196,00.html
- NBER Tax Simulator: http://www.nber.org/~taxsim/taxsim-calc9/index.html

exercises

1. Suppose that your income level is $55,000. Using the tax table for 2010 (Table 27.2.1 "Revised 2010 Tax Rate Schedules"), what are your marginal and average tax rates?
2. Suppose that taxes paid were equal to a constant plus a tax rate times income. Devise a tax schedule such that the marginal tax rate is 25 percent and the average tax at $10,000 is $2,000. What is the constant?

3. In times of inflation, the nominal income of households increases. What happens over time to their marginal and average tax rates?
4. Our tax function has a constant marginal tax rate at all levels of income. Explain why this means that the average tax rate is also constant. Is the average tax rate higher, lower, or equal to the marginal tax rate in this case?
5. We noted earlier that the average tax rate for someone earning $100,000 was 67 percent in 1963. However, there has been considerable inflation between 1963 and the present day. What is the equivalent in current dollars of an income of $100,000 in 1963? (Look at the toolkit if you need a reminder of how to convert from nominal to real variables.)
6. Suppose that autonomous consumption is 600 and the marginal propensity to consume is 0.9. Graph the consumption and savings functions first with disposable personal income on the horizontal axis and then with GDP on the axis. If there is a change in taxes, how would that affect these graphs?
7. What is the difference between the marginal propensity to consume and the marginal propensity to spend?
8. Why is a temporary tax cut likely to have a smaller impact on real GDP than a permanent tax cut?
9. Using the logic of consumption smoothing, if a tax cut from 10 years ago will expire next year, what will a household do now in anticipation of the coming tax change?
10. If labor supply is known to be relatively insensitive to changes in the real wage, what does this imply about the argument that cuts in tax rates can lead to revenue increases?

Economics Detective

1. Pick some country other than the United States. Can you find the income tax rates for that country? How do they compare with those in the United States?
2. Go to the IRS web page. Suppose that you are a member of a married household with total household income of $55,000. What are your marginal and average tax rates? Compare these to the tax rates on individuals. Which group faces the higher marginal income tax rate? What effects might this have on their behavior?
3. In the summer of 2010, the George W. Bush tax cuts were about to expire. What would the change in tax rates be if the tax cuts had been allowed to expire?
4. Go to the Bureau of Economic Analysis website (http://www.bea.gov). Click on the link “Personal Income and Outlays” and find out what has happened recently to personal income and disposable income. Have they been increasing or decreasing?

Spreadsheet Exercises

1. Using a spreadsheet, enter the data for disposable income and consumption from Table 27.3.2 "Consumption and Income in the 1960s (Seasonally Adjusted, Annual Rates)". Now enter a formula to calculate the average propensity to consume and another to calculate the marginal propensity to consume. Check that your answers are the same as in Table 27.3.2 "Consumption and Income in the 1960s (Seasonally Adjusted, Annual Rates)".
2. Suppose autonomous consumption is 6000 and the marginal propensity to consume is 0.9. Furthermore, suppose the tax rate is 30 percent. Create a spreadsheet where the first column is income, ranging from $0 to $100,000, by increments of $1,000. Create columns showing the taxes paid at each income level, the level of disposable income at each income level, and consumption at each income level. Graph the relationship between consumption and income. What is the slope of the line? Experiment with changing the marginal propensity to consume and the tax rate. Explain how changing these parameters affect the relationship between consumption and income.

This page titled 27.6: End-of-Chapter Material is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

28: Social Security

28.1: Death and Taxes
28.2: Individual and Government Perspectives on Social Security
28.3: A Model of Consumption
28.4: Social Security in Crisis?
28.5: The Benefits and Costs of a Social Security System
28.6: Social Security in the Real World
28.7: End-of-Chapter Material

This page titled 28: Social Security is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

28.1: Death and Taxes 28.1: Death and TaxesEdit Update titlecancel

Benjamin Franklin famously wrote that “in this world nothing is certain but death and taxes.” The current chapter is about both.

If you are like most readers of this book, you are at the very beginning of your working life, and you have probably given little thought to your retirement. In the early years of work, you might be asked to make some decisions regarding a company pension plan, but it is still unlikely that you will spend much time thinking about how you will live when your working life is over. This is normal; none of us is very good at imagining at the age of 20 what our life will be like when we are 70. (Even at the age of 30, or 40, or 50, it is hard to imagine life at age 70.)

One message of this chapter, though, is that even though it is hard to think that far ahead, it is also smart to try to do so. From the very beginning of your working life, you will be making decisions that affect your life in retirement. And from the very beginning of your working life, those decisions will—or should—be influenced by something called the Social Security system.

Social Security was born in the Great Depression. Many people suffered tremendous economic hardship in the 1930s. As part of President Franklin D. Roosevelt’s New Deal in the 1930s, the US government established several systems to alleviate such hardships. Social Security—one of the most important—was designed to provide financial assistance to the elderly. More than 170 other countries, big and small, rich and poor, also have social security systems. To take a few random examples, you will find social security in operation in Mexico, France, the United Kingdom, Kiribati, Laos, Azerbaijan, Chile, Andorra, Burkina Faso, Egypt, Cyprus, Paraguay, and Slovenia.

The Social Security system will give you money when you are older, but it takes money from you when you are working. So even if it is hard to think about the effect that Social Security will have on your income in the distant future, it is very easy to see the effect it has when you are working. Workers are required to make Social Security contributions—one of the many kinds of tax that we all pay—with the promise that they will receive reimbursement from the system when they are older. The state of the US Social Security system is therefore something that you should think about long before you receive payouts. Decisions about your personal saving and consumption right now are, or at least should be, directly influenced by your current tax contributions and expectations of your future Social Security payouts.

Opinion polls reveal that Social Security is one of the most well-supported government programs in the United States. Yet the casual reader of the newspapers could be forgiven for thinking that the system is perpetually in crisis. In the 1980s, for example, there was discussion of serious difficulties with the funding of Social Security. A commission headed by Alan Greenspan (who later became chairman of the Federal Reserve Bank) identified problems with the system and recommended a large number of changes, including some increases in Social Security tax rates. These reforms were supposed to ensure the solvency of Social Security well into the future. Yet, a few decades later, proposals for major reforms of Social Security are back under discussion. The exact form that Social Security will take in the coming decades is an open question that will continue to play a major role in political debate.

We explain the details of the system more carefully later in this chapter, but the basic idea is the following. The government taxes current workers and uses those revenues to pay retired workers. When the system was originally set up, the idea was that payments to retired people in a given year would be (approximately) funded by taxes on those working during that year, so the system would be roughly in balance. For many years, this “pay-as-you-go” structure worked fine. In some years, payments to workers were larger than tax receipts, and in some years, they were smaller, but on the whole there was an approximate match between payments and receipts.

In the 1980s, policymakers first began to pay serious attention to the fact that there was a problem with the pay-as-you-go structure. Demographic changes mean that the system is not balanced in the very long run. The number of retirees relative to the number of workers will increase substantially over the next two decades, and without changes, the time will come when tax revenues will no longer be sufficient to match the obligations of the system.

This is not a looming crisis. The best estimates suggest that the system will no longer be able to pay the full amount of benefits near the middle of the century, although there is disagreement on the exact date. The most recent Social Security Trustees report (http://www.ssa.gov/OACT/TRSUM/index.html) predicted this date as 2036, whereas in 2005 the Congressional Budget Office (www.cbo.gov/ftpdocs/60xx/doc6074/02-09-Social-Security.pdf) gave a date of about 2054.See the discussion at Charles P. Blahous III and Robert D. Reischauer, “Status of the Social Security and Medicare Programs,” Social Security Administration, 2011, accessed July 20, 2011, http://www.ssa.gov/OACT/TRSUM/index.html; Douglas Holtz-Eakin, “CBO Testimony,” Congressional

Budget Office, February 9, 2005, accessed July 20, 2011, www.cbo.gov/ftpdocs/60xx/doc6074/02-09-Social-Security.pdf. Of course, changes will almost certainly be made well before this crisis point. But what form should those changes take?

How should we reform Social Security?

This question matters to every single one of us. As workers, we all pay into the Social Security system, and we all anticipate receiving benefits when we are retired. The current discussion will determine both the level of taxes we pay now and the benefits we will receive in the future.

The average person could be forgiven for thinking that the debate over Social Security is complicated, arcane, and impossible to understand without an immense amount of study. In fact, the basics of the system are quite straightforward, and the most important elements of the discussion can be understood using very little economics. In this chapter, we demystify the arguments about Social Security. This will make it easier for you to understand why you pay Social Security contributions, what you can expect to get in the future, and whether the politicians and talking heads are making any sense when they discuss various reforms.

Road Map

At the heart of the economic analysis of Social Security is a very straightforward idea: “forced saving.” Individuals are required to give up some of their income now—income that they could, if desired, have used for current consumption—and, in return, they are promised income in the future. Understanding Social Security from the individual perspective means understanding the impact of this forced saving on individual choices.

Meanwhile, we also need to understand how Social Security looks from a government perspective. Social Security contributions are a source of government revenue, and Social Security payments are an example of a government transfer. These revenues and payments enter into the government’s budget constraint.

From the perspective of an individual, there is a disconnect in time between taxes and payments. Individuals pay taxes during their working years and receive transfers during their retirement years. But from the perspective of the government, taxes and payments take place at the same time. In any given year there are some individuals who are working and paying taxes, and the money they pay into the system is paid right back out to others who are in retirement.

To address questions about reforming the Social Security program, we therefore need to understand (1) the structure of the program and (2) how it interacts with individual choices about consumption and saving. We study how individuals respond to Social Security by using a model of consumption and saving that applies over an individual’s lifetime. Once we understand how individuals make these choices, we ask how Social Security affects their decisions. Then we think about how the government fits into the picture. We study these flows into and from the Social Security program using the government budget constraint, to link changes in the program with changes in taxes.This tool is used elsewhere in the book in other applications, such as Chapter 27 and Chapter 29. In the end, we are able to see how individuals’ consumption and saving decisions are influenced by their beliefs about government behavior.

This page titled 28.1: Death and Taxes is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

28.2: Individual and Government Perspectives on Social Security 28.2: Individual and Government Perspectives on Social SecurityEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. How do households respond to variations in income over their lifetime?
2. What is the government's budget constraint in a pay-as-you-go system?
3. What is the effect of Social Security on lifetime income?

We begin with the individual perspective on Social Security.

Social Security: The Individual's View

Household incomes tend to vary quite a lot, but households like their consumption to be approximately constant over time. Households therefore use their savings to smooth out the variations in their income.We also discuss consumption smoothing in Chapter 27. For the purposes of understanding Social Security—both its problems and its reforms—we need to examine this idea of consumption smoothing more rigorously. Because Social Security is a program to provide for consumption in retirement, we must carefully lay out the decisions of a household over the entire lifetime of its members. By so doing, we can determine the likely effect of a promise of a transfer in the future on behavior today. Bear in mind that these transfers may be far in the future: for a 25-year-old worker, we are thinking about money that won't be received for another 40 years or so.

As is often the case in economics, we start by looking at the simplest setup we can imagine. We do this not because we think we can answer every question with a simple model, but because we must make sure we understand the fundamentals of Social Security before we worry about the complexities. So, to keep things simple, we examine the life of a single individual—that is, think of the household as containing just one person. Hence we do not have to worry about multiple wage earners (who might be of different ages), and we do not have to worry about how to incorporate children (who grow up and leave the household) into our story. In this chapter, we use the terms *individual* and *household* interchangeably. We call our individual Carlo.

Carlo thinks about his income and consumption over his entire lifetime. Because he has the possibility of saving and borrowing, his income and consumption need not be equal in any given year. Carlo faces a **lifetime budget constraint**, however; in the end, his lifetime spending is limited by his lifetime income. The **life-cycle model of consumption** examines Carlo's decisions about how much to consume each year, given this budget constraint.

We begin with a simple numerical example. Suppose Carlo is 20 years old and very well informed about his future. He knows that he is going to work for 45 years—that is, up to age 65. He knows that, every year, he will receive income of $40,000, excluding Social Security contributions. He has to pay Social Security contributions on this income at a rate of 15 percent. Thus he knows he will pay $6,000 each year to Social Security, and his after-tax income is $\$40,000 - \$6,000 = \$34,000$. After he retires at age 65, he knows he will receive a Social Security payment of $18,000 each year until he dies, 15 years later, on his 80th birthday (of heart failure, brought on by the exertion of blowing out all those candles).

To decide on his lifetime consumption and saving patterns, Carlo needs to know what his lifetime resources are. We know that, in general, a dollar today is not worth the same amount as a dollar next year—or 60 years from now—because of interest rates and inflation. We sidestep that problem for the moment by imagining that the **real interest rate** is zero. In this case, it is legitimate to add together dollars from different time periods.

So Carlo earns $34,000 per year for each of his 45 working years and obtains $18,000 per year for his 15 retired years. His total lifetime resources are as follows:

$$lifetime\ income = income\ during\ working\ years + income\ during\ retirement\ years = (\$34,000 \times 45) + (\$18,000 \times 15) = \$1,530,000 + \$270,000 = \$1,800,000.$$

Over his life, therefore, he has $1.8 million to spend. Figure 28.2.1 shows his lifetime income. His total lifetime resources are obtained by adding together the two rectangles labeled "Lifetime Income from Working" and "Lifetime Social Security Income." The height of each rectangle gives his income, and the width of each rectangle gives the number of years for which he earns that income.

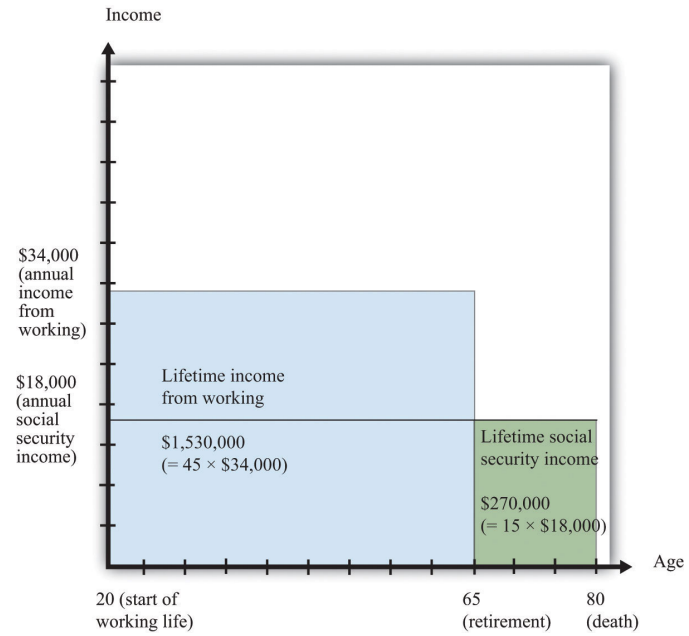


Figure 28.2.1: Lifetime Income

For 45 years, Carlo earns $34,000 per year (for a total of $1,530,000), and for 15 years, Carlo receives $18,000 per year in Social Security payments (for a total of $270,000). His total lifetime income is $1,800,000.

Carlo's lifetime budget constraint says that his lifetime consumption must equal his lifetime income,

$$lifetime\ consumption = lifetime\ income.$$

If Carlo wants to keep his consumption perfectly smooth, he will consume exactly the same amount in each of his 60 remaining years of life. In this case, his consumption each year is given by the following equation:

Figure 28.2.2 shows Carlo's consumption. The area of the rectangle is a measure of Carlo's lifetime consumption since—as before—the height of the rectangle is his consumption per year, and the width is the number of years.

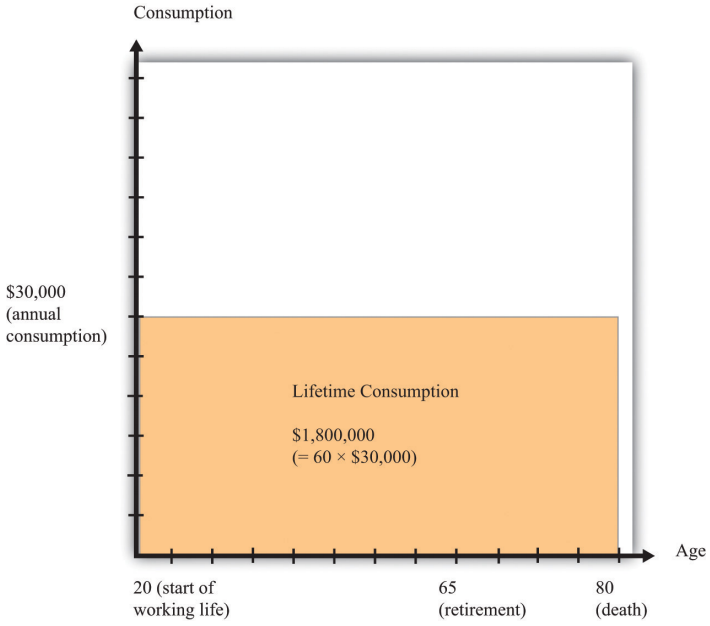


Figure 28.2.2: Lifetime Consumption

For the 60 years of his remaining life, Carlo spends $30,000 per year, making a total of $1,800,000 during his working years.

In each of his working years, Carlo earns $34,000 but consumes only $30,000. Thus he saves $4,000 every year. When he is 21 years old, he therefore has $4,000 in the bank. When he is 22 years old, he has $8,000 in the bank. By the time he retires at age 65,

he has saved $180,000 (= 45 × $4,000).

During his retirement years, Carlo starts to draw on his savings. Social Security pays him $18,000, so he needs to take an additional $12,000 from his savings to have $30,000 in consumption. At age 66, therefore, he has savings of $180,000 − $12,000 = $168,000. For each of his retirement years, his savings are reduced by a further $12,000. After his 15 years of retirement, he has reduced his savings by $12,000 × 15 = $180,000 and dies at the age of 80 with exactly zero in the bank.

Figure 28.2.3 combines Figure 28.2.1 and Figure 28.2.2 and shows Carlo's income and consumption. The difference between income and consumption in Carlo's working years is his saving. Notice the rectangles labeled "saving" and "dissaving." One way of understanding his lifetime budget constraint is that these two rectangles must be equal in area.

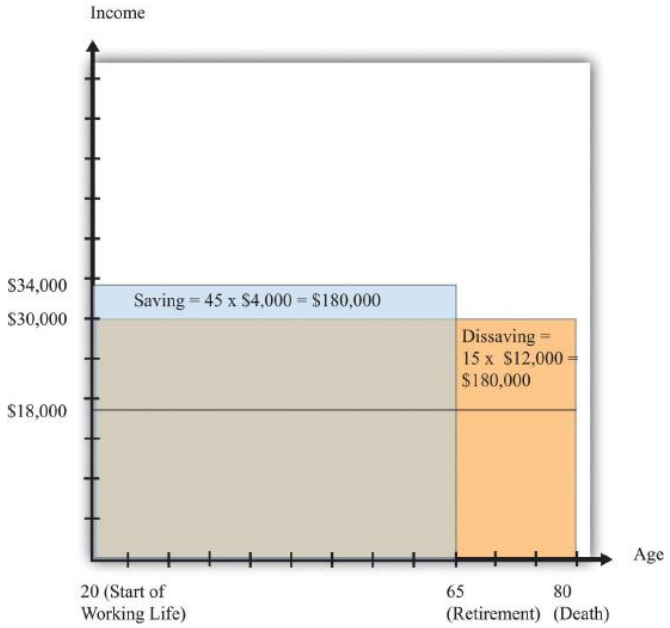


Figure 28.2.3: Lifetime Consumption and Saving. During his 45 working years, Carlo saves $4,000 per year. During his 15 retirement years, Carlo dissaves at a rate of $12,000 per year.

Figure 28.2.4 shows his wealth over his lifetime. It increases from zero to $180,000 and then decreases again to zero. The fact that he ends his life with exactly zero wealth is just another way of saying that he exactly satisfies his lifetime budget constraint.

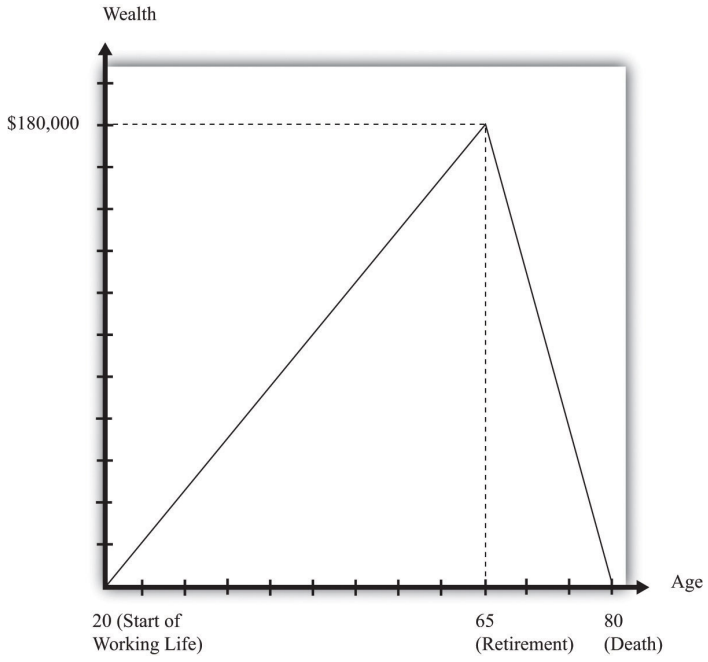


Figure 28.2.4: Lifetime Wealth Accumulation

Over his working life, Carlo builds up his stock of wealth, so he has $180,000 in his bank account when he retires. During retirement, he dissaves, exactly using up the $180,000 he accumulated while working.

Social Security: The Government's View

Now let us shift perspective and examine the Social Security system from the perspective of the government. The original intention was that Social Security would be (approximately) pay-as-you-go. Under a strict pay-as-you-go system, the inflows to the government in the form of tax revenues are exactly balanced by outflows to retired people. In any given year, in other words, the government takes money from those that are working and transfers all that money—not a cent more, not a cent less—to those who are retired.

Under this system the government does not maintain any kind of "savings accounts" for individuals: it taxes you when you work and transfers the revenues to retirees at the same time. The government promises to make payments to you after you retire, with these payments being financed by those who will then be working.

Let us pause for a moment here. We have to determine how to capture transfers across different generations in the economy in as simple a setup as possible. The easiest way to do this is to suppose that everyone in the economy is just like Carlo. That is, every working person in the economy earns $40,000 and pays $6,000 into the Social Security system. Every retired person receives a Social Security payment of $18,000 per year.

Let us further suppose that there is the same number of people of every age in the economy. In each year the same number of people is born, and—like Carlo—they all live to exactly the age of 80. Like Carlo, everyone works for 45 years (from age 20 to 65) and is retired for 15 years (from age 65 to 80). If we wanted to calibrate this roughly to the US economy (that is, make the numbers in the example a bit more realistic), we might suppose that there are 4 million people born every year. Since everyone lives to the same age, this means that there are 4 million 20-year-olds, 4 million 21-year-olds, and so forth, up to 4 million 79-year-olds. (This implies a total population of 320 million, which is close to the size of the actual US population.)

Having made these simplifications, it is a short step to realize that we might as well just suppose that there is only one person of each age. The basic structure of the economy will be the same, but the math will be much easier. (If you can prefer, though, you can multiply both sides of every equation that follows by 4 million.)

Given this demography, what do the government finances look like? Every year, the government collects $6,000 each in Social Security revenues from 45 working people, so that the total revenues are given by the following equation:

$$Social\ Security\ revenues = 45 \times \$6,000 = \$270,000.$$

Meanwhile, the government pays out $18,000 each year to 15 people:

$$Social\ Security\ payments = 15 \times \$18,000 = \$270,000.$$

You can see that we have chosen the numbers for our example such that the Social Security system is in balance: revenues equal receipts. A system like this one would indeed be pay-as-you-go.

The Effect of a Change in Social Security Benefits

Now, what would happen in this example if the government decided it wanted to increase Social Security payments by $3,000 per year? The total increase in payments would equal $45,000 since all 15 retired individuals would receive the extra $3,000. If the government is required to keep the Social Security system in balance, then it would also be obliged to increase Social Security contributions by $1,000 per worker (since there are 45 workers). How would Carlo (and everybody else like him) feel about this change?

Remember that Carlo had income before Social Security of $40,000 per year and had to pay $6,000 per year in Social Security contributions. Now he will have to pay $7,000 in contributions, so his income after tax is equal to $33,000. In his retirement years, however, Carlo will now get $21,000 a year instead of $18,000. His lifetime resources are therefore as follows:

$$lifetime\ income = (\$33,000 \times 45) + (\$21,000 \times 15) = \$1,485,000 + \$315,000 = \$1,800,000.$$

Carlo's lifetime resources are exactly the same as they were before. Of course, this means that Carlo would choose exactly the same amount of consumption as before: $30,000 a year. However, his saving behavior would be different. He would now only save $3,000 a year. At the time of retirement, he would have saved a total of $135,000. Over the remaining 15 years of his life, Carlo

would draw on his savings at the rate of $9,000 per year, which—combined with his Social Security payment of $21,000—would ensure that he had $30,000 to spend in his retirement years. His saving and dissaving are illustrated in Figure 28.2.5.

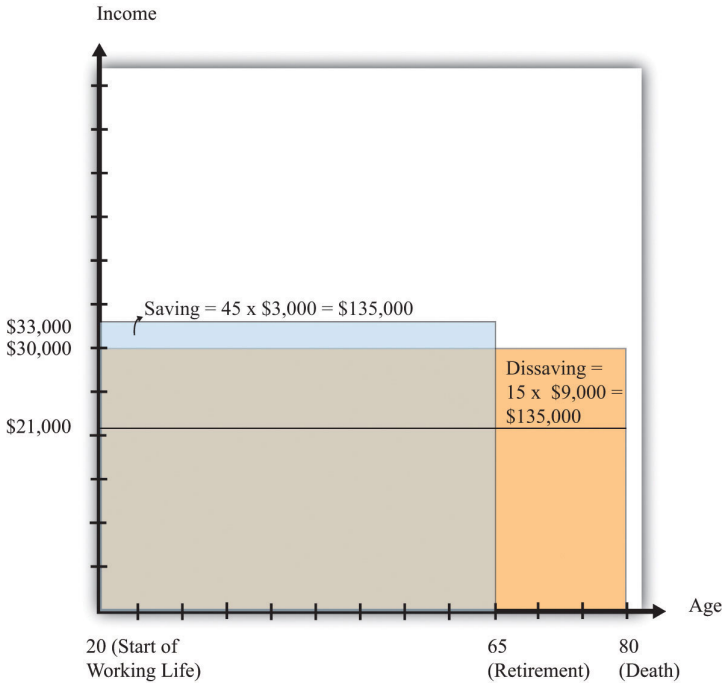


Figure 28.2.5: Lifetime Consumption and Saving

For his 45 working years, Carlo saves $3,000 a year. For his 15 retirement years, Carlo dissaves at a rate of $9,000 a year.

We have discovered a rather remarkable conclusion: *the change in the government's Social Security scheme has no effect on Carlo's lifetime resources or lifetime consumption*. From Carlo's point of view, the change means that the government is saving more on Carlo's behalf, and therefore he does not need to save so much for himself. Carlo's saving declines by exactly the same amount as the increase in Social Security taxes ($1,000) per year; likewise, his dissaving declines by exactly the same amount as the increase in Social Security payments.

Another example is even more striking. Suppose there were no Social Security system at all. Then Carlo would receive $40,000 a year for 45 years but nothing at all in his retirement years. His lifetime resources would equal

$$lifetime\ income = \$40,000 \times 45 = \$1,800,000,$$

which is again the same as before. To enjoy lifetime consumption of $30,000 a year, Carlo would save $10,000 in every working year and dissave $30,000 in every retirement year.

These numerical examples suggest an extraordinary conclusion: Social Security seems to be completely irrelevant for Carlo. No matter what the scheme looks like, Carlo has the same lifetime resources and same lifetime consumption. This is an amazing and perhaps even shocking finding. We have used some economics to analyze one of the most important government programs, one that is a source of constant scrutiny and debate (and, not incidentally, requires substantial resources for its administration). Not only have we found no reason to expect a crisis in Social Security, we have found that it is irrelevant.

Should we now pack up and go home, saying that "economists have analyzed Social Security and it is actually a nonissue"? We hope it is obvious that the answer is no. After all, all we have done so far is present a numerical example. The example *suggests* that Social Security might be irrelevant under certain circumstances, but it certainly does not *prove* that it is irrelevant in general.

This is how economists very often work. A simple numerical example can help us understand the operation of a complicated system like Social Security and can lead to some suggestive conclusions. Our next task is to determine whether the conclusion from our example holds more generally. We will first see that the result does not depend only on the particular numbers that we chose. We will then try to understand exactly where the conclusion comes from.

Key Takeaways

1. Households respond to variations in income over their lifetime through adjustments in saving to smooth consumption.
2. In a pay-as-you-go system, the government's payments to Social Security recipients exactly matches the revenues received from workers.
3. The integration of the government and household budget constraints implies that in a pay-as-you-go system, Social Security influences household saving but leaves lifetime consumption and income unchanged.

Exercises

1. What uncertainties did Carlo face over his lifetime?
2. In what parts of the discussion did we use the assumption that the real interest rate was zero?

This page titled 28.2: Individual and Government Perspectives on Social Security is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

28.3: A Model of Consumption 28.3: A Model of ConsumptionEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What is the life-cycle model of consumption?
2. What are the effects of changes in a pay-as-you-go system?

Chapter 28, 28.2 Section "Individual and Government Perspectives on Social Security" examined an explicit example of what Social Security implies for households and for the government. We can take away the following insights from this example:

- Households decide on consumption and saving taking into account their lifetime income.
- Lifetime income includes both taxes paid during working years and benefits received during retirement.
- From the government's view, taxes received and benefits paid need to balance, at least over long periods of time.
- In the example, the Social Security program was irrelevant: individuals had the same lifetime income and thus consumption opportunities regardless of the Social Security taxes paid and benefits received.

We now go beyond our numerical example and give a more general analysis of how an individual's lifetime consumption choices are influenced by Social Security.

Household Budget Constraints

We first consider the budget constraints faced by an individual or household (remember that we are using the two terms interchangeably). There are two household budget constraints. The first applies in any given period: ultimately, you must either spend the income you receive or save it; there are no other choices. For example,

$$disposable\ income = consumption + household\ savings.$$

Households also face a lifetime budget constraint. They can save in some periods of their life and borrow/dissave in other periods, but over the course of any household's lifetime, income and spending must balance. The simplest case is when real interest rates equal zero, which means that it is legitimate simply to add together income and consumption in different years. In this case the lifetime budget constraint says that

$$total\ lifetime\ consumption = total\ lifetime\ income.$$

If real interest rates are not zero, then the budget constraint must be expressed in terms of discounted present values. The household's lifetime budget constraint is then

$$discounted\ present\ value\ of\ lifetime\ consumption = discounted\ present\ valueof\ lifetime\ income.$$

If the household begins its life with some assets (say a bequest), we count this as part of income. If the household leaves a bequest, we count this as part of consumption. As in our earlier numerical example, we can think about the lifetime budget constraint in terms of the household's assets. Over the course of a lifetime, the household can save and build up its assets or dissave and run down its assets. It can even have negative assets because of borrowing. But the lifetime budget constraint says that the household's consumption and saving must result in the household having zero assets at the end of its life.

Toolkit: Section 31.4 and Section 31.5

You can review both the household's intertemporal budget constraint and the concept of discounted present value in the toolkit.

To see how this budget constraint works, consider an individual who knows with certainty the exact number of years for which she will work (her *working years*) and the exact number of years for which she will be retired (her *retirement years)*. While working, she receives her *annual disposable income*—the same amount each year. During retirement, she receives a Social Security payment that also does not change from year to year. As before, suppose that the real interest rate is zero.

Her budget constraint over her lifetime states that

$$\begin{aligned} &total\ lifetime\ consumption = total\ lifetime\ income \\ &= working\ years \times disposable\ income + retirement\ years \times Social\ Security\ payment. \end{aligned}$$

Our numerical example earlier was a special case of this model, in which

$$disposable\ income = \$34,000 working\ years = 45 retirement\ years = 15,$$

and

$$Social\ Security\ payment = \$18,000.$$

Plugging these values into the equation, we reproduce our earlier calculation of lifetime income (and hence also lifetime consumption) as $(\$45 \times \$34,000) + (15 \times \$18,000) = \$1,800,000$.

The Life-Cycle Model of Consumption

Economists often use a **consumption function** to describe an individual's consumption/saving decision:

$$consumption = autonomous\ consumption + marginal\ propensity\ to\ consume \times disposable\ income.$$

The **marginal propensity to consume** measures the effect of current income on current consumption, while autonomous consumption captures everything else, including past or future income.

The **life-cycle model** explains how households make consumption and saving choices over their lifetime. The model has two key ingredients: (1) the household budget constraint, which equates the discounted present value of lifetime consumption to the discounted present value of lifetime income, and (2) the desire of a household to smooth consumption over its lifetime.

Toolkit: Section 31.32 and Section 31.34

You can review the consumption function, consumption smoothing, and the life-cycle model in the toolkit.

Let us see how this model works. According to the life-cycle model of consumption, the individual first calculates her lifetime resources as

$$working\ years \times disposable\ income + retirement\ years \times Social\ Security\ payment.$$

(We continue to suppose that the real interest rate is zero, so it is legitimate simply to add her income in different years of her life.) She then decides how much she wants to consume in every period. Consumption smoothing starts from the observation that people do not wish their consumption to vary a lot from month to month or from year to year. Instead, households use saving and borrowing to smooth out fluctuations in their income. They save when their income is high and dissave when their income is low.

Perfect consumption smoothing means that the household consumes exactly the same amount in each period of time (month or year). Going back to the consumption function, perfect consumption smoothing means that the marginal propensity to consume is (approximately) zero.With perfect consumption smoothing, changes in current income will lead to changes in consumption only if those changes in income lead the household to revise its estimate of its lifetime resources. If a household wants to have perfectly smooth consumption, we can easily determine this level of consumption by dividing lifetime resources by the number of years of life. Returning to our equations, this means that

This is the equation we used earlier to find Carlo's consumption level. We took his lifetime income of $1,800,000, noted that lifetime income equals lifetime consumption, and divided by Carlo's 60 remaining years of life, so that consumption each year was $30,000. That is really all there is to the life-cycle model of consumption. Provided that income during working years is larger than income in retirement years, individuals save during working years and dissave during retirement.

This is a stylized version of the life-cycle model, but the underlying idea is much more general. For example, we could extend this story and make it more realistic in the following ways:

- Households might have different income in different years. Most people's incomes are not constant, as in our story, but increase over their lifetimes.
- Households might not want to keep their consumption exactly smooth. For example, if the household expects to have children, then it would probably anticipate higher consumption—paying for their food, clothing, and education—and it would expect to have lower consumption after the children have left home.
- The household might start with some assets and might also plan to leave a bequest.
- The real interest rate might not be zero.
- The household might contain more than one wage earner.

Working through the mathematics of these cases is more complicated—sometimes a lot more complicated—than the calculations we just did, and so is a topic for advanced courses in macroeconomics. In the end, though, the same key conclusions continue to hold even in the more sophisticated version of the life-cycle model:

- A household will examine its entire expected lifetime income when deciding how much to consume and save.
- Changes in expected future income will affect current consumption and saving.

The Government Budget Constraint

The household's budget constraints for different years are linked by the household's choices about saving and borrowing. Over the household's entire lifetime, these individual budget constraints can be combined to give us the household's lifetime budget constraint. Similar accounting identities apply to the federal government (and for that matter, to state governments and local governments as well).

In any given year, money flows into the government sector, primarily from the taxes that it imposes on individuals and corporations. We call these **government revenues**. The government also spends money. Some of this spending goes to the purchase of goods and services, such as the building of roads and schools or payments to teachers and soldiers. Whenever the government actually *buys* something with the money it spends, we call these **government purchases (or government expenditures)**. Some of the money that the government pays out is not used to buy things, however. It takes the form of **transfers**, such as welfare payments and Social Security payments. Transfers mean that dollars go from the hands of the government to the hands of an individual. They are like negative taxes. Social Security payments are perhaps the most important example of a government transfer.

Any difference between government revenues and government expenditures and transfers represents saving by the government. Government saving is usually referred to as a **government surplus**:

$$government\ surplus = government\ revenues - government\ transfers - government\ expenditures.$$

If, as is often the case, the government is borrowing rather than saving, then we instead talk about the **government deficit**, which is the negative of the government surplus:

$$government\ deficit = -government\ surplus = government\ transfers + government\ expenditures - government\ revenues.$$

Toolkit: Section 31.33 and Section 31.27

You can review the government budget constraint in the toolkit.

Applying the Tools to Social Security

The life-cycle model and government budget constraint can be directly applied to our analysis of Social Security. Let us go back to Carlo again. Carlo obtains pretax income and must pay Social Security taxes to the government. Carlo's disposable income in any given year is given by the equation

$$disposable\ income = income - Social\ Security\ tax.$$

Imagine that he receives no retirement income other than Social Security. Carlo's lifetime resources are given by the following equation:

$$\begin{aligned} lifetime\ resources = {} & working\ years \times income - working\ years \times Social\ Security\ tax \\ & + retirement\ years \times Social\ Security\ income. \end{aligned}$$

Now let us examine Social Security from the perspective of the government. To keep things simple, we suppose the only role of the government in this economy is to levy Social Security taxes and make Social Security payments. In other words, the government budget constraint is simply the Social Security budget constraint. The government collects the tax from each worker and pays out to each retiree. For the system to be in balance, the government surplus must be zero. In other words, government revenues must equal government transfers:

$$number\ of\ workers \times Social\ Security\ tax = number\ of\ retirees \times Social\ Security\ payment.$$

Now, here is the critical step. We suppose, as before, that all workers in the economy are like Carlo, and one worker is born every year. It follows that

$$number\ of\ workers = working\ years$$

and

$$number\ of\ retirees = retirement\ years.$$

But from the government budget constraint, this means that

$$working\ years \times Social\ Security\ tax = retirement\ years \times Social\ Security\ payment,$$

so *the second and third terms cancel in the expression for Carlo's lifetime resources*. Carlo's lifetime resources are just equal to the amount of income he earns over his lifetime *before* the deduction of Social Security taxes:

$$lifetime\ resources = income\ from\ working.$$

No matter what level of Social Security payment the government chooses to give Carlo, it ends up taking an equivalent amount away from Carlo when he is working. In this pay-as-you-go system, the government gives with one hand but takes away with the other, and the net effect is a complete wash. We came to this conclusion simply by examining Carlo's lifetime budget constraint and the condition for Social Security balance. We did not even have to determine Carlo's consumption and saving during each year. And—to reiterate—the assumption that there is just one person of each age makes no difference. If there were 4 million people of each age, then we would multiply both sides of the government budget constraint by 4 million. We would then cancel the 4 million on each side and get exactly the same result.

We have gained a remarkable insight into the Social Security system. The lifetime income of the individual *is independent of the Social Security system*. Whatever the government does to tax rates and benefit levels, provided that it balances its budget, there will be no effect on Carlo's lifetime income. Since consumption decisions are made on the basis of lifetime income, it also follows that the level of consumption is independent of variations in the Social Security system. Any changes in the Social Security system result in changes in the level of saving by working households but nothing else. As we saw in our original numerical example, individuals adjust their saving in a manner that cancels out the effects of the changes in the Social Security system.

The model of consumption and saving we have specified leads to a very precise conclusion: the household neither gains nor loses from the existence of the Social Security system. The argument is direct. If the well-being of the household depends on the consumption level over its entire lifetime, then Social Security is irrelevant since lifetime income (and thus consumption) is independent of the Social Security system.

Key Takeaways

1. The life-cycle model of consumption states that the household chooses its consumption during each period of life subject to a budget constraint that the discounted present value of lifetime income must equal the discounted present value of lifetime consumption.
2. If the household chooses to perfectly smooth consumption, then consumption during each period of life is equal to the discounted present value of income divided by the number of years in a lifetime.
3. In general, a household's lifetime income and consumption are independent of the taxes and benefits of a pay-as-you-go Social Security system. Changes to the system lead to adjustments in saving rather than consumption.

Exercises

1. What are the two types of budget constraints that a household faces?
2. If working years increased by five and retirement years decreased by five, what would happen to lifetime income?

This page titled 28.3: A Model of Consumption is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

28.4: Social Security in Crisis? 28.4: Social Security in Crisis?Edit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What is the current state of the Social Security system in the United States?
2. What are some of the policy choices being considered?

The Social Security system in the United States went into deficit in 2010: tax receipts were insufficient to cover expenditures. This was largely because the recession led to reduced receipts from the Social Security tax. However, the Social Security Board of Trustees warns that "[a]fter 2014, cash deficits are expected to grow rapidly as the number of beneficiaries continues to grow at a substantially faster rate than the number of covered workers." "A Summary of the 2011 Annual Reports: Social Security and Medicare Boards of Trustees," Social Security Administration, accessed June 24, 2011, http://www.ssa.gov/OACT/TRSUM/index.html.

It is hard to reconcile these statements with the model that we developed in 28.2 Section "Individual and Government Perspectives on Social Security" and 28.3 Section "A Model of Consumption". If Social Security is an irrelevance, why is there so much debate about it, and why is there so much concern about its solvency? The answer is that our model was *too* simple. The framework we have developed so far is a great starting point because it tells us about the basic workings of Social Security in a setting that is easy to understand. Don't forget, though, that our discussion was built around a pay-as-you-go system in a world where the ratio of retirees to workers was not changing. Now we ask what happens if we complicate the demography of our model to make it more realistic.

The Baby Boom

During the period directly following World War II, the birthrate in many countries increased significantly and remained high for the next couple of decades. People born at this time came to be known—for obvious reasons—as the baby boom generation. The baby boomers in the United States and the United Kingdom are clearly visible in Figure 28.4.1, which shows the age distribution of the population of those countries. If babies were being born at the same rate, you would expect to see fewer and fewer people in each successive age group. Instead, there is a bulge in the age distribution around ages 35–55. (Interestingly, there is also a second baby boomlet visible, as the baby boomers themselves started having children.)

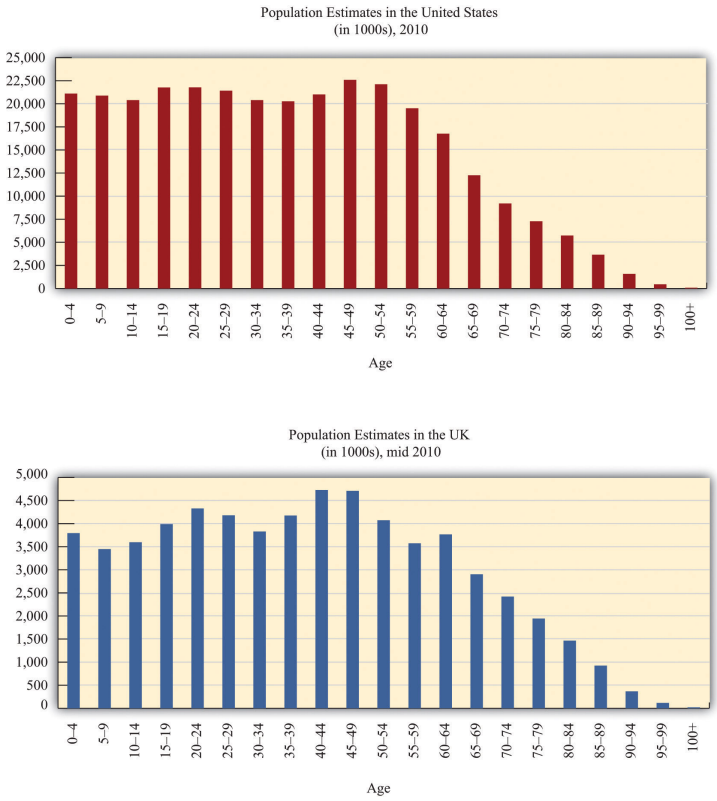


Figure 28.4.1: The Baby Boom in the United States and the United Kingdom. If the same number of people were born every year, then a bar chart of population at different age groups would show fewer and fewer people in each successive age group. Instead, as these pictures show, the United States and the United Kingdom had a "baby boom": an unusually large number of children were born in the decades immediately following World War II. In 2010, this generation is in late middle age.

Source: US Census Bureau, International Data Base, www.census.gov/population/international/data/idb/informationGateway.php.

Figure 28.4.2 presents the equivalent US data for 1980, 1990, and 2000, showing the baby boom working its way through the age distribution.

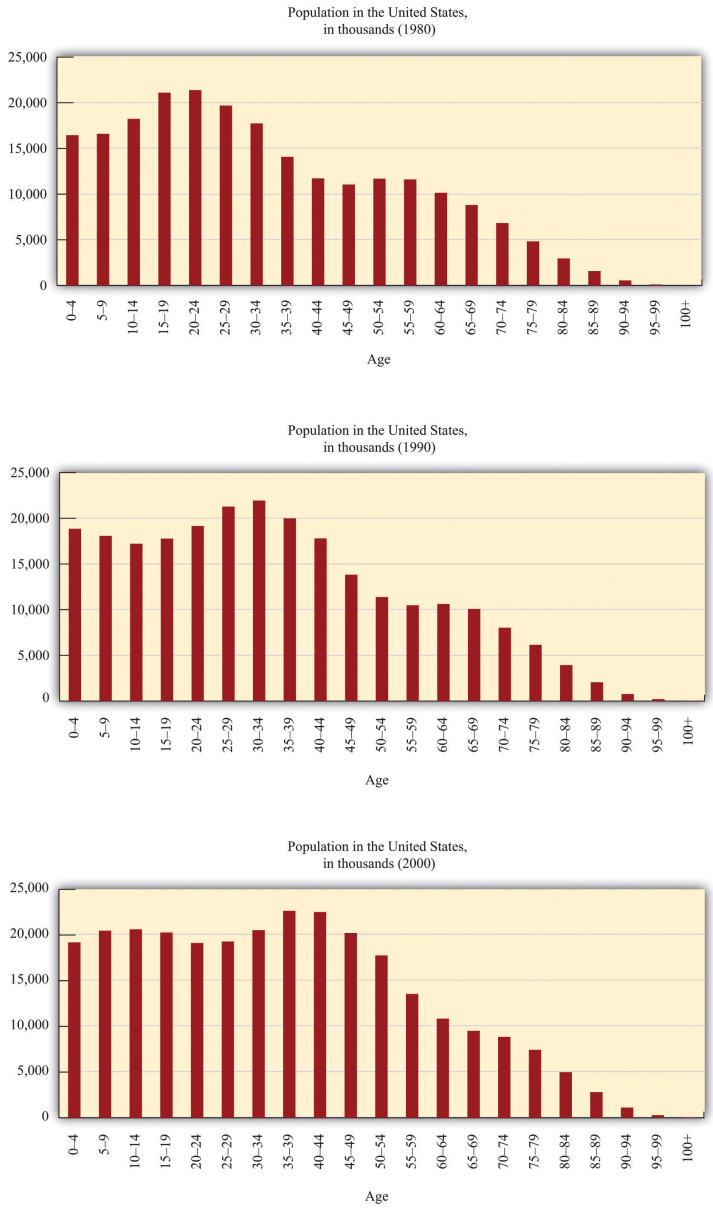


Figure 28.4.2: *The US Baby Boom over Time*. These pictures show the age distribution of the population as the baby boom generation gets older. The "bulge" in the age distribution shifts rightward. In 1980, the baby boomers were young adults. By 2000, even the youngest baby boomers were in middle age.

Source: US Census Bureau, International Data Base, www.census.gov/population/international/data/idb/informationGateway.php.

As the baby boom generation makes its way to old age, it is inevitable that the **dependency ratio**—the ratio of retirees to workers—will increase dramatically. In addition, continuing advances in medical technology mean that people are living longer than they used to, and this too is likely to cause the dependency ratio to increase. The 2004 *Economic Report of the President* predicted that the dependency ratio in the United States will increase from 0.30 in 2003 to 0.55 in 2080. *Economic Report of the President* (Washington, DC: GPO, 2004), accessed July 20, 2011, www.gpoaccess.gov/usbudget/fy05/pdf/2004_erp.pdf. Roughly speaking, in other words, there are currently about three workers for every retiree, but by 2080 there will only be two workers per retiree.

A Baby Boom in Our Model

In our framework, we assumed that there was always one person alive at each age. This meant that the number of people working in any year was the same as the working life of an individual. Likewise, we were able to say that the number of people retired at a point in time was the same as the length of the retirement period.

Here is a simple way to represent a baby boom: Let us suppose that, *in one year only*, two people are born instead of one. When the extra person enters the work force, the dependency ratio will decrease—there is still the same number of retirees, but there are more workers. If Social Security taxes are kept unchanged and the government continues to keep the system in balance every year, then the government can pay out higher benefits to retirees. For 45 years, retirees can enjoy the benefits of the larger workforce.

Eventually, though, the baby boom generation reaches retirement age. At that point the extra individual stops contributing to the Social Security system and instead starts receiving benefits. What used to be a boon is now a problem. To keep the system in balance, the government must *reduce* Social Security benefits.

Let us see how this works in terms of our framework. Begin with the situation before the baby boom. We saw earlier that the government budget constraint meant that Social Security revenues must be the same as Social Security payments:

$$number\ of\ workers \times Social\ Security\ tax = number\ of\ retirees \times Social\ Security\ payment.$$

If we divide both sides of this equation by the number of retirees, we find that

The first expression on the right-hand side (number of workers/number of retirees) is the inverse of the dependency ratio.

- **When the baby boom generation is working.** Once the additional person starts working, there is the same number of retirees, but there is now one extra worker. Social Security revenues therefore increase. If the government continues to keep the system in balance each year, it follows that the annual payment to each worker increases. The dependency ratio has gone down, so payments are larger. The government can make a larger payment to each retired person while still keeping the system in balance. Retirees during this period are lucky: they get a higher payout because there are relatively more workers.
- **When the baby boom generation retires.** Eventually, the baby boom generation will retire, and there will be one extra retiree each year until the baby boom generation dies. Meanwhile, we are back to having fewer workers. So when the baby boom generation retires, the picture is reversed. Social Security payments are higher than in our baseline case, and revenues are back to where they were before the baby boomers started working. Because there are now more retirees relative to workers—that is, the dependency ratio has increased—retirees see a cut in Social Security benefits.

If the *Economic Report of the President* figures are to be believed, the coming increase in the dependency ratio means that Social Security payments would have to decrease by about 45 percent if the Social Security budget were to be balanced every year. The reality is that this simply will not happen. First, the Social Security system does not simply calculate payouts on the basis of current Social Security receipts. In fact, there is a complicated formula whereby individuals receive a payout based on their average earnings over the 30 years during which they earned the most. Kaye A. Thomas, “Understanding the Social Security Benefit Calculation,” Fairmark, accessed July 20, 2011, www.fairmark.com/retirement/socsec/pia.htm. Of course, that formula could be changed, but it is unlikely that policymakers will completely abandon the principle that payments are based on past earnings. Second, retired persons already make up a formidable political lobby in the United States. As they become more numerous relative to the rest of the population, their political influence is likely to become even greater. Unless the political landscape changes massively, we can expect that the baby boom generation will have the political power to prevent a massive reduction in their Social Security payments.

Social Security Imbalances

To completely understand both the current situation and the future evolution of Social Security, we must make one last change in our analysis. Although the Social Security system was roughly in balance for the first half-century of its existence, that is no longer the case. Because payments are calculated on the basis of past earnings, it is possible for revenues to exceed outlays or be smaller than outlays. This means that the system is not operating on a strict pay-as-you-go basis.

When the government originally established Social Security, it set up something called the Social Security Trust Fund—think of it as being like a big bank account. Current workers pay contributions into this account, and the account also makes payments to retired workers. Under a strict pay-as-you-go system, the balance in the trust fund would always be zero. In fact, in some years payments to workers are smaller than tax receipts, in which case the extra goes into the Trust Fund. In other year payments exceed receipts, and the difference is paid for out of the Trust Fund.

To be more precise,

$$tax\ revenues = number\ of\ workers \times Social\ Security\ taxes = number\ of\ workers \times tax\ rate \times income$$

and

$$Social\ Security\ payments = number\ of\ retirees \times Social\ Security\ payment.$$

If tax revenues exceed payments, then the system is running a surplus: it is taking in more in income each period than it is paying out to retirees. Conversely, if payments exceed revenues, the system is in deficit. In other words,

$$Social\ Security\ surplus = number\ of\ workers \times tax\ rate \times income - number\ of\ retirees \times Social\ Security\ payment.$$

For the first half-century of Social Security, there was an approximate match between payments and receipts, although receipts were usually slightly larger than payments. In other words, rather than being exactly pay-as-you-go, the system typically ran a small surplus each year. "Trust Fund Data," Social Security Administration, January 31, 2011, accessed July 20, 2011, http://www.ssa.gov/OACT/STATS/table4a1.html. Over the first half-century of the program, the Trust Fund accumulated slightly less than $40 billion in assets. This might sound like a big number, but it amounts to only a few hundred dollars per worker. The Social Security Trust Fund contains the accumulated surpluses of past years. It gets bigger or smaller over time depending on whether the surplus is positive or negative. For example,

$$Trust\ Fund\ balance\ this\ year = Trust\ Fund\ balance\ last\ year + Social\ Security\ surplus\ this\ year.$$

(Strictly, that equation is true provided that we continue to suppose that the real interest rate is zero.) If tax revenues exceed payments, then there is a surplus, and the Trust Fund increases. If tax revenues are less than payments, then there is a deficit (or, to put it another way, the surplus is negative), so the Trust Fund decreases.

The small surpluses that have existed since the start of the system mean that the Trust Fund has been growing over time. Unfortunately, it has not been growing fast enough, and in 2010, the fund switched from running a surplus to running a deficit. There are still substantial funds in the system—almost a century's worth of accumulated surpluses. But the dependency ratio is so high that those accumulated funds will disappear within a few decades.

Resolving the Problem: Some Proposals

We can use the life-cycle model of consumption/saving along with the government budget constraint to better understand proposals to deal with Social Security imbalances.

We saw that the surplus is given by the following equation:

$$Social\ Security\ surplus = number\ of\ workers \times tax\ rate \times income - number\ of\ retirees \times Social\ Security\ payment.$$

The state of the Social Security system in any year depends on five factors:

1. The level of income
2. The Social Security tax rate on income
3. The size of the benefits
4. The number of workers
5. The number of retirees

Other things being equal, increases in income (economic growth) help push the system into surplus.The effect of economic growth is lessened because of the fact that Social Security payments are linked to past earnings. Higher growth therefore implies higher payouts as well as higher revenues. Still, on net, higher growth would help Social Security finances. A larger number of the population of working age also tends to push the system into surplus, as does a higher Social Security tax. On the other hand, if benefits are higher or there are more retirees, that tends to push the system toward deficit.

Increasing Taxes or Decreasing Benefits

Many of the proposals for reforming Social Security can be understood simply by examining the equation for the surplus. Remember that the number of workers × the tax rate × income is the tax revenue collected from workers, whereas the number of retirees × the Social Security payment is the total transfer payments to retirees. If the system is running a deficit, then to restore balance, either revenues must increase or payouts must be reduced.

The tax rate and the amount of the payment are directly under the control of the government. In addition, there is a ceiling on income that is subject to the Social Security tax ($106,800 in 2011). At any time, Congress can pass laws changing these variables. It could increase the tax rate, increase the income ceiling, or decrease the payment. If we simply think of the problem as a mathematical equation, then the solution is easy: either increase tax revenues or decrease benefits. Politics, though, is not mathematics. Politically, such changes are very difficult. Indeed, politicians often refer to increases in taxes and/or reductions in

benefits as a political “third rail” (a metaphor that derives from the high-voltage electrified rail that provides power to subway trains—in other words, something not to be touched).

Another way to increase revenue is through increases in GDP. If the economy is expanding and output is increasing, then the government will collect more tax revenues for Social Security. There are no simple policies that guarantee faster growth, however, so we cannot plan on solving the problem this way.

Delaying Retirement

We have discussed the tax rate, the payment, and the level of income. This leaves the number of workers and the number of retirees. We can change these variables as well. Specifically, we can make the number of workers bigger and the number of retirees smaller by changing the retirement age. This option is frequently discussed. After all, one of the causes of the Social Security imbalance is the fact that people are living longer. So, some ask, if people live longer, should they work longer as well?

Moving to a Fully Funded Social Security System

The financing problems of Social Security stem from a combination of two things: demographic change and the pay-as-you-go approach to financing. Suppose that, instead of paying current retirees by taxing current workers, the government were instead simply to tax workers, invest those funds on their behalf, and then pay workers back when they are retired. Economists call this a **fully funded Social Security system**. In this setup, demographic changes such as the baby boom would not be such a big problem. When the baby boom generation was working, the government would collect a large amount of funds so that it would later have the resources to pay the baby boomers their benefits.

As an example, Singapore has a system known as the Central Provident Fund, which is in effect a fully funded Social Security system. Singaporeans make payments into this fund and are guaranteed a minimum return on their payments. In fact, Singapore sets up three separate accounts for each individual: one specifically for retirement, one that can be used to pay for medical expenses, and one that can be used for specific investments such as a home or education.

Some commentators have advocated that the United States should shift to a fully funded Social Security system, and many economists would agree with this proposal. “Economic Letter,” Federal Reserve Bank of San Francisco, March 13, 1998, accessed July 20, 2011, http://www.frbsf.org/econrsrch/wklyltr/wklyltr98/el98-08.html. This letter discussed the transition to a fully funded Social Security system. Were it to adopt such a system, the US government would not in the future have the kinds of problems that we currently face. Indeed, the Social Security reforms of the 1980s can be considered a step away from pay-as-you-go and toward a fully funded system. At that time, the government stopped keeping the system in (approximate) balance and instead started to build up the Social Security Trust Fund.

But this is not a way to solve the *current* crisis in the United States. It is already too late to make the baby boomers pay fully for their own retirement. Think about what happened when Social Security was first established. At that time, old workers received benefits that were much greater than their contributions to the system. That generation received a windfall gain from the establishment of the pay-as-you-go system. That money is gone, and the government cannot get it back.

Suppose the United States tried to switch overnight from a pay-as-you-go system to a fully funded system. Then current workers would be forced to pay for Social Security twice: once to pay for those who are already retired and then a second time to pay for their own retirement benefits. Obviously, this is politically infeasible, as well as unfair. Any realistic transition to a fully funded system would therefore have to be phased in over a long period of time.

Privatization

Recent discussion of Social Security has paid a lot of attention to privatization. Privatization is related to the idea of moving to a fully funded system but with the additional feature that Social Security evolves (at least in part) toward a system of private accounts where individuals have more control over their Social Security savings. In particular, individuals would have more choice about the assets in which their Social Security payments would be invested. Advocates of this view argue that individuals ought to be responsible for providing for themselves, even in old age, and suggest that private accounts would earn a higher rate of return. Opponents of privatization argue, as did the creators of Social Security in the 1930s, that a privatized system would not provide the assistance that elderly people need.

Some countries already have social security systems with privatized accounts. In 1981, Chile’s pay-as-you-go system was virtually bankrupt and replaced with a mandatory savings scheme. Workers are required to establish an account with a private pension company; however, the government strictly regulates these companies. The system has suffered from compliance problems,

however, with much of the workforce not actually contributing to a plan. In addition, it turns out that many workers have not earned pensions above the government minimum, so in the end it is not clear that the private accounts are really playing a very important role. Recent reforms have attempted to address these problems, but it remains unclear how successful Chile's transition to privatization will be.

As with the move to a fully funded Social Security system, a big issue with privatization is the transition period. If, for example, the government announced a plan today to privatize Social Security, it would have to deal with the fact that many retired people would no longer have Social Security income. Furthermore, many working people would have already paid into the program. Thus proposals to privatize Social Security must include a plan for dealing with existing retirees and those who have paid into the system through payroll taxes.

Some recent discussion has suggested, implicitly or explicitly, that privatization would help solve the current Social Security imbalance. This is misleading. By cutting off the payroll tax revenues, privatization makes the problem worse in the short run, not better. Although privatization is certainly a proposal that can be discussed on its own merits, it should be kept separate from the debate about how to balance existing Social Security claims with revenues.

Key Takeaways

1. Many studies predict that, if there are no policy changes, the Social Security system will be bankrupt by the middle of this century. A main cause of this problem is demographic change: fewer workers are supporting more retirees, and life expectancies have increased.
2. Some possible policy remedies include raising taxes on workers, reducing benefits, and increasing the retirement age.

Exercises

1. What is the dependency ratio? Why might it change over time?
2. What is the Social Security Trust Fund?

This page titled 28.4: Social Security in Crisis? is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

28.5: The Benefits and Costs of a Social Security System 28.5: The Benefits and Costs of a Social Security SystemEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What are the benefits of having a Social Security system?
2. How does a Social Security system help someone deal with the uncertainties of life?
3. What are the effects of Social Security on national saving?

We have seen how demographic changes in the economy, combined with the pay-as-you-go form of Social Security, are leading to funding problems within the US system. The United States is not alone; many other countries also have pay-as-you-go systems and are facing similar demographic challenges. We have also examined some ways of resolving these financing problems. Yet we have not addressed another more basic question: why have a Social Security system at all? After all, our analysis suggests that people may adjust their private saving behavior in a way that undoes the effects of Social Security. What advantages and disadvantages of Social Security have we so far missed?

The Uncertainties of Life

A century or two ago, if you were unlucky enough to fall into serious poverty, there was very little in the way of government help, even in the richest countries. You were likely to end up in the poorhouse (sometimes called a workhouse or an almshouse), where you obtained the bare minimum of shelter and food in exchange for grueling work. For those who were old and poor, the poorhouse was a place to die an ignominious death:

Numerous as are the old men's homes, old ladies' homes, and homes for aged couples that are supported by private charity, they are yet, as every worker among the poor knows, too few to meet the demand. Our almshouses are also practically homes for the aged poor. Some almshouse inmates became paupers before they were aged, but many of them led independent and self-respecting lives, and even put by something for the future while physically able to earn wages. When wages ceased, savings, if any were made, were used up or else lost in unwise investments, and at the end almshouse relief and the pauper's grave were preferred to exposure and starvation. Henry Seager, *Social Insurance: A Program of Social Reform*, Chapter V—"Provision of Old Age," 1910, accessed August 9, 2011, http://www.ssa.gov/history/pdf/seager5.pdf.

Social Security in the United States and other countries was set up largely to save old people from this fate.

Carlo did not face any of the problems suggested by the quotation. In Carlo's world there was no uncertainty: working and retirement income were known at the start of his working life, and his dates of retirement and death were also known with certainty. Carlo had no risk of using up all his savings before he died, or of losing his money in "unwise investments." But Carlo's world is not the world in which we live. In practice, individuals face enormous uncertainty both about their lifetime income and their consumption needs in retirement.

The mere fact that we live in an uncertain world is not, in and of itself, a reason for the government to intervene. Private insurance markets might be available that allow individuals to purchase insurance to cover themselves against these kinds of risks. As an example, many people have disability insurance that they either purchase individually or obtain through their employer. Disability insurance means that if you are unlucky enough to suffer an accident or illness that prevents you from working, you will still receive income. It is also possible to purchase **annuities** (which are sort of a reverse life insurance): these are assets that pay out a certain amount each year while you are alive and allow you to insure yourself against the uncertain time of your death.

Early discussions of Social Security highlighted the insurance role of the program. During the Great Depression, it became clear that insurance provided through markets was woefully incomplete. Thus the government created a variety of safety nets, financed by public funds. Social Security was one of these programs. An early pamphlet on Social Security summarizes this view:

In general, the Social Security Act helps to assure some income to people who cannot earn and to steady the income of millions of wage earners during their working years and their old age. In one way and another taxation is spread over large groups of people to carry the cost of giving some security to those who are unfortunate or incapacitated at any one time. The act is a foundation on which we have begun to build security as States and as a people, against the risks which families cannot meet one by one.Mary

Ross, "Why Social Security?" Bureau of Research and Statistics, 1937, accessed July 20, 2011, http://www.ssa.gov/history/whybook.html.

Financial sophistication has increased markedly since the 1930s, but insurance markets are still far from perfect, so most people agree that the government should continue to provide the insurance that private markets fail to deliver. As President George W. Bush's Council of Economic Advisors wrote, "To protect against this risk [of living an unusually long time], a portion of the retirement wealth that a worker has accumulated must be converted into an annuity, a contract that makes scheduled payments to the individual and his or her dependents for the remainder of their lifetimes." *Economic Report of the President* (Washingon, DC: GPO, 2004), accessed July 20, 2011, www.gpoaccess.gov/usbudget/fy05/pdf/2004_erp.pdf, p. 130. Once we acknowledge two things—(1) there is major uncertainty in life, and (2) insurance markets are lacking—we see a clear role for Social Security.

The Complexity of Optimization

There is another reason to think that our analysis of Carlo was much too simple. For Carlo, it was quite straightforward to determine his optimal level of consumption: all he had to do was to calculate his lifetime income, divide by the number of years of life that he had left, and he knew his optimal level of consumption.

We said earlier that the basic idea of this life-cycle model continues to hold even in a more complicated world, where incomes are not constant, real interest rates are not zero, and consumption needs may vary over one's lifetime. If you have a PhD in economics, you even learn to solve these problems in a world of uncertainty.

Yet when one considers all the uncertainties of life, the problem certainly becomes very complex. Most individuals do not have PhDs in economics, and most people—even including those with economics PhDs—are not able to forecast their income and consumption needs very accurately. As a result, it seems likely that many people are not capable of making good decisions when they are thinking about consumption and saving over their entire lifetimes. As stated in the 2004 *Economic Report of the President*, "Some individuals may not be capable of making the relevant calculations themselves and may not be able to enlist the service of a financial professional to advise them." *Economic Report of the President* (Washingon, DC: GPO, 2004), accessed July 20, 2011, www.gpoaccess.gov/usbudget/fy05/pdf/2004_erp.pdf, p. 130. Social Security can therefore be seen as a program that provides assistance to individuals unable to make optimal decisions on their own.That said, figuring payments under the current social security system is not easy either. To understand why, check out the information on benefits at the Social Security Administration website. "Your Retirement Benefit: How It Is Figured," Social Security Administration, January 2011, accessed July 20, 2011, http://www.ssa.gov/pubs/10070.html.

In general, economists believe both that people are aware of their own self-interests and are capable of making good decisions. Economists tend to be suspicious of arguments that suggest that the government can make better decisions for people than they can make for themselves. At the same time, research by economists and psychologists suggests that individuals are subject to biases and errors of judgment in their decision making. And if government paternalism makes sense anywhere, then it is likely to be in the context of lifetime saving decisions. After all, we are not talking about deciding which kind of coffee to buy or what price to set for a product this month. There is no room for learning from your mistakes, there are no second chances, and the consequences of error are enormous. In life, you only get old once.

Distortions and Administrative Costs

The key arguments in favor of Social Security are therefore that it provides some insurance that may not be available through private markets and protects people in the face of their inability to make sound decisions when they are planning for the distant future. But just because there are some shortcomings of private insurance and annuity markets, we should not presume that government can do things better. Against the benefits of the Social Security system must also be set some costs.

First, any government program requires resources to operate. It costs about 1 percent of the benefits paid to administer the Social Security system. This is a direct cost of the program. Second—and more interestingly in terms of economics—whenever we have a government scheme that affects the taxes that people pay, there will be some distortionary effects on people's willingness to work. Taxes lower the relative price of leisure compared to consumption goods, which may induce people to work less. Because Social Security imposes a tax on the incomes of working people, it distorts their choices. This is another cost of the Social Security system.

The Effect on National Savings

There is another effect of Social Security that is much more subtle. It reduces the savings of the nation as a whole. This means less capital and ultimately lower living standards. The intuition is as follows. When individuals save, they make funds available in the financial markets for firms to borrow. Thus saving leads to investment and a buildup of the economy's capital stock. But as we saw, Social Security reduces the individual incentive to save. People don't need to save if the government will provide for them in retirement. Furthermore, the taxes being collected by the government are not being used to finance capital investment either; they are being paid out to old workers.

A pay-as-you-go system thus tends to reduce overall national saving. In a fully funded Social Security system, this is not an issue, and indeed this is one of the most compelling arguments in favor of a gradual shift to a fully funded system.

Redistributive Effects of Social Security

Social Security redistributes income in ways that may not be desirable. After all, those who benefit the most from Social Security are those who live the longest. Thus the scheme effectively redistributes money from the unlucky people who die young to the lucky ones who live for a very long time. This is a politically charged argument, for life expectancy is correlated with poverty, race, and sex. The life expectancy of poor African American men is significantly lower than the life expectancy of rich white American women, for example. Social Security may redistribute resources, from poor African American men to rich white American women.

Key Takeaways

1. Some benefits of a Social Security system arise from the provision of insurance over the uncertainties of life and in helping people make once in a lifetime choices that are very complex.
2. Through the Social Security system, retirees receive benefits until they die. This is a form of insurance to deal with the uncertainties of life.
3. Since a pay-as-you-go Social Security system provides income during retirement years, it reduces the incentive for households to save.

Exercises

1. How does Social Security help people who are unable to make choices on their own?
2. In what ways does Social Security redistribute resources across households?

This page titled 28.5: The Benefits and Costs of a Social Security System is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

28.6: Social Security in the Real World 28.6: Social Security in the Real WorldEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What aspects of the real world are highlighted, and which are missed in our simple framework?
2. Why do people disagree about Social Security reform?

Our discussion of Social Security deliberately used a simple framework. Using that framework, we first showed that, in the simplest case, the Social Security system actually has no effect on the lifetime consumption of households. We also explained that, once we move away from this simple setup, there are some arguments both for and against a Social Security system.

Complications

The world is much more complicated than our simple framework, and we need to make sure that our analysis has not left out some important feature of the real world that would change our conclusion. In this section, we briefly discuss some complications to our model. Some of these complications provide some additional reasons to support a Social Security system; others identify additional costs of the system. However, these additional costs and benefits are much less important than those we have already identified.

Positive Real Interest Rates

We based all our discussion on an assumption that the real interest rate is zero. When the real interest rate is zero, it is legitimate to add real income in different years and consumption in different years. With a real interest rate of zero, adding income levels in different periods is not a problem. But if the real interest rate is positive, this is not correct. To add income and consumption in different years, we have to calculate discounted present values.

Toolkit: Section 31.5

You can review discounted present value in the toolkit.

Suppose you will receive some income next year. The value of that this year is given by the following equation:

Income earned in the future has a lower value from the perspective of today. The mathematics of the lifetime budget constraint is harder once we allow for nonzero interest rates, so we will not go through the formal calculations here. Without going through all the details of the analysis, what can we conclude?

The main observation is a rather surprising one. Once we introduce a positive real interest rate, the Social Security system makes people worse off. Remember that we concluded earlier that the system had no effect on the total resources in the hands of the household. Households are taxed when they are young, though, and get that money returned to them when they are old. With positive real interest rates, they would strictly prefer the money when they were young.

This result seems odd. A Social Security system allows the government, in effect, to borrow from the future, taxing younger generations to pay older generations. So how does it end up making people worse off? A key part of the answer is that, when the system was first introduced, the first generation of old people obtained benefits without having to make contributions. In the past, therefore, the introduction of the Social Security system did make one group of people better off.

Economic Growth

As we know, most economies grow over time. We neglected this in our analysis. Economic growth has two implications for Social Security: one unimportant and one more significant. First, economic growth is another reason why individuals' incomes increase over the course of their lifetimes. We have already observed that this does not change the fundamental idea of lifetime consumption smoothing: you still add lifetime income in both working and nonworking years and then divide by the number of years of life to find the optimal level of consumption.

More interestingly, economic growth also means that Social Security payments increase over time. As the income of workers increases because of economic growth, so too does the amount of tax collected by the government. If the Social Security system is

in balance at all times, Social Security payments must also increase. Thus when workers are retired, they continue to enjoy the benefits of economic growth. (In fact, if the growth rate of the economy happened to be the same as the real interest rate, the effect of positive economic growth would exactly offset the negative effect of real interest rates.) Normally, the effect of economic growth partially offsets the negative effect of positive real interest rates.

Access to Credit Markets

In our setting, individuals were able to save without difficulty at the market real interest rate (which was zero in our basic formulation). In the jargon of economics, individuals have good access to credit markets. Yet many individuals in reality have a limited ability to borrow and lend.In our example, individuals wanted to save and not to borrow because they obtained income early in life. If we made more realistic assumptions about the patterns of wages over the lifetime, we would typically find that people want to borrow at certain times of their lives. For example, people often borrow early in life to finance their education. There is ample evidence that many people do not actively participate in stock markets: they do not hold mutual funds or shares of individual companies' stocks. Such individuals typically save by putting money in a bank, and the interest they earn is relatively low. In particular, it is lower than the interest that the Social Security Trust Fund can earn.

Social Security in effect allows the government to do some saving on behalf of individuals at a better interest rate than they themselves can earn. Thus individuals who do not have good access to capital markets can be made better off by access to a Social Security system.

This is in some ways the exact opposite of the argument for privatization. Supporters of privatization argue that if individuals can make their own investment decisions, they can earn a better interest rate than is provided by Social Security. They point out that, on average, the stock market provides a better rate of return than is provided by the system. This argument is correct: people may be able to do better. We need to recognize, though, that these higher returns would come at the cost of higher risk—which brings us right back to the original argument for why we need a Social Security system.

Moral Hazard

Finally, because Social Security serves as a form of insurance, it is subject to problems that are faced by all insurance systems. One of these goes by the name **moral hazard**, which simply means that the presence of insurance may cause people to change their behavior in bad ways. For example, if people have fire insurance, they may be less likely to keep a fire extinguisher in their homes. Similarly, because people know that the government will provide them with Social Security, they have less incentive to manage their own saving in a careful manner.

Why Do People Disagree about Social Security?

President George W. Bush's suggestions for reforming the Social Security system encountered a lot of opposition and rapidly became a partisan issue in US politics. Yet it seems as if Social Security is a program that we could analyze completely and carefully using the tools of economics. Why is a basic economic program such as Social Security so politicized?

Some people, of course, will view any proposal from the perspective of politics. There are undoubtedly people who supported President George W. Bush's proposals not on their merits or demerits but just because they support the Republican Party. Likewise, there are surely Democrats who opposed the president's proposals simply because they came from a Republican. But leaving such extreme partisan viewpoints aside, there are still good reasons why reasonable people might have different opinions on Social Security:

- **People differ in their assessment of the importance of market failure in insurance markets.** A key argument for Social Security is that private markets do not permit people to insure themselves against the risk of poverty in old age. Insurance and annuity markets do exist, so some people argue that this failure of markets is no longer very significant. At the same time, it requires financial sophistication to take advantage of these markets. Many people do not have the expertise to use these markets or access to financial professionals who could advise them.
- **People differ in their beliefs about whether individuals can make good decisions about lifetime consumption and savings.** Economists generally think that individuals are the best judges of their own well-being. As a consequence, economists are suspicious of arguments that suggest that the government knows better than you do how you should make your own private decisions (such as how to manage your money). However, the decision making required for lifetime financial planning is very complicated, and the consequences of error are so severe, that many economists nonetheless think that failures of individual decision making are a good reason to support Social Security.

- **People differ in their beliefs about how much government should be involved in people's lives.** Some people are, in general, philosophically opposed to significant government involvement in individual decisions. Even if individuals make poor decisions about their lifetime consumption and savings and end up poor, these people would argue that individuals should bear the consequences of their own mistakes, and government should not bail them out. Others tend to the view that government has a critical role to play in protecting the unfortunate and unlucky.
- **People have different views about fairness and equality.** Some people have the view that an important function of government is to protect the worst off in society and to redistribute some resources from those who are relatively rich to those who are poorer. Such people tend to be strong supporters of programs such as Social Security because it protects those who, through bad luck or poor decisions, would otherwise end their lives in poverty. Others disagree, saying that government should not be involved in redistribution of resources. They also point out, as we observed earlier, that Social Security, by its very nature, benefits those who live for a long time, so it is not a good deal for groups with lower life expectancies.

Beyond Social Security

You may have heard in the news that discussion of the need to reform Social Security applies to other government programs. In particular, if a part of the Social Security program is a growing imbalance in the age distribution, then other programs that support transfers to older people are potentially in trouble as well.

A leading example of this is the Medicare program. You can find information about this program at Medicare.gov: http://www.medicare.gov/default.asp. This program provides health care to the elderly. A second example is Medicaid, which is also a publically funded program, administered at the state level, to provide health care; this program is intended to provide assistance to poor people. "Medicaid Program—General Information," US Department of Health and Human Services, June 16, 2011, accessed July 20, 2011, www.cms.hhs.gov/MedicaidGenInfo. These programs, like Social Security, entail large outlays by the government. In his testimony in June 2008 to the Senate Finance Committee, Peter Orszag, the director of the CBO, stated the following: "The Congressional Budget Office (CBO) projects that total federal Medicare and Medicaid outlays will rise from 4 percent of GDP [gross domestic product] in 2007 to 12 percent in 2050 and 19 percent in 2082, which, as a share of the economy, is roughly equivalent to the total amount that the federal government spends today. The bulk of that projected increase in health care spending reflects higher costs per beneficiary rather than an increase in the number of beneficiaries associated with an aging population." "The Long-Term Budget Outlook and Options for Slowing the Growth of Health Care Costs," Congressional Budget Office, June 17, 2008, accessed July 20, 2011, http://www.cbo.gov/doc.cfm?index=9385. This quote contains two key ideas. First, it seems likely that outlays for these two programs will be growing rapidly over the next 50 or so years. From the CBO projections, the share of spending on Medicare and Medicaid grows while the share of spending on Social Security is basically constant after 2020. This comes from figure 1 of the following testimony: "The Long-Term Budget Outlook and Options for Slowing the Growth of Health Care Costs," Congressional Budget Office, June 17, 2008, accessed September 20, 2011, http://www.cbo.gov/doc.cfm?index=9385. Second, in contrast to Social Security, the problem is not only demographics. Instead, as noted in the testimony, a significant part of the increased cost of these programs comes from the increases in treatment per individual, rather than the number of individuals.

Thus as you use the tools provided in this chapter to ponder Social Security, keep in mind that other programs have similar budgetary challenges. Long-term solutions are needed either to finance the projected increase in outlays or to reduce the costs of these programs.

Key Takeaways

1. The framework we presented captures the idea that saving is used to smooth consumption over a lifetime, and lifetime income includes taxes paid during working years together with retirement benefits. The framework did not allow for positive real interest rates or economic growth. It also ignored uncertainties of life.
2. Much of the disagreement about Social Security can be traced to a debate about its value in terms of providing insurance over uncertain lifetimes and the ability of individuals to act in their own interests when making consumption and saving choices.

Exercises

1. Give two reasons why there is disagreement about Social Security reform.
2. What does it mean not to have access to credit markets?
3. What other government programs are facing budgetary problems? Are the sources of these problems the same as Social Security?

This page titled 28.6: Social Security in the Real World is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

28.7: End-of-Chapter Material 28.7: End-of-Chapter MaterialEdit Update titlecancel

In Conclusion

Throughout the world, people contribute to and benefit from social security programs like that in the United States. Yet, owing to demographic changes and other factors, the US Social Security system as we currently know it is unlikely to survive. The challenges faced by the United States are present in many other countries with similar demographics. In much of the developed world, the ratio of workers to retirees will decrease over the next decades. Armed with the tools of this chapter, you are now equipped to understand the implications of proposed changes to Social Security programs, both in the United States and the rest of the world.

Our analysis of Social Security combines two tools often used in macroeconomics. The first is the life-cycle model of consumption/saving, which provides insights into how individuals and households make consumption and saving decisions over long time horizons. We saw that people do not have to match their consumption to their spending each year; instead they can save or borrow to keep their consumption relatively smooth over their lifetimes. However, they must still satisfy a budget constraint over their entire lifetime.

The second is the government budget constraint. We first examined the case where the government kept the Social Security system in balance. In this case, revenues and payments were equal each year. Then we examined the case where the government did not necessarily match revenues and spending. In this case, there is still an accounting of government flows that links surpluses and deficits today with future obligations.

Our discussion illustrates a very important fact about how the economy works: household behavior typically responds to government policy. In the case of Social Security, we saw that households reduce their saving when the government saves on their behalf.

Key Links

- Obama Council of Economic Advisors on Social Security: www.whitehouse.gov/issues/seniors-and-social-security
- Social Security Administration
  - Tax rates: http://www.ssa.gov/OACT/ProgData/taxRates.html
  - Tax base limits: http://www.ssa.gov/OACT/COLA/cbb.html#Series
  - Programs in other countries: http://www.ssa.gov/international/links.html
- The Central Provident Fund in Singapore: mycpf.cpf.gov.sg/Members/home.htm
- Social security in China: www.gov.cn/english/official/2005-07/28/content_18024.htm

exercises

1. Suppose that disposable income is $50,000, working years is 50, retirement years is 20, and the Social Security payment is $20,000. What is the lifetime income for this household?
2. Suppose a household lives for two periods, working and earning disposable income of $10,000 in the first period and obtaining retirement income of $5,000 in the second period. Suppose that the real interest rate is not 0 percent (as in our example of Carlo) but rather is 10 percent. What is the discounted present value of the household's lifetime income? (Refer to the toolkit if you need a reminder of how to calculate a discounted present value.) How would you write the lifetime budget constraint when the real interest rate is not 0 percent?
3. Some rapidly growing countries, such as China, have a very high saving rate. Everything else being the same, explain why a household in a rapidly growing economy would tend to have a low and not a high saving rate. The social security system in China is not very generous. Explain how this would help you to understand the high saving rate in China.
4. Using the life-cycle model, how would the level of consumption respond to an increase in
   1. retirement relative to working years?
   2. the annual labor income during working years?
   3. payments of Social Security during retirement relative to income earned during working years?
5. The equation for lifetime earnings is key to understanding the effects of Social Security. Explain in your words why the last two terms on the right side of that equation disappear using the government budget constraint.

6. Suppose you expect to live for 50 more years. Suppose also that, because the company you work for had a successful year, you get a $50,000 bonus. If you smooth your consumption perfectly, how much of your bonus will you spend this year, and how much will you save? (You can assume the real interest rate is zero.)
7. Suppose you expect to live for 50 more years. Suppose also that, because you have done an excellent job this year, you get a $2,000 raise. This means you expect that your income will be $2,000 higher every year. If you smooth your consumption perfectly, how much of this raise will you spend this year, and how much will you save? (You can assume the real interest rate is zero.)
8. Suppose you expect to live for 50 more years. Suppose also that, because the company you work for had a successful year, your boss tells you (and you believe her!) that you will get a $50,000 bonus one year from now. If you smooth your consumption perfectly, what will happen to your consumption and saving this year? (You can assume the real interest rate is zero.)
9. Why do you think that the Singaporean government allows people to withdraw funds from the government saving scheme in order to buy a house or apartment but not in order to take a vacation?
10. Suppose a government institutes a pay-as-you-go social security scheme. Explain why the first generation of recipients are clear beneficiaries from the scheme.
11. Give two reasons why households do not smooth their consumption perfectly.

Economics Detective

1. Find the most recent Social Security Administration release. What is the current status of the program? When is it forecasted to go bankrupt?
2. Pick a country other than the United States. What is the social security system like there? What is its current status?
3. Go to http://www.ssa.gov/OACT/COLA/cbb.html#Series. What does the contribution and benefits base mean? Using the correcting for inflation tool, what has happened to the contribution and benefit bases in real terms over the past 20 years?
4. Go to the Social Security Administration (http://www.ssa.gov/pubs/10070.html) to figure out how to calculate the benefits for someone about to retire in your group of family or friends.

Spreadsheet Exercises

1. Consider the life of Carlo, as summarized in Figure 28.7.1. Write a spreadsheet program to reproduce the calculations of lifetime income and consumption made in that figure. Introduce a real interest rate of 5 percent into your program. Recalculate the discounted present value of lifetime income. What will Carlo consume each period of his life?
2. Use your spreadsheet program from Problem 1 to determine how changes in Social Security affect consumption and saving. Do this first with a real interest rate of 0 and then with a 5 percent real interest rate. Compare your results.

This page titled 28.7: End-of-Chapter Material is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

29: Balancing the Budget

29.1: $24,211,075,873,711.89
29.2: Deficits and Debt
29.3: The Causes of Budget Deficits
29.4: The Benefits of Deficits
29.5: The Costs of Deficits
29.6: The Ricardian Perspective
29.7: End-of-Chapter Material

This page titled 29: Balancing the Budget is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

29.1: $24,211,075,873,711.89 29.1: $24,211,075,873,711.89Edit Update titlecancel

…is a big number. It is the total amount of US government debt outstanding as of April 13, 2020. This number, which changes every day, is reported by the US Treasury Department at the Treasury Direct website (www.treasurydirect.gov/NP/BPDLogin?application=np).

The debt of the United States is the subject of a growing political storm in Washington. Indeed, in August 2011 there seemed to be a very real possibility that the US Congress would refuse to raise the "debt ceiling"—an upper limit on the size of the government debt. Had that occurred, the government would no longer have been able to fulfill all its obligations. Many commentators believe that the US government is facing a crisis with respect to its budget policies—specifically, the fact that the government is running persistent budget deficits. The issues are not the stuff of dry academic debate. If you are a typical reader of this book, you will be working and paying taxes over the next 50 years. *Yours* is the future generation that will be called on to deal with present-day deficits; debates about government deficits today are debates about your standard of living. If deficits matter to anyone, they should matter to you.

Just like a household, a government has income and outlays. If a household's outlays exceed its income, then it must borrow to finance its spending. And if a household borrows repeatedly, it builds up debt. The same is true of governments. If a government spends more than its income, then it is running a deficit that must be financed by borrowing. Repeated government deficits lead to the existence of a stock of government debt.

In recent decades, the US federal government has run a deficit much more often than not. The federal government has been in deficit for all but 4 years between 1960 and 2011. As a consequence, the stock of debt outstanding in the United States has increased from $290 billion to more than $14 trillion.

Most of us cannot really conceptualize what this sum means. We can try visual images: if we stacked up 14 trillion dollar bills, we would get a pile half a million miles high—more than twice the distance to the moon. But it is easiest to get a handle on the deficit if we divide by the number of people in the economy to obtain the debt per person. As of August 9, 2011, according to the US National Debt Clock (www.brillig.com/debt_clock), this number is $46,905.36. This means that if the government wanted to pay off its debt today, each and every woman, man, and child in the United States would have to be taxed by this amount, on average, to pay off the obligations of the government.

US citizens hold more than half of the debt—about 60 percent. So if the government were to pay off its debt, the majority would end up being redistributed in the economy from taxpayers to holders of US government bonds. Foreigners hold the remaining 40 percent, so this money would be transferred from US taxpayers to citizens of other countries. The US government is not proposing to pay off the existing debt, however. To the contrary, the government is projected to run budget deficits for the foreseeable future, meaning that the stock of debt, and the obligation of future generations, will continue to grow. These forecasts are available from the Congressional Budget Office (CBO; http://www.cbo.gov).

In response to concern over government deficits, one proposal has arisen over and over again: a balanced-budget amendment to the US Constitution. Such a measure would simply make deficits illegal. A balanced-budget amendment came within one vote of passing in a 1997 US Senate vote, and one was passed by the US House of Representatives in 1997. Another bill was introduced by a group of US House members in 2003. Here is part of the text of the 2003 bill:

SECTION 1. Total outlays for any fiscal year shall not exceed total receipts for that fiscal year, unless three-fifths of the whole number of each House of Congress shall provide by law for a specific excess of outlays over receipts by a rollcall vote.

SECTION 2. The limit on the debt of the United States held by the public shall not be increased, unless three-fifths of the whole number of each House shall provide by law for such an increase by a rollcall vote.

SECTION 3. Prior to each fiscal year, the President shall transmit to the Congress a proposed budget for the United States Government for that fiscal year in which total outlays do not exceed total receipts. Although such bills are typically termed "balanced-budget amendments," they often, as is the case here, permit surpluses. US House of Representatives, "H.J.RES.22—Proposing a Balanced-Budget Amendment to the Constitution of the United States," February 13, 2003, accessed July 20, 2011, thomas.loc.gov/cgi-bin/query/z?c108:H.J.RES.22:.

The Tea Party, which rose to some political prominence in the United States in 2010, campaigned in favor of a balanced-budget amendment as well. In July 2011, the House of Representatives passed HR 2560, called the Cut, Cap, and Balance Act, which (among other things) called for a constitutional amendment to balance the budget to be transmitted to the states for their

consideration.The Cut, Cap, and Balance Bill is presented at “Bill Text Versions: 112th Congress (2011–2012) H.R.2560,” THOMAS: The Library of Congress, accessed September 20, 2011, thomas.loc.gov/cgi-bin/query/z?c112:H.R.2560: For ongoing discussion, read Rep. Mike Coffman, “Balanced Budget Amendment Caucus,” accessed July 20, 2011, coffman.house.gov/index.php?option=com_content&view=article&id=257&Itemid=10. This bill was not passed by the Senate. Whether this political activity will ever generate a constitutional amendment remains an open question and a point of debate in the 2012 election.

The discussion of constitutional limits on budget deficits is not limited to the United States. In 2009, Germany amended its constitution to limit federal budget deficits to 0.35% of GDP by 2016. This limit applies when the German economy is operating near its potential output. The regulations allow the German government to run deficits during recessions but require surpluses in times of high economic activity. The fiscal situation in Germany is described in “Reforming the Constitutional Budget Rules in Germany,” Federal Ministry of Finance—Economics Department, September 2009, accessed September 20, 2011, http://www.kas.de/wf/doc/kas_21127-1522-4-30.pdf?101116013053.

Should the government be forced to balance its budget each year, as such measures suggest? There are certainly good reasons why households sometimes incur debt—to pay for a house, a new car, or advanced education. Perhaps the same is true of governments. We should not presume that deficits are harmful without first trying to understand why they occur. Others have even argued that deficits are neither good nor bad but are simply unimportant. Indeed, Vice President Cheney is reported to have said that “[President] Reagan proved that deficits don’t matter.” Quoted in Ron Suskind, *The Price of Loyalty: George W. Bush, the White House, and the Education of Paul O’Neill* (New York: Simon and Schuster), 291.

So are deficits bad for the economy, good for the economy, or just irrelevant? Our goal in this chapter is to understand the economic effects of government budget deficits so that we can evaluate competing claims such as these and ultimately help you answer the following question.

Should the government be forced to balance its budget?

Road Map

We go through five steps in our evaluation of the merits of a balanced-budget amendment:

1. We make sure that we know what we are talking about. “Debt” and “deficit” are technical terms with precise meanings. We go through their definitions carefully.
2. We examine the causes of the deficit in an accounting sense. Specifically, we examine how and why the budget deficit depends on the state of the economy. We can then explore the implications—again in an accounting sense—of a balanced-budget law.
3. We progress to a deeper understanding of why deficits occur. We examine *why* governments choose to run deficits. At this point, we examine possible benefits of deficits to the economy.
4. We examine why deficits might be harmful to the economy.
5. We examine the argument for why deficits might be irrelevant.

This page titled 29.1: $24,211,075,873,711.89 is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

29.2: Deficits and Debt 29.2: Deficits and DebtEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What is the difference between the deficit and the debt?
2. What are the links between the deficit and the debt?
3. What are the budget constraints faced by the government?

We begin by being careful and precise about terminology. The terms *deficit* and *debt* are sometimes used sloppily in everyday discourse; as a consequence, much nonsense is spoken about fiscal policy. We must first make sure that we understand exactly what these terms mean. The CBO (http://www.cbo.gov/showdoc.cfm?index=6060&sequence=13) has a glossary of terms on its web page.

Budget Deficit: Definition

The **government deficit** is the difference between **government outlays** and **government revenues**. Inflows and outflows are part of the circular flow of income. Revenues flow to the government when it imposes taxes on households and firms and when it collects money through various other fees. For our purposes here, we do not need to distinguish all the different kinds of taxes, and we do not worry about whether they are paid by firms or by households. All that matters is that, in the end, some of the income generated in the economy flows to the government. The government also collects Social Security payments, which are discussed in more detail in Chapter 28. These are just another kind of tax.

Money flows out in the form of government purchases of goods and services and government transfers. **Government purchases** include things like roads, streetlamps, schools, and missiles. They also include wage payments for government employees—that is, the purchase of the services of teachers, soldiers, and civil servants. Outlays also occur when government gives money to households. These are called transfer payments, or **transfers** for short. Examples include unemployment insurance, Social Security payments, and Medicare payments. Finally, transfers include the interest payments of the government on its outstanding obligations.

The outlays of the government and its revenues are not always equal. The difference between government purchases and transfers and government revenues represents a government deficit, as set out in the following definition:

$$\begin{aligned} &government\ deficit = outlays–revenues = government\ purchases + transfers - taxrevenues \\ &= government\ purchases - (tax\ revenues - transfers) = government\ purchases - net\ taxes. \end{aligned}$$

Often we find it useful to group taxes and transfers together as “net taxes” and separate out government purchases, as in the last line of our definition.

When outflows are less than inflows, then we say there is a **government surplus**. In other words, a negative government deficit is the same thing as a positive government surplus, and a negative government surplus is the same thing as a positive government deficit:

$$government\ surplus = -government\ deficit.$$

A government surplus is sometimes called “government savings.” When the government runs a deficit, borrowing from the financial markets funds such spending. When the government runs a surplus, these funds flow into the financial markets and are available for firms to borrow.

To illustrate the calculation of the deficit, we examine some made-up numbers in Table 29.2.1 "Calculating the Deficit". Our equation defining the deficit tells us that we can calculate it two ways. Look, for example, at year 3. The level of government spending is 200, tax receipts are 160, and transfers are 20.

- We can add together purchases and transfers to get total outlays of the government, which is 220. Then we can subtract revenues of 160 to find that the deficit is 60.
- We can subtract transfers from tax receipts to get the amount of net taxes. Here, net taxes are 140. We subtract this from purchases of 200 to find a deficit of 60.

Obviously, we get the same answer either way; it is just a matter of how we group the different terms together. It might seem natural to group transfers with government expenditures (since they are both outlays). Conceptually, though, transfers are more like taxes, in that they represent a flow of dollars that is not matched by a flow of goods or services. The difference is that taxes flow from into the government; transfers flow the other way. Government expenditures are very different: they represent purchases of real gross domestic product (real GDP) produced in the economy, thus contributing to the overall demand for output.

Table 29.2.1: Calculating the Deficit

Year	Government Purchases	Tax Revenues	Transfers	Net Taxes	Deficit
1	50	30	10	20	30
2	100	160	40	120	−20
3	200	160	20	140	60
4	200	220	20	200	0
5	140	160	20	140	0

In Table 29.2.1 "Calculating the Deficit", the deficit varies considerably over time. It is low in year 1, negative in year 2 (in other words, there is a surplus), high in year 3, and zero in years 4 and 5. Between year 1 and year 2, government purchases and transfers increased, but tax revenues increased even more. In fact, they increased sufficiently to turn the deficit into a surplus. Between years 2 and 3, government purchases increased, and transfers decreased. However, the decrease in transfers was less than the increase in government purchases, so total government outlays increased substantially. Tax revenues stayed constant, so the government went back into deficit.

In years 4 and 5 the government ran a balanced budget. If we compare year 4 to year 3, we see that the budget could be balanced by raising taxes (from 160 to 220) and leaving outlays unchanged. Conversely, by comparing year 5 to year 3, we see that the budget could be balanced by cutting spending and leaving taxes unchanged. A balanced budget is consistent with high taxes and high spending or low taxes and low spending. It is the combination of low taxes and high spending that give us a deficit. Table 29.2.1 "Calculating the Deficit" makes it clear that changes in the deficit can be explained only by examining all components of the government budget constraint.

The Single-Year Government Budget Constraint

We begin with the government budget constraint as it operates in a single year. This budget constraint can be seen in terms of the flows into and from the government sector in the circular flow, as shown in Figure 29.2.1 (which explicitly shows that taxes come from households and firms). Later we discuss a second government budget constraint that links spending and revenues over longer periods of time.

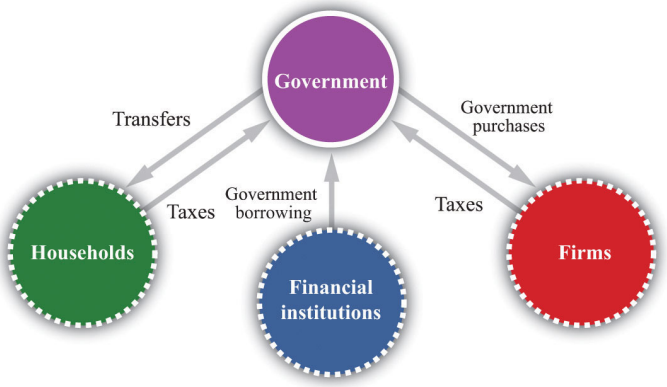


Figure 29.2.1: The Government Sector in the Circular Flow. The inflows into the government sector come from taxes and borrowing from the financial sector. The outflows comprise government purchases and government transfers.

You might be wondering how it is possible for the government to have outlays that exceed its revenues. The answer is given by the government budget constraint. The **government budget constraint** says that the deficit, which is the difference between outlays

and revenues, must be financed by borrowing. If outlays exceed revenues in a given year, then the government must somehow make up the difference. It does so by borrowing from the public. In this sense, the government is no different from a household. Each of us can, like the government, spend more than we earn. When we do, we must either borrow from someone or draw on our savings from the past.

The government borrows by issuing **government debt**. This debt can take several forms. The government has many types of obligations, ranging from short-term Treasury Bills to longer-term bonds. For our analysis, we do not need to distinguish among these different assets.

Toolkit: Section 31.33 and Section 31.27

You can review the government budget constraint and the circular flow of income in the toolkit.

The Deficit: Recent Experience

Table 29.2.2 "Recent Experience of Deficits and Surpluses (Billions of Dollars)" shows some actual numbers for the United States: receipts, outlays, and the federal budget deficit in current dollars for fiscal years 1990 to 2010.Government budget numbers in the United States are reported for a “fiscal year,” which runs from October to September. Thus fiscal year 2000 ran from October 1, 1999, to September 30, 2000.

Table 29.2.2: Recent Experience of Deficits and Surpluses (Billions of Dollars)

Fiscal Year	Receipts	Outlays	Surplus or Deficit (−)
1990	1,032.0	1,253.1	−221.0
1991	1,055.1	1,324.3	−269.2
1992	1,091.3	1,381.6	−290.3
1993	1,154.5	1,409.5	−255.1
1994	1,258.7	1,461.9	−203.2
1995	1,351.9	1,515.9	−164.0
1996	1,453.2	1,560.6	−107.4
1997	1,579.4	1,601.3	−21.9
1998	1,722.0	1,652.7	69.3
1999	1,827.6	1,702.0	125.6
2000	2,025.5	1,789.2	236.2
2001	1,991.4	1,863.2	128.2
2002	1,853.4	2,011.2	−157.8
2003	1,782.5	2,160.1	−377.6
2004	1,880.3	2,293.0	−412.7
2005	2,153.9	2,472.2	−318.3
2006	2,406.9	2,655.1	−248.2
2007	2,568.0	2,728.7	–160.7
2008	2,524.0	2,982.5	–458.6
2009	2,105.0	3,517.7	–1,412.7
2010	2,161.7	3,455.8	–1,294.1

Source: “Historical Budget Tables,” Congressional Budget Office, January 2011, accessed September 20, 2011, http://www.cbo.gov/ftpdocs/120xx/doc12039/HistoricalTables[1].pdf.

In the early 1990s, the government ran a deficit of about $200–300 billion every year. (Note that a negative number in the last column corresponds to a government deficit.) In the mid-1990s, however, the deficit began to decrease. Both outlays and receipts were increasing, but receipts were increasing more quickly. By 1998, the federal budget was in surplus, and it reached a peak of $236 billion in 2000. Thereafter, revenues decreased for several years, while spending continued to increase. By 2002, the budget had gone back into deficit again, and by the middle of the decade, the deficit was at record levels.

As is evident from Table 29.2.2 "Recent Experience of Deficits and Surpluses (Billions of Dollars)", the budgetary picture changed dramatically with the onset of the severe recession in 2008. Revenues decreased and outlays increased so that the budget deficit widened considerably, to more than $1 trillion in both 2009 and 2010.

If you look at data on the government budget, you will see that the federal budget is divided into “on-budget” and “off-budget” items. Table 29.2.3 "On-Budget, Off-Budget, and Total Surplus, 2010 (Billions of Dollars)" shows these numbers for fiscal year 2010. The Congressional Budget Office defines off-budget items as follows. “Spending or revenues excluded from the budget totals by law. The revenues and outlays of the two Social Security trust funds (the Federal Old-Age and Survivors Insurance Trust Fund and the Disability Insurance Trust Fund) and the transactions of the Postal Service are off-budget.” Congressional Budget Office, Glossary, accessed October 19, 2011, www.cbo.gov/doc.cfm?index=2727&type=0&sequence=14

The transactions of the US Postal Service are not that important, so you can essentially think of the off-budget items as being the Social Security system. Since the Social Security system was in surplus over much of this period, the on-budget deficit is larger than the total. From Table 29.2.3 "On-Budget, Off-Budget, and Total Surplus, 2010 (Billions of Dollars)", the total government deficit of $1,294 billion in 2010 reflects an on-budget deficit and a small off-budget surplus.

The idea behind the separate budgeting is that Social Security represents a known set of future government obligations. For this reason the government has, in effect, set aside a separate account for Social Security revenues and outlays (much as you, as an individual, might decide you want a separate account for your savings). We discussed the Social Security Trust Fund, as this account is called, in Chapter 28. At least in theory, this separates the debate about Social Security from the debate about current government spending and receipts. Many policy discussions do focus just on the “on-budget” accounts. In the end, though, all these monies flow either into or from the federal government. The humorist Dave Barry once remarked that what distinguishes off-budget items is that “*these are written down on a completely different piece of paper from the regular budget*.” Dave Barry, “The Mallomar Method,” DaveBarry.com, March 24, 1991, accessed August 28, 2011, http://www.davebarry.com/misccol/mallomar.htm. What is more, there are other known future obligations, such as Medicare, that are not treated separately. The on-budget/off-budget distinction is really no more than an accounting fiction, and in terms of the overall economic effects of the deficit, it is better to focus on the total.

Table 29.2.3: On-Budget, Off-Budget, and Total Surplus, 2010 (Billions of Dollars)

	Receipts	Outlays	Surplus or Deficit (−)
On-Budget	1,530.1	2,9091.1	−1,371.1
Off-Budget	631.7	554.7	77.0
Total	2,161.7	3,455.8	−1,294.1

Source: US Treasury, Financial Management Service, October 2010 Statement, accessed September 20, 2011, www.fms.treas.gov/mts/mts0910.txt.

There are mixed messages to take away from Table 29.2.2 "Recent Experience of Deficits and Surpluses (Billions of Dollars)". The experience of budget surpluses in the 1990s tells us that budget balancing is possible. At the same time, more recent experience suggests that substantial changes in receipts and/or outlays are now needed to balance the budget. To explore this somewhat further, look at Table 29.2.4 "Federal Outlays, 2010 (Billions of Dollars)", which shows various outlays for 2010. As we already know, total spending for that year was $3.5 trillion. National defense, Social Security, and health-care programs together account for $2.2 trillion, or about 63 percent of the total outlays. Other nondiscretionary spending—largely outlays such as retirement payments to federal employees, unemployment insurance, housing assistance, and food stamps—accounts for a further $401 billion. Interest payments account for $196 billion. These categories together account for more than 80 percent of federal outlays.

Table 29.2.4: Federal Outlays, 2010 (Billions of Dollars)

Item	Amount	Total Outlays (%)
Defense	689	19.9
Nondefense Discretionary Spending	658	19.0
Social Security	701	20.3
Health Care Programs (including Medicare and Medicaid)	810	23.4
Other Nondiscretionary Spending	401	11.6
Interest Payments	196	5.7
Total	3,456	100.0

Source: Compiled from data in CBO, “The Budget and Economic Outlook: An Update,” August 2011, accessed September 20, 2011, www.cbo.gov/ftpdocs/123xx/doc12316/08-24-BudgetEconUpdate.pdf. Totals do not add up because of rounding errors.

Just looking at those numbers should make it clear that it is very difficult to balance the budget simply by cutting federal spending. Almost everyone agrees that there is waste in the federal government, and there are programs that could and almost certainly should be abolished. (This is not to say that you could find even a single program that everyone would want to abolish. Every program benefits someone, after all. But there are certainly programs that most people would agree are wasteful.) However, the vast majority of the budget is taken up with either essential functions of government or programs that enjoy huge political popularity. Few politicians would sign up for closing the public schools, the abolition of unemployment insurance, or the cancellation of veterans’ benefits.

The budget accounts distinguish between mandatory and discretionary spending. Many of the big items listed in Table 29.2.4 "Federal Outlays, 2010 (Billions of Dollars)" fall into the mandatory category—that is, outlays that are required by existing law. Less than 40 percent of outlays in 2010 were discretionary, and half of those were national defense spending. The remaining outlays were mandatory spending or payment of interest on the outstanding debt. “Budget and Economic Outlook: Historical Budget Data,” Congressional Budget Office, January 2011, accessed July 20, 2011, http://www.cbo.gov/ftpdocs/120xx/doc12039/HistoricalTables[1].pdf.

If the government were to pass a balanced-budget amendment, in other words, the hard job of cutting spending or raising taxes would remain. Recall Section 3 of the amendment that we quoted in the chapter opener: “the President shall transmit to Congress a proposed budget…in which total outlays do not exceed total receipts.” Even with a balanced-budget amendment, the president would still have to propose either major cuts in existing popular programs or increases in taxes. However, such an amendment might provide “political cover” for the president and Congress: they could explain their support for unpopular spending cuts or tax increases by saying that the balanced-budget amendment gave them no choice.

The Intertemporal Government Budget Constraint

We discussed in 29.2 Section "The Single-Year Government Budget Constraint" that the single-period government budget constraint links spending and revenues to the deficit (or surplus) of the government each year. There is a second constraint faced by the government, called the **intertemporal budget constraint**, linking deficits in one year to deficits in other years.

When you take out a loan, you will ultimately have to repay it. The same is true of the government; when it takes out a loan, it will ultimately have to repay the loan as well. If the government chooses to pay for its expenditures today by borrowing instead of through current taxes, then it will need additional taxes at some point in the future to pay off its loan. The intertemporal budget constraint is just a fancy way of saying that, like everyone else, the government has to pay off its loans at some point. Actually, there is one way in which the government is different from private individuals. For practical purposes, we expect that the government will go on forever. This means that the government could always have a stock of outstanding debt. However, there are practical limits on this stock—for one thing, households will not lend unlimited amounts to the government. Thus it is generally fair to say that additional borrowing by the government will have to be repaid. As a consequence, tax and spending decisions at different dates are linked. Although governments can borrow or lend in a given year, the government’s total spending over time must be matched by revenues.

To express the intertemporal budget constraint, we introduce a measure of the deficit called the **primary deficit**. The primary deficit is the difference between government outlays, *excluding interest payments on the debt*, and government revenues. The **primary surplus** is equal to the minus of the primary deficit and is the difference between government revenues and government outlays, excluding interest payments on the debt. In our example in Table 29.2.1 "Calculating the Deficit", the deficit in year 1 was 30. If payment of interest on outstanding debt was 5, then the primary deficit would be 25, and the primary surplus would be −25.

The intertemporal budget constraint says that if the government has some existing debt, it must run surpluses in the future so that it can ultimately pay off that debt. Specifically, it is the requirement that

$$current\ debt\ outstanding = discounted\ present\ value\ of\ future\ primary\ surpluses.$$

This condition means that the debt outstanding today must be offset by primary budget surpluses in the future. Because we are adding together flows in the future, we have to use the tool of discounted present value. If, for example, the current stock of debt is zero, then the intertemporal budget constraint says that the discounted present value of future primary surpluses must equal zero.

Toolkit: Section 31.5

You can review the meaning and calculation of discounted present value in the toolkit.

Linking the Debt and the Deficit

The stock of debt is linked directly to the government budget deficit. When the government runs a budget deficit, it finances the deficit by issuing new debt. The deficit is a *flow*, which is matched by a *change in the stock* of government debt:

$$change\ in\ government\ debt(in\ given\ year) = deficit(in\ given\ year).$$

If there is a government surplus, then the change in the debt is a negative number, so the debt decreases. The total government debt is simply the accumulation of all the previous years' deficits. From this equation, the stock of debt in a given year is equal to the deficit over the previous year plus the stock of debt from the start of the previous year. (In this discussion, we leave aside the fact that the government may finance part of its deficit by issuing new money. In the United States and most other economies, this is a minor source of funding for the government. See Chapter 26 for more discussion. More precisely, then, every year, $change\ in\ government\ debt = deficit - change$ in money supply.Written this way, the equation tells us that the part of the deficit *that is not financed by printing money* results in an increase in the government debt.)

To see the interactions between deficits and the stock of debt in action, examine Table 29.2.5 "Deficit and Debt", which takes the deficit numbers from Table 29.2.1 "Calculating the Deficit" and calculates the corresponding debt. We suppose that there is initially zero debt at the beginning of year 1. The deficit of 30 in the first year means that there is outstanding debt of 30 at the end of that year. In the second year, there is a budget surplus of 20. This reduces the debt, but it is not sufficient to bring the debt all the way back to zero. Outstanding debt at the end of the year is 10. In the third year, the deficit of 60 must be added to the existing debt of 10, so the debt at the end of the year is 70.

Table 29.2.5: Deficit and Debt

Year	Deficit	Debt (Start of Year)	Debt (End of Year)
1	30	0	30
2	−20	30	10
3	60	10	70
4	0	70	70
5	0	70	70

In years 4 and 5, the government runs a balanced budget: the deficit is zero. But the stock of debt stays unchanged. The debt is equal to the accumulation of all the deficits. Eliminating deficits (for example, by a balanced-budget amendment) means that the debt stays at its existing level. *Eliminating deficits is not the same thing as paying off the debt.*

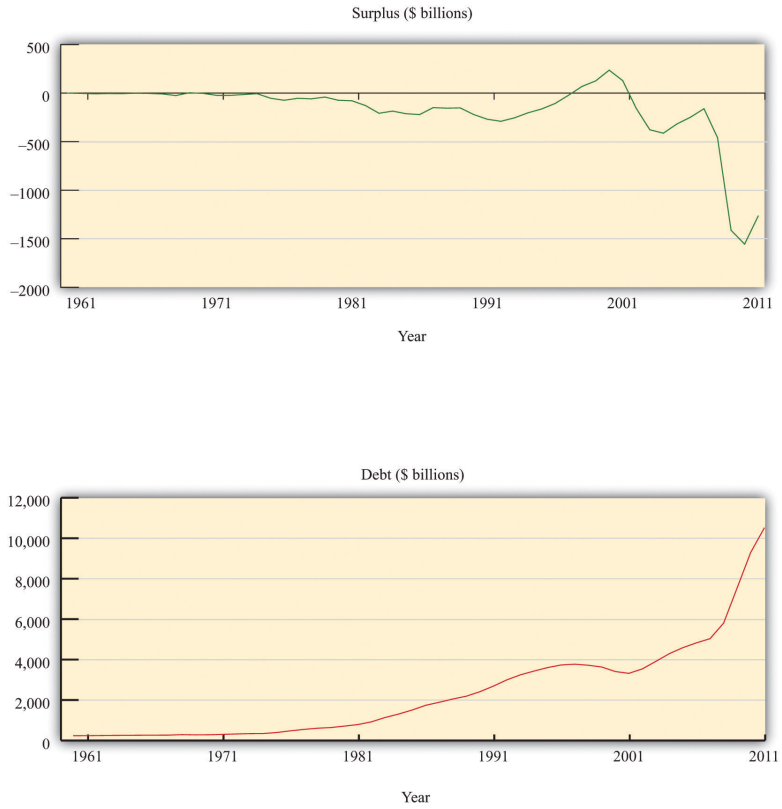


Figure 29.2.2: US Surplus and Debt, 1962–2010. Source: Congressional Budget Office.

The experience of the US deficit and debt held by the public since 1962 is summarized in Figure 29.2.2. The surplus is shown in the upper figure, and the level of debt is shown in the lower figure. All values are in current dollars. At the far left of the graph, we see that the US government ran relatively small deficits (negative surpluses) in the 1960s and early 1970s. As a result, the debt increased slowly. From the mid-1970s to the mid-1990s, deficits were substantial, so the amount of debt outstanding grew rapidly. As we saw earlier, there was a brief period of surplus in the late 1990s and a corresponding decrease in the debt, but deficit spending recommenced during the George W. Bush administration (2001–2008). The debt increased again.

Although an analysis of deficits and debt is often presented using data similar to those in Figure 29.2.2, this figure is incomplete in two ways: (1) these numbers are not corrected for inflation (they are current dollar figures), and (2) there is no sense of how large the deficit and the debt are relative to the aggregate economy. Figure 29.2.3 remedies both defects by showing the surplus and the debt as a fraction of nominal GDP. Because nominal GDP is also measured in dollars, these ratios are just numbers. We see that the deficit has been a relatively stable fraction of GDP, averaging about 2.7 percent of GDP. The debt level has averaged about 36 percent over the period.

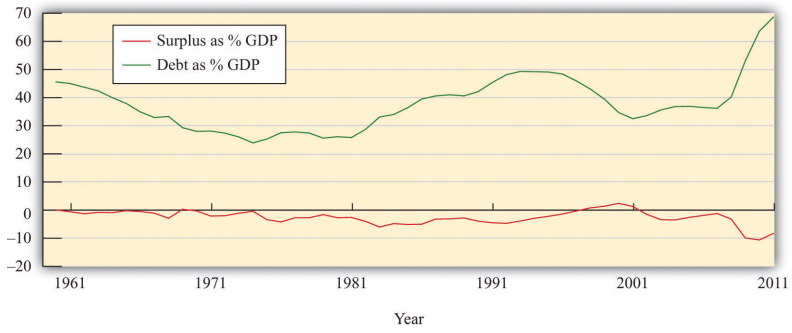


Figure 29.2.3: US Surplus and Debt as a Fraction of GDP, 1962–2010. Source: Congressional Budget Office and *Economic Report of the President.*

The federal debt is now in excess of $14 trillion. So if the United States were to pass a balanced-budget amendment binding on the federal government, to take effect in 2012, say, the stock of debt would thereafter remain fixed at well over $14 trillion. To reduce

the stock of debt outstanding, the deficit must be negative: the change in the stock of debt will be negative only if the government runs a surplus.

Moreover, the government must pay interest on its outstanding debt. Recall that when the government runs up debt, it is borrowing from the general public. The debt of the government is an asset from the perspective of households: it is one of the ways in which people can hold their saving. Holders of government bonds earn interest on these assets. Look again at Table 29.2.4 "Federal Outlays, 2010 (Billions of Dollars)". In the United States, interest payments on the debt amounted to $184 billion in 2005. Interest payments on the debt amount to more than half of the deficit. Balancing the budget therefore means that, once we exclude interest payments, spending plus transfers would have to be much smaller than tax revenues. If there is outstanding debt, a balanced budget means that the government must run a primary surplus.

To summarize, we have discovered three things about a balancing the budget:

1. A balanced budget means that the deficit equals zero.
2. A balanced budget means that the debt is constant.
3. If there is existing debt, a balanced budget means that the government must run a primary surplus.

Who Holds the Debt?

Given that the US government makes such large interest payments on outstanding debt, who receives those payments? US government debt is held by households, firms, and governments in many countries. Table 29.2.6 "Foreign Holdings of US Treasury Securities as of August 2008 (Billions of Dollars)" lists some of the foreign countries holding US Treasury securities (bills, bonds, and notes) in two different months: August 2008 and May 2011.

Table 29.2.6: Foreign Holdings of US Treasury Securities as of August 2008 (Billions of Dollars)

Country	Holding as of August 2008	Holdings as of May 2011
Japan	585.9	912.4
China	541.0	1159.8
oil exporters	179.8	229.8
Mexico	33.5	27.7
Canada	27.7	90.7
total	2,740.3	4,514.0

Source: “Major Foreign Holders of Treasury Securities,” US Department of the Treasury, July 18, 2011, accessed July 20, 2011, http://www.treasury.gov/resource-center/data-chart-center/tic/Documents/mfh.txt.

In May 2011, the total foreign ownership of US Treasury securities was more than 45 percent of the total privately held US public debt (“privately held” means we are excluding debt held by the Federal Reserve System). As you can see from Table 29.2.6 "Foreign Holdings of US Treasury Securities as of August 2008 (Billions of Dollars)", the ownership of US debt has changed significantly over the past few years. Japan was the largest holder of US debt in August 2008, but more recently China has taken its place.

You might wonder how these countries came to hold such a large fraction of US debt. Part of the answer goes back to the interaction between trade and capital flows between the United States and the rest of the world. The key is the link between trade deficits and borrowing from abroad:

$$borrowing\ from\ other\ countries = imports - exports = trade\ deficit.$$

This equation tells us that whenever a country runs a trade deficit, it must finance that deficit by borrowing from abroad. The United States has been running trade deficits since the early 1970s. Consequently, foreign countries have been accumulating US assets, and government debt is one important such asset.

Observers sometimes comment on the fact that a substantial fraction of government debt is “owed to ourselves” (that is, it is held by US citizens) and therefore less of a cause for concern than the fraction that is owned by foreigners. Does this reasoning make sense? The answer is “not very much.” To see why, consider a US citizen who owns some US government bonds. Now imagine

that she sells those bonds to a German bank and uses the proceeds to buy some General Motors (GM) shares that are currently owned by a French investment bank.

All that has happened here is some rebalancing of portfolios. One individual decided to shift her assets around, so she now owns GM shares instead of government bonds. Likewise, the German bank decided it wanted more US bonds in its portfolio, whereas the French investment bank decided it wanted fewer GM shares. These kinds of transactions go on all the time in our economy.

Our hypothetical citizen is just as wealthy as she was before; she is simply holding her wealth in a different form. The same is true for the German and French financial institutions. Yet foreigners hold more of the national debt than previously. Domestic or foreign ownership of the debt can change with no implications for the overall indebtedness of individuals or the country. It is more meaningful to look at the amount of *foreign debt* that has been accumulated by a country as a result of its borrowing from abroad. Foreign debt represents obligations that will have to be repaid at some future date.

Commentators sometimes express worry over the fact that foreign central banks—notably those of Japan and China—own substantial amounts of US debt. There is a legitimate concern here: if one or more of those banks suddenly decided they no longer wanted to hold that debt, then there might be a large change in US interest rates and resulting financial instability. But the real issue is not that the debt is foreign owned. Rather, it is that a large amount of debt is held by individual institutions big enough to move the market.

At the same time, the Chinese are equally concerned about the value of the US government debt they hold. In their view, they traded away goods and services for pieces of paper that are claims to be paid by the US government. These claims are in nominal terms (in dollars). Hence any change in the exchange rate changes the value of this debt to the Chinese. If, for example, the dollar depreciates relative to the Chinese renmimbi (RMB), then the real value (in terms of Chinese goods and services) of this debt is reduced.

The RMB/dollar exchange rate was 8.28 in January 2000. A holder of a US dollar bill could obtain 8.28 RMB in exchange. This rate was 8.07 in January 2006. However, by June 2011, the exchange rate was 6.48. This means that someone who exchanged RMB for dollars in 2000 and then sold those dollars for RMB in June 2011 lost about 20 percent in nominal terms.

Key Takeaways

1. The deficit is the difference between government outlays and government revenues. It is a flow. The debt is a measure of the stock of outstanding obligations of the government at a point in time.
2. The change in the debt between two dates is equal to the deficit incurred during the time between those two dates.
3. The government faces a single-year constraint that its deficit must be financed by issuing new debt. The government also faces an intertemporal budget constraint that its debt at a point in time must equal the discounted present value of future primary surpluses.

Exercises

1. What is the difference between the budget deficit and the primary deficit?
2. If the government runs a surplus, does this mean the stock of debt must be negative?
3. Is it legal for residents of other countries to hold US debt?
4. Table 29.2.2 "Recent Experience of Deficits and Surpluses (Billions of Dollars)" is in current dollars. What does that mean?

This page titled 29.2: Deficits and Debt is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

29.3: The Causes of Budget Deficits 29.3: The Causes of Budget DeficitsEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. How does fiscal policy affect the budget deficit?
2. How does the state of the economy affect the budget deficit?
3. How do we determine whether a budget deficit results from fiscal policy or the state of the economy?

Now that we have defined budget deficits, budget surpluses, and the government debt, it is time to examine what determines these economic variables. The budget deficit reflects two forces: the stance of fiscal policy and the state of the economy.

Fiscal policy refers to the choice by the government of (1) its levels of spending on goods and services, (2) its transfers to households, and (3) the tax rates it sets on households and firms. Most countries have different levels of government, so some tax and spending decisions are made for the whole country, whereas others are made locally. In principle, we can include all levels of government in our discussion. This means that, in the United States, "government" can refer to the totality of local government, state government, and the federal government. In practice, though, it is the decisions of the federal government that have the main impact on the overall fiscal policy of the country. The same is true in other countries—local government decisions are not usually very important for the overall stance of fiscal policy.

Tools of Fiscal Policy

There are two aspects of fiscal policy: government spending and tax/transfer policy. These fiscal policy choices determine the deficit.In other chapters we examine the effects of government spending on the aggregate economy. For example, Chapter 22 explained how changes in government spending can sometimes be used to stimulate the overall economy.

Government Spending

Over long periods of time, government spending increases as an economy gets richer. Over shorter periods of time, however, the level of government spending is not closely influenced by the overall level of economic activity. For this reason, we typically suppose that government spending is an **exogenous variable** that is determined "outside" our framework of analysis. We illustrate this in Figure 29.3.1.

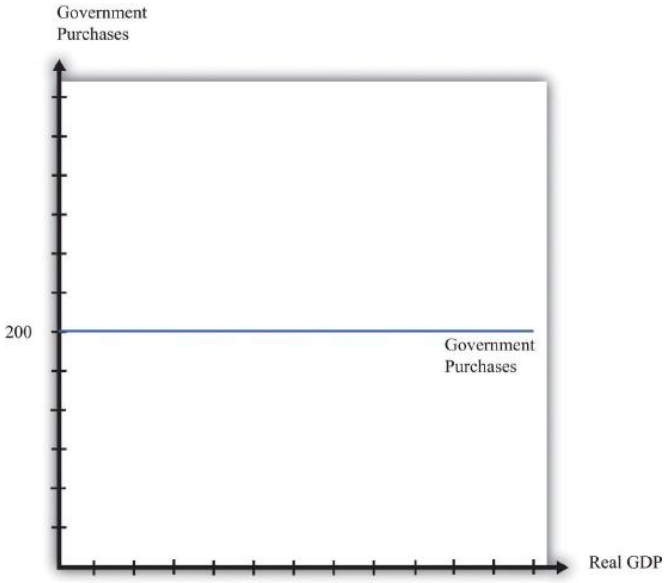


Figure 29.3.1: Government Spending. We suppose that government spending is independent of the level of gross domestic product (GDP), which means that it shows up as a horizontal line.

Taxation

Our interest here is in deficits and the debt rather than the details of taxation, so we take a very simple approach to taxation. We assume that there is a constant tax rate that applies to all levels of income and abstract away from all the other complexities of the tax schedule. This view of the tax and transfer system is summarized by the following equation:

$$net\ taxes = tax\ rate \times income.$$

We illustrate this relationship in Figure 29.3.2. The slope of the line is the tax rate. In other words, for every dollar increase in income, net tax receipts increase by the amount of the tax rate.

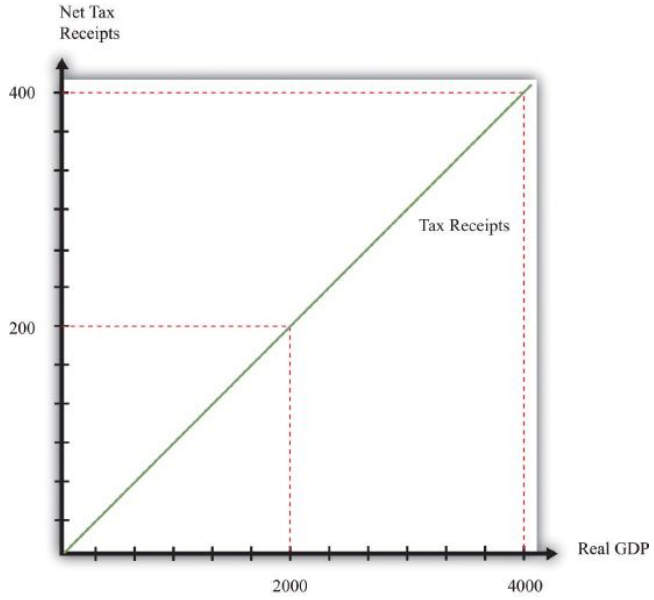


Figure 29.3.2: *The Tax Function*. Net tax receipts depend on the state of the economy. When income is higher, the government collects more in taxes and pays out less in transfers.

Taxes depend positively on income because of the way the tax code is written. Conversely, transfers (such as unemployment insurance or Medicare payments) tend to depend negatively on income: when people are richer, they are less likely to need transfers from the government. The tax rate in the figure captures the overall effect: higher income increases net tax revenues both because people pay more taxes and because they receive fewer transfers.

Table 29.3.1 "Tax Receipts and Income" provides an example of tax receipts at different levels of income, when the tax rate is 10 percent. At the level of an individual household, taxes increase and transfers decrease as the household's income increases. At the level of the entire economy, exactly the same thing is true. As real GDP increases, tax receipts increase and transfers decrease. Increased income, holding the tax rate fixed, leads to increased tax receipts. At the same time, increases in the tax rate lead to higher tax receipts at each level of income. Thus there are two factors determining tax receipts in the economy: the tax rate and the overall level of economic activity.

Table 29.3.1: Tax Receipts and Income

Income	Tax Rate	Tax Receipts
0	0.1	0
100	0.1	10
500	0.1	50
1,000	0.1	100
2,000	0.1	200
5,000	0.1	500

The Budget Deficit and the State of the Economy

As the level of economic activity—real GDP—increases, the tax receipts of the government also increase. To determine the deficit, we need to know both the current fiscal policy (as summarized by the level of government purchases and the tax rate) and the level of economic activity. Building on the example in Table 29.3.1 "Tax Receipts and Income", suppose that government purchases are 200 and the tax rate is 10 percent. The relationship between the level of economic activity (GDP) and the deficit is given in Table 29.3.2 "Deficit and Income". In this example, the level of GDP must reach 2,000 before the budget is in balance (Figure 29.3.3).

Table 29.3.2: Deficit and Income

GDP	Government Purchases	Tax Receipts	Deficit
0	200	0	200
100	200	10	190
500	200	50	150
1,000	200	100	100
2,000	200	200	0
5,000	200	500	−300

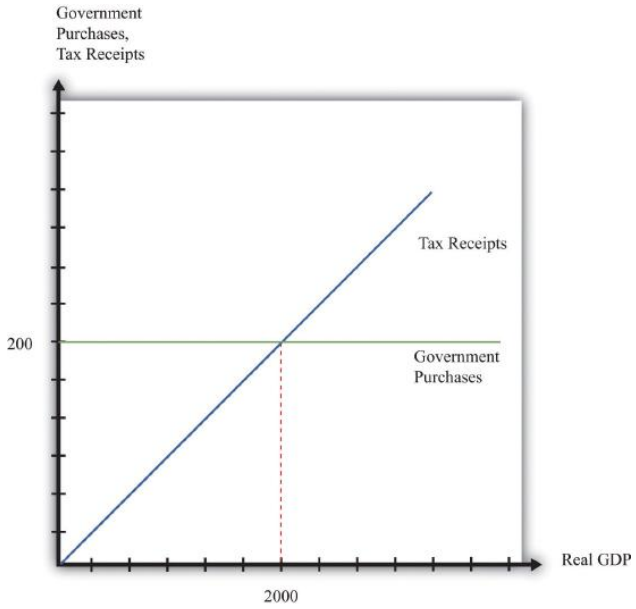


Figure 29.3.3: Government Spending and Tax Receipts. Tax receipts increase as income increases, whereas government spending is unaffected by the level of GDP.

The dependence of the deficit on real GDP and the stance of fiscal policy are summarized in Figure 29.3.4, which graphs the numbers from Table 29.3.2 "Deficit and Income". The deficit/surplus is measured on the vertical axis, and real GDP is measured on the horizontal axis. The deficit/surplus line is drawn for a given tax rate. As real GDP increases, the deficit decreases. Thus the line in Figure 29.3.4 has a negative slope.

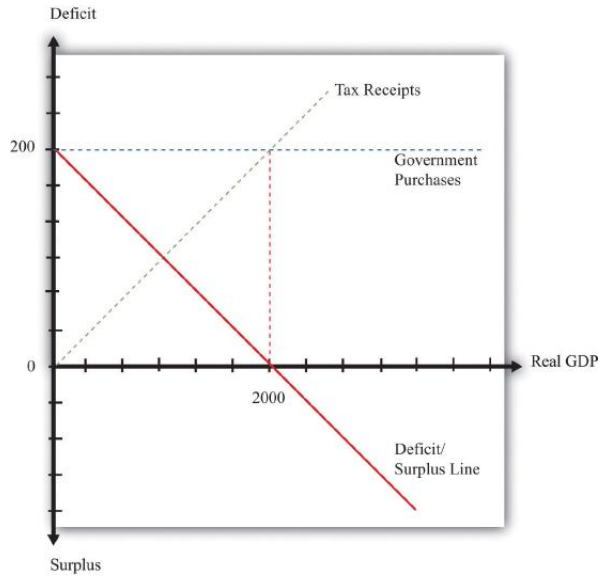


Figure 29.3.4: Deficit/Surplus and GDP. The deficit equals government purchases minus net tax receipts. The deficit is positive when GDP is low, but the budget goes into surplus when GDP is sufficiently high.

The deficit/surplus is the difference between the level of government purchases and the level of receipts. There is a particular level of economic activity such that the budget is exactly in balance. In our example, this level of GDP is 2,000. The deficit is zero when income is 2,000 because that is the point at which government purchases equal tax revenues. For levels of income in excess of this level of GDP, the government budget is in surplus. In Figure 29.3.4, we see that the budget deficit/surplus line crosses the horizontal axis when GDP is 2,000.

Increases in government purchases or reductions in the tax rate are examples of **expansionary fiscal policy**. Decreases in government purchases or increases in the tax rate are called **contractionary fiscal policy**. Expansionary fiscal policy increases the deficit for a given level of real GDP. An increase in government spending shifts the deficit line upward, as shown in Figure 29.3.5. With a decrease in the tax rate, by contrast, the intercept stays the same, but the line rotates upward. The effect is still to increase the deficit at all positive levels of income.

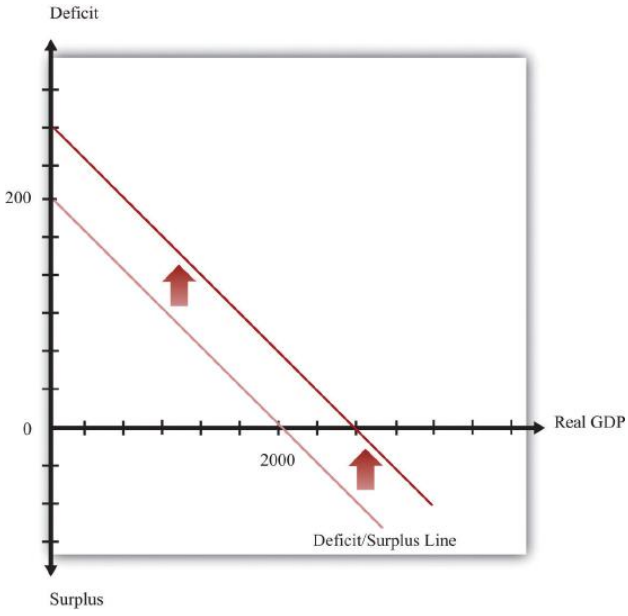


Figure 29.3.5: Expansionary Fiscal Policy. Expansionary fiscal policy causes the deficit to increase at all levels of income, so the deficit line shifts upward. This picture illustrates the case of an increase in government purchases.

Cyclically Adjusted Budget Deficit

Given that the deficit depends on both the level of real GDP and the stance of fiscal policy, it is useful to have a way to distinguish these two influences. Put differently, it is helpful to know if the deficit is large because of the level of economic activity or because of the choices of government spending and taxes. This distinction came to the forefront in the 2004 presidential election in the United States. One of the issues raised in the debates between President George W. Bush and Senator Kerry was how the forecasted surplus from 2000 turned into the massive deficits of 2004. Were the deficits caused by the state of the economy or the policy decisions undertaken by President George W. Bush? To answer such questions, we need to decompose changes in the deficit into changes due to fiscal policy and changes due to the level of economic activity.

The Congressional Budget Office (CBO; http://www.cbo.gov) produces a measure of the budget deficit, called the **cyclically adjusted budget deficit**, for this purpose. The CBO first calculates a measure of **potential output**—the level of GDP when the economy is at full employment. Then it calculates the outlays and revenues of the federal government *under the assumption that the economy is operating at potential GDP*. The deficit is calculated by subtracting revenues from outlays. For obvious reasons, the cyclically adjusted budget deficit is also sometimes called the full-employment deficit."The Cyclically Adjusted and Standardized Budget Measures," Congressional Budget Office, April 2008, accessed July 20, 2011, cbo.gov/ftpdocs/90xx/doc9074/StandBudgetTOC.2.1.htm.

Figure 29.3.6 illustrates this idea. We first calculate the level of potential output and then use the deficit line to tell us the cyclically adjusted budget deficit or surplus for the economy. The figure shows two possibilities. In the first case, there is a government deficit when actual output is equal to potential output. In the second case, there is a government surplus when output is equal to potential output. Of course, the practical calculations are somewhat trickier than this picture suggests, but the idea is straightforward.

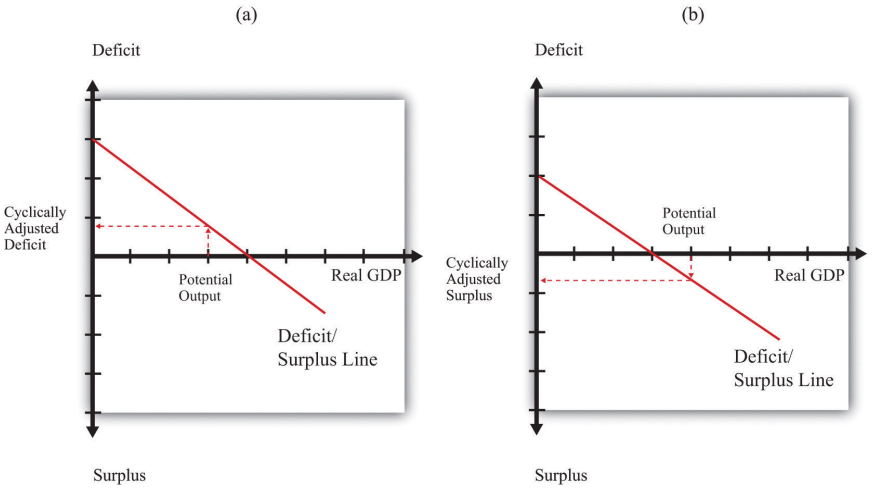


Figure 29.3.6: The Cyclically Adjusted Budget Deficit. To determine the cyclically adjusted deficit or surplus in an economy, calculate the level of potential output and then use the deficit/surplus line to determine what the deficit or surplus would be at that level of output. In panel (a), the economy has a cyclically adjusted deficit, whereas in panel (b), it has a cyclically adjusted surplus.

Figure 29.3.7 and Figure 29.3.8 show that there are two distinct reasons why a government might go from surplus into deficit—as happened in 2002, for example. Suppose that, last year, the economy was at potential output and there was a cyclically adjusted *surplus* (point A). Now imagine that this year there is a government deficit. One possibility is that the economy went into recession, as in Figure 29.3.7, point B. This is called a **cyclical deficit** because it is due to the state of the business cycle. Another is that the stance of fiscal policy has changed—for example, because of an increase in government spending, as in Figure 29.3.8, point C. The CBO calls this a **standardized deficit (or structural deficit)**.A key simplification in these pictures is that the level of potential GDP is independent of taxes and government spending. Chapter 27 explains why potential output itself might be affected by the tax code.

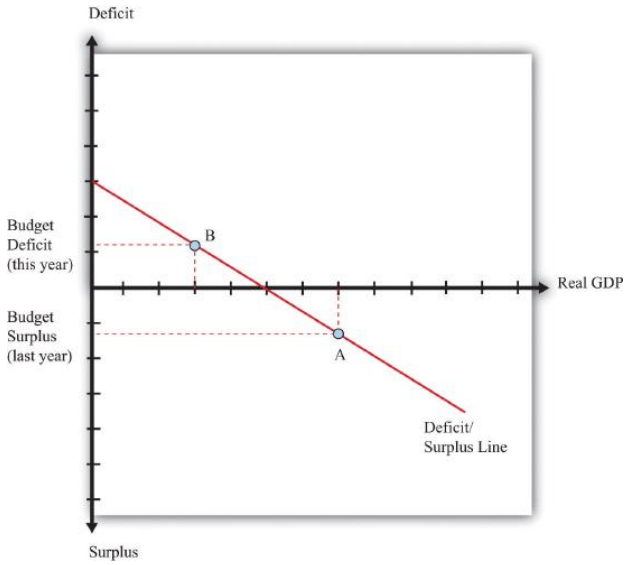


Figure 29.3.7: Cyclical Deficit. The economy went from surplus (A) to deficit (B) because of recession. Real GDP declines, tax receipts decrease, and the budget goes into deficit. The economy moves along the deficit/surplus line.

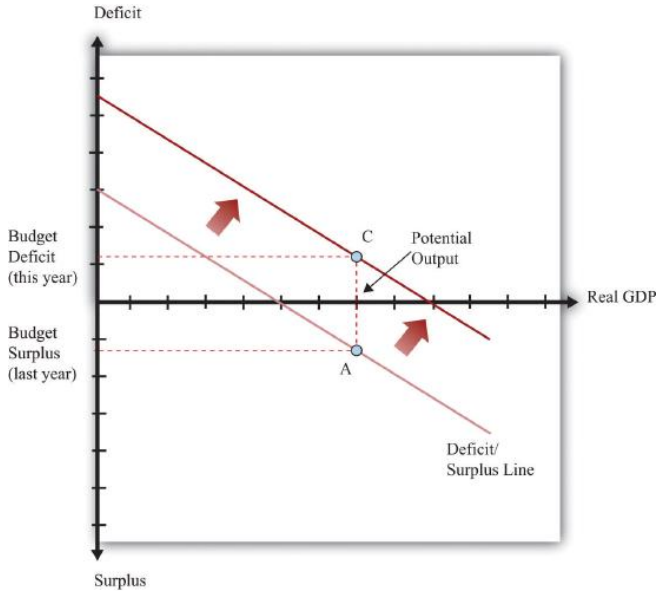


Figure 29.3.8: Structural Deficit. The economy went from surplus (A) to deficit (C) because of changes in fiscal policy. Real GDP does not change: it is at potential output in both cases. The deficit/surplus line shifts upward.

Cyclical Deficits and a Balanced-Budget Requirement

We have identified two factors that determine the size of the deficit: the stance of fiscal policy and the state of the economy. We can use this information to learn more about the effects of a balanced-budget amendment on the economy.

Suppose that the economy is at potential output. A balanced-budget requirement would say that the economy must be neither in surplus nor in deficit at this point. In other words, a balanced-budget requirement describes the overall stance of fiscal policy. The deficit/surplus line must be shifted to ensure that it passes through the horizontal axis at potential output, as shown in Figure 29.3.9.

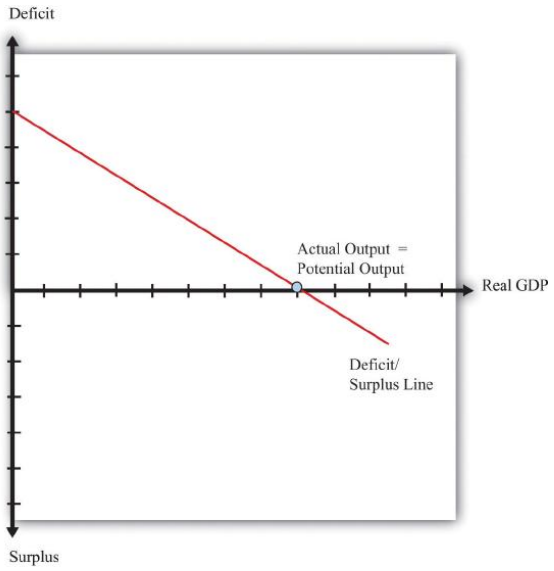


Figure 29.3.9: Balanced-Budget Requirement. A balanced-budget requirement implies that the full-employment deficit/surplus must be zero. The deficit/surplus line must pass through zero when real GDP equals potential output.

Now suppose that, for some reason, the economy goes into recession. In Figure 29.3.10, this means that output goes from potential output to some lower level. We know that this leads to a deficit, which is shown as a shift from point A to point B. Under a balanced-budget rule, the government is not allowed to let this situation persist. Instead the government must respond by increasing taxes or cutting spending, moving the economy from point B to point C. Similarly, if the economy went into a boom, this would tend to lead to a surplus. The government would be forced to cut taxes or increase spending to bring the budget back into balance. A balanced-budget amendment would force the government to conduct *procyclical* fiscal policy.In fact, the effects of a balanced-budget amendment would be even worse. The countercyclical fiscal policy would cause GDP to decrease even further, thus requiring even bigger cuts in spending or increases in taxes.

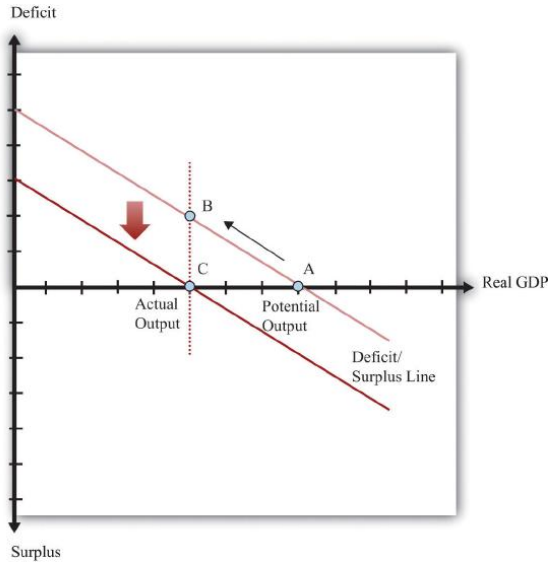


Figure 29.3.10: Recession with a Balanced-Budget Amendment. If the economy were to go into recession, a balanced-budget requirement would force the government to increase taxes or cut spending to bring the budget back into balance.

Key Takeaways

1. At a given level of GDP, an expansionary fiscal policy increases the budget deficit, and a contractionary fiscal policy decreases the budget deficit.
2. As the level of economic activity increases, tax revenues increase, transfers decrease, and the budget deficit decreases.
3. By examining the cyclically adjusted budget deficit, it is possible to evaluate how much of the budget deficit is due to the state of the economy and how much is due to the stance of fiscal policy.

Exercises

1. In Table 29.3.2 "Deficit and Income", why do tax receipts increase with real GDP?
2. What do we know about fiscal policy if the cyclically adjusted budget deficit is negative?
3. If the budget is in deficit, what do we know about the level of real GDP compared to potential GDP?

This page titled 29.3: The Causes of Budget Deficits is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

29.4: The Benefits of Deficits 29.4: The Benefits of DeficitsEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. When do countries run government budget deficits?
2. Why might a country incur a government budget deficit?

To evaluate the merits of a balanced-budget amendment, we need to know why governments run deficits in the first place. After all, governments may have good reasons for these policies. We have seen one explanation for deficits: governments run deficits because of economic downturns. Reductions in gross domestic product (GDP), other things being equal, lead to increases in the budget deficit. We are more concerned with why governments *choose* to run persistent structural deficits, though. We first look to history for clues.

Government Debt: A Historical Perspective

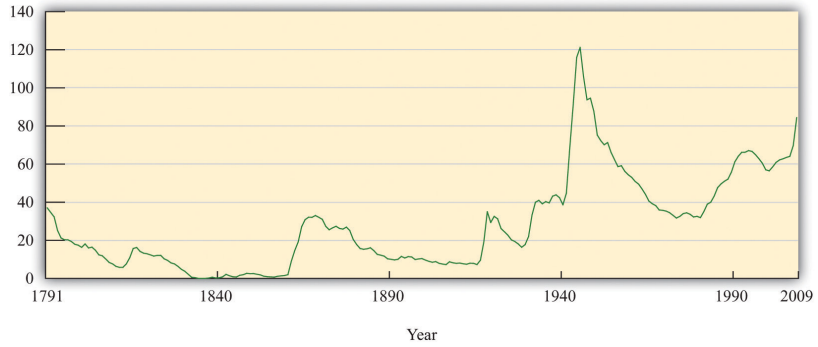


Figure 29.4.1: Ratio of US Debt to GDP, 1791–2009. Source: Debt data from *http://www.treasurydirect.gov/govt/reports/pd/histdebt/histdebt.htm; GDP data from https://eh.net/.*

Figure 29.4.1 shows the ratio of US federal government debt to GDP from 1791 to 2009. The US Civil War in the 1860s, World War I in 1917, and World War II in the early 1940s all jump out from this figure. These are periods in which the stock of US federal debt soared. During the Civil War, the stock of debt was $64,842,287 in 1860 and peaked at $2,773,236,174 in 1866. The debt level was more than 40 times higher in 1866 than in 1860.

In 1915 (after World War I had started but before the United States had entered the war), the stock of debt was $3,058,136,873.16, not much more than the level in 1866. By 1919, the level of the debt was $27,390,970,113.12, an increase of almost 800 percent. During World War II, there was again a significant buildup of the debt. In 1940, the level of debt outstanding was $42,967,531,037.68, or about 42 percent of GDP. By 1946, this had increased by about 527 percent to $269,422,099,173.26. In 1946, the outstanding debt was 121 percent of GDP.

There are two other periods that show a significant buildup of the debt relative to GDP. The first is the Great Depression. This buildup was *not* due to a big increase in borrowing by the government. Rather, it was largely driven by the decline in the level of GDP (the denominator in the ratio). The second is the period from the 1980s to the present. The buildup of the debt in the 1980s was unprecedented in peacetime history.

Figure 29.4.1 also shows a dramatic asymmetry in the behavior of the debt-to-GDP ratio. Although the increases in this ratio are typically rather sudden, the decreases are much more gradual. Look again at the rapid increase in the debt-to-GDP ratio around the Civil War. After the Civil War ended, the debt-to-GDP ratio decreased but only slowly. As seen in the figure, the debt-to-GDP ratio decreased for about 45 years, from 1870 to 1916. Part of this decrease was due to the growth in GDP over the 45 years, and part was due to a decrease in nominal debt outstanding until around 1900.

Why Do Governments Run Deficits?

It is evident that during periods of war the debt is higher. What underlies this relationship between wars and deficits? War is certainly expensive. Take, for example, the conflicts in Iraq and Afghanistan. Congress has already appropriated about $1 trillion for these wars, and a Congressional Budget Office study projected the conflicts would eventually cost the United States about $2.4

trillion. When government purchases increase due to a war, a government can either increase taxes to pay for the war or issue government debt. Remember that when the government runs a deficit to pay for a war, it is borrowing from the general public. The government's intertemporal budget constraint reminds us that—since government debt is ultimately paid for by taxes in the future—the choice is really between taxing households now or taxing them later. History tells us that deficits have been the method of choice: governments have chosen to tax future generations to pay for wars.

There are two arguments in favor of this policy:

1. **Fairness.** Any gains from winning a war will be shared by future generations. Hence the costs should be shared as well: the government should finance the war with debt so that future generations will repay some of the obligations. To take an extreme case, suppose a country is fighting for its right to exist. If it wins the war, future generations will also benefit.
2. **Tax smoothing.** A good fiscal policy is one where tax rates are relatively constant. In the face of a rapid increase in spending, such as a war, the best policy is one that pays for the spending increase over many periods of time, not in one year.

Taxation is expensive to the economy because it distorts economic decisions, such as saving and labor supply. The amount people want to work depends on their real wage, after taxes. So if tax rates are increased to finance government spending, this reduces the benefit from working. Put differently, increased income taxes increase the price of consumption relative to leisure. The fact that people work less when taxes increase is a distortionary effect of taxation. Instead of bunching all this distortionary taxation into a short amount of time, such as a year, it is more efficient for the government to spread the taxation over many years. This is called **tax smoothing**. So by running a budget deficit, the government imposes relatively small distortions over many years rather than imposing large distortions within a single year.

> **Toolkit: Section 31.3**
>
> For more analysis of the choice underlying labor supply, you can review the labor market in the toolkit.

Similar arguments apply to other cases in which governments engage in substantial spending. Imagine that the government is considering putting a large amount of resources into cancer research. The discovery of a successful cancer treatment would, of course, benefit many generations of citizens. Because households would share the gains in the future, the costs should be shared as well. By running a budget deficit, the government is able to distribute the costs across generations of citizens in parallel with the benefits. From the perspective of both fairness and efficiency, there are some gains to deficit spending.

More generally, we might want to make a distinction among different types of government purchases, just as we do among private purchases. We know that the national accounts distinguish consumption purchases (broadly speaking, things from which we get short-run benefit, such as food and movies) from investment purchases (things that bring long-term benefit, such as factories and machinery). Likewise, we might want to distinguish between government consumption, such as wages of employees at the Department of Motor Vehicles, from government investment, such as spending on cancer research. We could then argue that it makes more sense to borrow to finance government investment rather than government consumption.

Although a very nice idea in principle, this approach to the government accounts often founders on the practicalities and the politics of implementation. First, it is not at all clear how to classify many government expenditures. Was a launch of the space shuttle consumption or investment? What about the wages of teachers in the public schools? What about the money spent on national parks? Second, politicians would have a strong incentive to classify expenditures as investment rather than consumption, to justify deferring payment.

Another benefit of deficits is that they can play a role in economic stabilization. Chapter 22 spelled out in detail the role for fiscal policy in stabilizing output. In the short run, the level of economic activity can deviate from potential GDP. As a consequence, aggregate expenditures play a role in determining the level of output. Fiscal policy influences the level of aggregate expenditures. Changes in government purchases directly affect aggregate expenditures because they are a component of spending, and changes in taxes indirectly affect aggregate demand through their effect on consumption. Hence deficit spending can help to stabilize the economy.

In summary, there are several arguments for allowing governments to run deficits. We would forswear these benefits if we were to adopt a balanced-budget amendment.One of the arguments for deficits—funding wars—is an explicit exception (and the only such exception) written into the bill from that we quoted earlier. But we conclude by noting that there is a further, much less benign, reason for government deficits: they may benefit politicians even if they do not benefit the country as a whole. Deficits allow politicians to provide benefits to constituents today and leave the bill to future generations. If politicians and voters care more about

current benefits than future costs, then they have a strong incentive to incur large deficits and let future generations worry about the consequences.

Deficits around the World

Do other countries also run deficits in the way that the United States does? Table 29.4.1 "Budget Deficits around the World, 2005*" summarizes the recent budgetary situation for several countries around the world. With the exception of Argentina, all the countries were running deficits in 2005.The table deliberately does not express the deficits relative to any measure of economic activity in the country. Thus it is hard to say whether these deficits are large or small. An Economics Detective exercise at the end of the chapter encourages you to look at this question.

Table 29.4.1: Budget Deficits around the World, 2005*

Country	Revenues	Expenditures	Deficit
Argentina	42.6	39.98	−2.62
China	392.1	424.3	32.2
France	1,006	1,114	108
Germany	1,249	1,362	113
Italy	785.7	861.5	75.8

* Data are in millions of US dollars.

Source: CIA Fact Book, http://www.cia.gov/cia/publications/factbook/fields/2056.html.

France, Germany, and Italy are of particular interest. These three countries are part of the European Union (EU). In January 1999, when the Economic and Monetary Union was formed, a restriction on the budget deficits of EU countries went into effect. This measure was contained in legislation called the *Stability and Growth Pact*. This pact is discussed in detail in “Stability and Growth Pact,” European Commission Economic and Financial Affairs, accessed September 20, 2011, http://ec.europa.eu/economy_finance/sgp/index_en.htm. Its main component is a requirement that member countries keep deficits below a threshold of 3 percent of GDP. This threshold is not set to zero to allow countries the ability to deal with fluctuations in real GDP. In other words, although the EU does not impose a strict balanced-budget requirement, it does impose limits on member countries. In recent years, however, these limits have been exceeded. For example, in 2005, Germany’s deficit was more than 4.5 percent of its GDP. During the past few years, Germany has been in a recession and, as highlighted by Figure 29.4.4, its deficit grew considerably. Instead of imposing contractionary fiscal policies to reduce its deficit, Germany allowed its deficit to grow outside the bounds set by the Stability and Growth Pact. The economic crisis of 2008 and subsequent recession that impacted many of the world economies had a further effect on the budget deficits of countries in Europe, contributing to severe debt crises and bailouts in Greece, Ireland, and Portugal. We examine what happened in these countries in Chapter 30.

Key Takeaways

1. Countries run government budget deficits when faced with large expenditures, such as a war.
2. By running a deficit, a government is able to spread distortionary taxes over time. Also, a deficit allows a government to allocate tax obligations across generations of citizens who all benefit from some form of government spending. Finally, stabilization policy often requires the government to run a deficit.

Exercises

1. What does it mean to say that a tax is “distortionary”?
2. What is the political benefit to deficit spending?
3. When does “fairness” provide a basis for running a deficit?

This page titled 29.4: The Benefits of Deficits is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

29.5: The Costs of Deficits 29.5: The Costs of DeficitsEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What is the crowding-out effect?
2. When is the crowding-out effect of government deficits large?

We now turn to the costs of deficit spending. (Although we refer to this as "deficit spending," the same arguments apply if we analyze the effects of a reduction in the government surplus.) First, we need to understand what happens in the financial sector of the economy if the government runs a deficit.

Savings and Investment

Earlier, we examined the circular flow of income in the government sector. Now we turn our attention to the circular flow in the financial sector, which is shown in Figure 29.5.1. We also examined this sector in Chapter 20. As with all sectors in the circular flow, the flows into and from the sector must match. In the case of the government sector earlier in the chapter, the balance of these flows is another way of saying that the government must satisfy its budget constraint. The rules of accounting tell us that, in the financial sector, the flows in must likewise match the flows out, but what is the underlying economic reason for this? The answer is that the flows are brought into balance by adjusting interest rates in the economy. We think of the financial sector of the economy as a large **credit market** in which the price is the **real interest rate**.

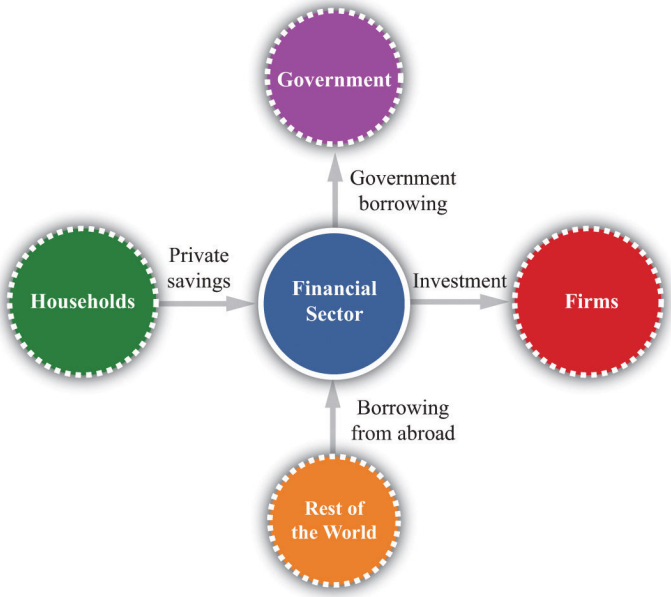


Figure 29.5.1: The Financial Sector in the Circular Flow. Funds flow into the financial sector as a result of household savings and borrowing from the rest of the world. Funds flow from the government sector (to finance the government deficit) and to the firm sector to finance investment.

Toolkit: Section 31.24

You can review the credit market in the toolkit.

The Credit Market

The supply of loans in the credit market comes from (1) private savings of households and firms, (2) savings or borrowing of governments, and (3) savings or borrowing of foreigners. Households generally respond to an increase in the real interest rate by saving more. Higher real interest rates also encourage foreigners to send funds to the domestic economy. **National savings** are defined as private savings plus government savings (or, equivalently, private savings minus the government deficit). The total

supply of savings is therefore equal to national savings plus the savings of foreigners (that is, borrowing from other countries). The demand for credit comes from firms who borrow to finance investment. As the real interest rate increases, investment spending decreases. For firms, a high interest rate represents a high cost of funding investment expenditures.

The matching of savings and investment in the aggregate economy is described by the following equations:

$$investment = national\ savings + borrowing\ from\ other\ countries$$

or

$$investment = national\ savings - lending\ to\ other\ countries.$$

The response of savings and investment to the real interest rate is shown in Figure 29.5.2. In equilibrium, the quantity of credit supplied equals the quantity of credit demanded. We have assumed that the country is borrowing from abroad, but nothing at all would change—other than the way we describe the supply curve—if the domestic economy were instead lending to other countries.

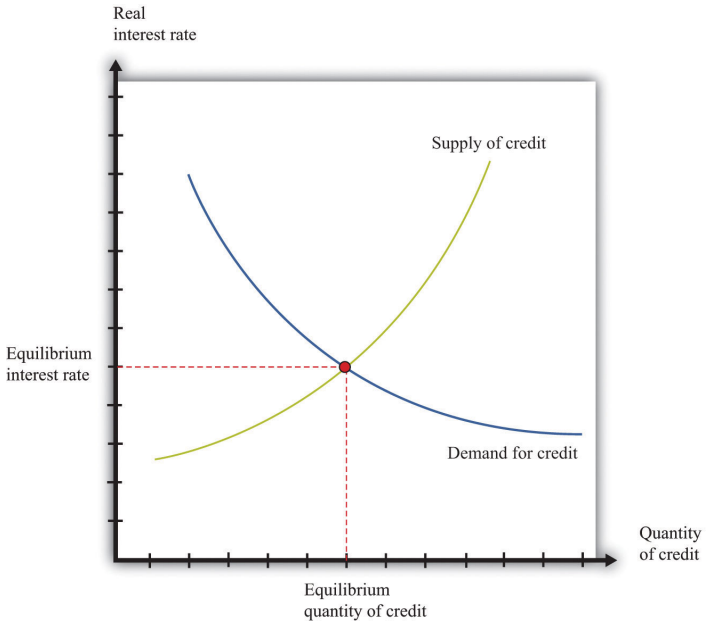


Figure 29.5.2: The Credit Market. Adjustment of the real interest rate ensures that the flows into and from the financial sector balance. The supply of loans comes from national savings plus borrowing from abroad. The demand for loans comes from firms seeking funds for investment.

Crowding Out

Armed with this framework, we can determine what happens to saving, investment, and interest rates when the deficit increases. Figure 29.5.3 begins with the credit market in equilibrium at point *A*. The increased government deficit is shown as a leftward shift of the national savings line. At each level of the real interest rate, the increased government deficit means that national savings is lower.

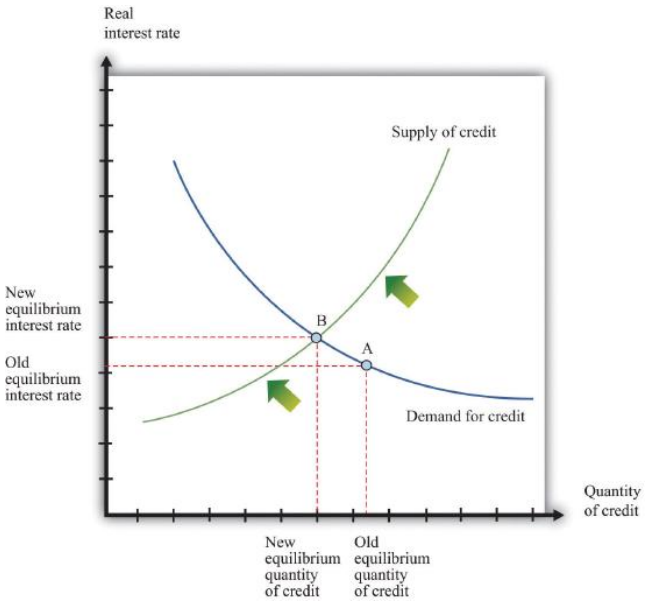


Figure 29.5.3: Crowding Out. An increase in the deficit means a reduction in saving, so the saving line shifts leftward and the new equilibrium entails a higher real interest rate and a lower level of investment. The equilibrium decrease in saving and investment is less than the initial decrease in government saving.

This shift in the savings line implies that the market for loans is no longer in equilibrium at the original interest rate. Real interest rates increase in response to the excess of investment over savings until the market is once again in equilibrium, at point *B* in Figure 29.5.3. Comparing *A* to *B*, we can see there are two consequences of the government deficit: (1) real interest rates increase, and (2) the amount of credit, and hence the level of investment, is lower. The reduction in investment spending caused by an increase in government spending is called **crowding out**. In addition, household spending on durable goods also decreases when interest rates increase: this is also an example of crowding out. To the extent that household spending on durables and investment are sensitive to changes in real interest rates, the crowding-out effect can be substantial.

Crowding out also operates through net exports. From Figure 29.5.3, we know that an increase in the deficit leads to an increase in interest rates. Increased interest rates have three effects:

1. They cause investment to decrease. This is the crowding-out effect.
2. They cause private saving to increase. Higher interest rates encourage people to save rather than consume.
3. They attract funds from other countries. Investors in other countries see the higher interest rates and decide to invest in the domestic economy.

The second and third effects explain why the supply of credit slopes upward in Figure 29.5.5. As a result, the decrease in investment is not as large as the increase in the deficit. The decrease in government saving is partly offset by an increase in private saving and an increase in borrowing from abroad. Increased borrowing from abroad must result in a decrease in net exports to keep the flows into and from the foreign sector in balance.

To understand these linkages, imagine that the United States sells additional government debt, some of which is purchased by banks in Europe, Canada, Japan, and other countries. These purchases of government debt require transactions in the foreign exchange market. If a bank in Europe purchases US government debt, there is an increased demand for dollars in the euro market for dollars, which leads to an **appreciation** in the price of the dollar. When the dollar appreciates, US citizens find that European goods and services are cheaper, whereas Europeans find that US goods and services are more expensive. US imports increase and exports decrease, so net exports decrease.

To summarize, an increased government deficit leads to the following:

- An increase in the real interest rate
- An appreciation of the exchange rate
- A reduction in investment and in purchases of consumer durables
- An increase in the trade deficit

Table 29.5.1 "Investment, Savings, and Net Exports (Billions of Dollars)" shows the US experience during the 1980s, when the US federal government ran a large budget deficit (the negative entries in the federal budget surplus column). The table also reveals that the United States ran a sizable *trade* deficit starting in 1983. This phenomenon became known as the twin deficits.

Table 29.5.1: Investment, Savings, and Net Exports (Billions of Dollars)

Year	Investment	Trade Surplus	National Saving	Budget Surplus	Error
1980	579.5	11.4	549.4	−23.6	41.5
1981	679.3	6.3	654.7	−19.4	30.9
1982	629.5	0.0	629.1	−94.2	0.4
1983	687.2	−31.8	609.4	−132.3	46.0
1984	875	−86.7	773.4	−123.5	14.9
1985	895	−110.5	767.5	−126.9	17.0
1986	919.7	−138.9	733.5	−139.2	47.3
1987	969.2	−150.4	796.8	−89.8	22.0
1988	1,007.7	−111.7	915.0	−75.2	−19
1989	1,072.6	−88.0	944.7	−66.7	39.9

Source: *Economic Report of the President* (Washington, DC: GPO, 2004), table B-32.

Even though recent years have also seen high deficits in the United States, interest rates have not increased, so we have not seen crowding out. This is because the Federal Reserve has also been operating in credit markets to keep interest rates low. Although crowding out is associated with fiscal policy, it also depends on what policies the monetary authority chooses to pursue.

When crowding out does occur, its long-term consequences may be significant. Lower investment translates, in the long run, into a lower standard of living. Chapter 21 explained how investment feeds into long-run economic growth. An increase in government spending means that the country has chosen to consume more now and less in the future. Similarly, crowding out of net exports means that the economy is borrowing more from other countries. This again means that the country has chosen to consume more now in exchange for debt that must be paid back later. The crowding-out effect is perhaps the most powerful argument in favor of a balanced-budget requirement.

Key Takeaways

1. Crowding out occurs when government deficits lead to higher real interest rates and lower investment. The high interest rates can also cause the domestic currency to appreciate, leading to a decrease in net exports.
2. The crowding-out effect is large when spending by households on durables and investment spending are sensitive to variations in the real interest rate and when exports are sensitive to changes in the exchange rate.

Exercises

1. Why do higher interest rates cause the currency to appreciate?
2. In using the credit market to study the effects of government deficits on real interest rates, what did we assume about household saving?

This page titled 29.5: The Costs of Deficits is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

29.6: The Ricardian Perspective 29.6: The Ricardian PerspectiveEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What is the Ricardian theory about the effects of deficits on interest rates and real gross domestic product (GDP)?
2. What is the evidence on the Ricardian theory?

Buried in our analysis of the crowding-out effect is a critical assumption. We argued that an increase in the government deficit would reduce national savings at every level of the interest rate. Implicitly, we assumed that the change in government behavior had no *direct* effect on private savings. Instead, there was an indirect effect: savings increased when the interest rate increased. But at any given level of interest rates, we assumed that private saving was unchanged.

Perhaps that is not the most reasonable assumption. Consider the following thought experiment:

- The government sends you and everyone else a check for $1,000, representing a tax cut.
- The government finances this increase in the deficit by selling government bonds.
- The government announces that it will increase taxes next year by the amount of the tax cut plus the interest it owes on the bonds that it issued.

What will be your response to this policy? A natural reaction is just to save the entire tax cut. After all, if the government cuts taxes in this fashion, then all it is doing is postponing your tax bill by one year. Your lifetime resources have not increased at all. Hence you can save the entire tax cut, accumulate the interest income, and use this income to pay off your increased tax liability next year.

The Household's Lifetime Budget Constraint

The household's **lifetime budget constraint** tells us that households must equate the discounted present values of income and expenditures over their lifetimes. We use it here to help us understand how households behave when there are changes in the *timing* of their income. In general, the budget constraint must be expressed in terms of discounted present values:

discounted present value of lifetime consumption = discounted present value of lifetime disposable income.

When the real interest rate is zero, life is simple. It is legitimate simply to add together income and consumption in different years. In this case, the lifetime budget constraint says that

$$total\ lifetime\ disposable\ income = total\ lifetime\ consumption.$$

The measure of income used in the household's budget constraint is lifetime disposable income. You can think of discounted lifetime disposable income as the difference between the discounted present value of income (before taxes) and the discounted present value of taxes. The effect of a government's tax policy is through the discounted present value of household taxes.

Toolkit: Section 31.34

You can review the life-cycle model of consumption in the toolkit.

Private Savings and Government Savings

In our earlier thought experiment, the increase in the government deficit was exactly offset by an increase in private savings. This implication is shown in Figure 29.6.1: nothing happens. The *composition* of national savings changes, so public savings decrease, and private savings increase. But these two changes exactly offset each other since the private sector saves the entire amount of the tax cut. As a result, the supply curve does not shift. Since national savings do not change, the equilibrium remains at point *A*, and there is no crowding-out effect. Economists call this idea **Ricardian equivalence**, after David Ricardo, the 19th century economist who first suggested such a link between public and private saving. Ricardian equivalence occurs when an increase in the government deficit leads to an equal increase in private saving and no change in either the real interest rate or investment.

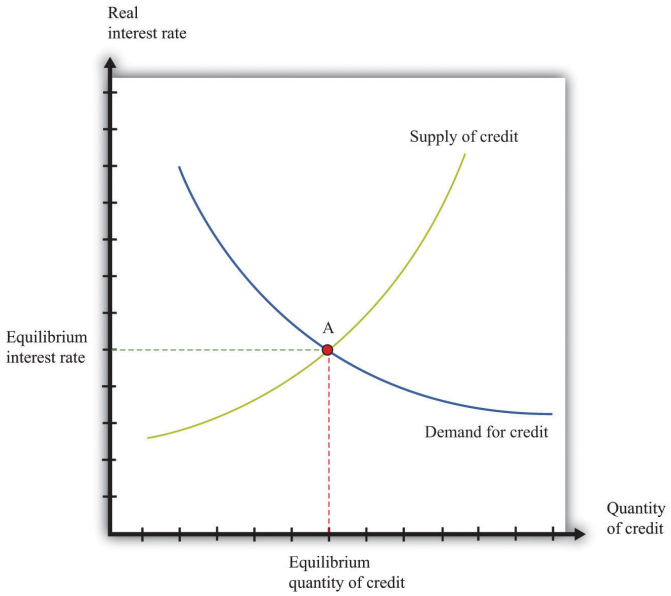


Figure 29.6.1: Ricardian Equivalence. An increase in the government deficit is equivalent to a decrease in government savings, which shifts national savings leftward. In a Ricardian world, private savings increases by an offsetting amount, so the final result is no change in national savings.

The Ricardian perspective can be summarized by two related claims:

1. The timing of taxes is irrelevant.
2. If government purchases are unchanged, tax cuts or increases should have no effect on the economy.

These claims follow from the government's intertemporal budget constraint and the household's lifetime budget constraint, taken together. The government's constraint tells us that a given amount (that is, a given discounted present value) of government spending implies a need for a given (discounted present value) amount of taxes. These taxes could come at all sorts of different times, with different implications for the deficit, but the total amount of taxes must be enough to pay for the total amount of spending. The household's lifetime budget constraint tells us that the timing of taxes may be irrelevant to households as well: they should care about the total lifetime (after-tax) resources that they have available to them.

The implications of the Ricardian perspective are not quite as stark if the increased deficit is due to increased government spending. Households should still realize that they have to pay for this spending with higher taxes at some future date. Lifetime household income will decrease, so consumption will decrease. However, consumption smoothing suggests that the decrease in consumption will be spread between the present and the future. The decrease in current consumption will be less than the increase in government spending, so national savings will decrease, as in the analysis in 29.5 Section "The Costs of Deficits". Since the Ricardian perspective says that the timing of taxes is irrelevant, the effect is the same as it would be if the taxes were also imposed today. So one way of thinking about this is to suppose that the government increases spending and finances that increase with current taxes.

If the Ricardian perspective is an accurate description of how people behave, then much of our analysis in this chapter becomes irrelevant. Deficits are not needed to spread out the costs of major government expenditures because households can do this smoothing for themselves. Changes in taxes have no effect on aggregate spending, so there is no crowding-out effect.

As for a balanced-budget amendment, it too would be much less significant in such a world. Ricardian households effectively "undo" government taxation decisions. However, the exact effect of an amendment would depend on how the government chose to ensure budget balance. Suppose the economy went into recession, so tax revenues decreased. There are two ways to restore budget balance. One is to increase taxes. According to the Ricardian perspective, this would have no effect on the economy at all. The other is to cut government purchases. As we have seen, this would have some effects.

Evidence

The Ricardian perspective seems very plausible when we consider a thought experiment such as a tax cut this year matched by a corresponding tax increase next year. At the same time, a typical tax cut is not matched by an explicit future tax increase at a specified date. Instead, a tax cut today means that *at some unspecified future date* taxes will have to be increased. Furthermore, the

Ricardian perspective requires that households have a sophisticated economic understanding of the intertemporal budget constraint of the government.

It is therefore unclear whether this Ricardian view is relevant when we evaluate government deficits. Do households understand the government budget constraint and adjust their behavior accordingly, or is this just an academic idea—theoretically interesting, perhaps, but of limited relevance to the real world? This is an empirical question, so we turn to the data. There are two natural ways to examine this question. The first is to determine the relationship between government deficits and real interest rates in the data. The second approach is to examine the relationship between government deficits and private saving.

Deficits and Interest Rates

We want to answer the following question: do increases in government deficits **cause** real interest rates to increase?

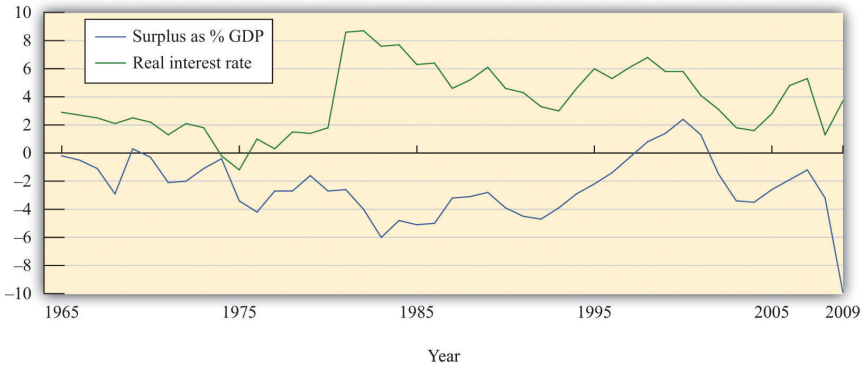


Figure 29.6.2: US Surplus/GDP Ratio and Real Interest Rate, 1965–2009. There is some evidence that declines in the government surplus are associated with higher real interest rates, contrary to the Ricardian view.

Source: Economic Report of the President, 2010, Tables B-63 and B-72.

Figure 29.6.2 shows two series. The first is the ratio of the US budget surplus to GDP, measured on the left axis. (Be careful—this is the surplus, not the deficit. The economy is in deficit when this series is negative.) The second is a measure of the real interest rate, measured on the right axis. The figure shows that interest rates do seem to increase when the surplus decreases and *vice versa*. We can compute the **correlation** between the surplus-to-GDP ratio and the real interest rate. For this data the correlation is −0.16. The minus sign means that when the surplus is above average, the real interest rate tends to be below its average value, consistent with the impression we get from the graph. However, the correlation is not very large.

The 1980s stand out in the figure. During this period, the budget deficit grew substantially, reflecting low economic activity as well as tax cuts that were enacted during the early years of the Reagan administration. Starting in 1982, real interest rates increased substantially, just as the budget deficit was widening. This is consistent with crowding out and contrary to the Ricardian perspective. We must be cautious about inferring causality, however. It is false to conclude from this evidence that an increase in the deficit *caused* interest rates to increase. It might be that some other force caused high interest rates and low economic activity. For example, as explained in Chapter 25, tight monetary policy (such as that enacted in the 1980s) leads to high interest rates and can push the economy into recession, leading to a deficit.

Toolkit: Section 31.23

You can review the definition of a correlation in the toolkit.

Government Deficits and Private Saving

According to the Ricardian perspective, increases in the government deficit should be matched by increases in private saving and *vice versa*. Private and government savings rates for the United States are shown in Figure 29.6.4. These calculations rely on data from the *Economic Report of the President* (Washington, DC: GPO, 2011), table B-32, accessed September 20, 2011, www.gpoaccess.gov/eop. The private saving rate equals private saving as a percentage of real GDP. The government saving rate essentially equals the government surplus as a percentage of GDP (there are some minor accounting differences that we do not need to worry about).

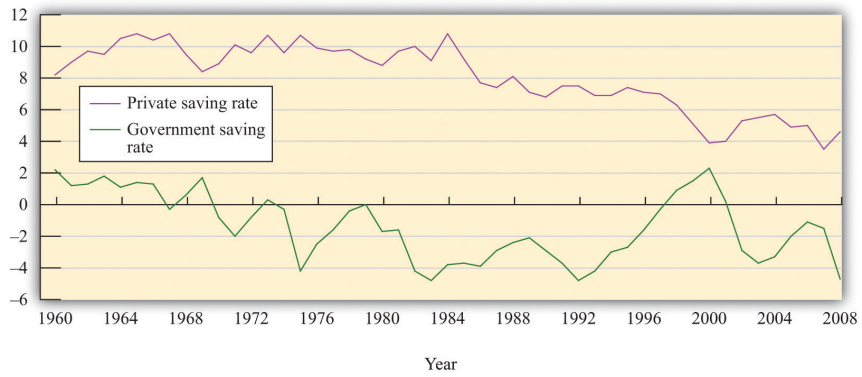


Figure 29.6.3: US Government and Private Savings Rates. There is some evidence that private and government saving move in opposite directions, as suggested by the Ricardian view.

Source: Calculations based on Economic Report of the President, Table B-32.

Private savings increased from the 1980 to 1985 period and decreased thereafter. Large deficits emerged during the early 1980s (negative government savings). At this time, there was an increase in the private savings rate. The government savings rate increased steadily during the 1990s, and, during this period, the private savings rate decreased. These data are therefore more supportive of the Ricardian view: private and government savings were moving in opposite directions.

Turning to international evidence, an Organisation for Economic Co-operation and Development study that examined 21 countries between 1970 and 2002 found that changes in government deficits were associated with partially offsetting movements in private saving. On average, the study found that changes in private savings offset about one-third to one-half of changes in the government deficit. See Luiz de Mello, Per Mathis Kongsrud, and Robert Price, "Saving Behaviour and the Effectiveness of Fiscal Policy," *Economics Department Working Papers* No. 397, Organisation for Economic Co-operation and Development, July 2004, accessed September 20, 2011, www.oecd.org/officialdocuments/displaydocumentpdf/?cote=eco/wkp(2004)20&doclanguage=en. Figure 29.6.4 and Figure 29.6.5 reproduce some figures from this study. In Spain and Greece, for example, we see patterns of savings that are consistent with the Ricardian perspective: private savings and government savings move in opposite directions. By contrast, the pictures for Ireland and France show little evidence of such an effect.

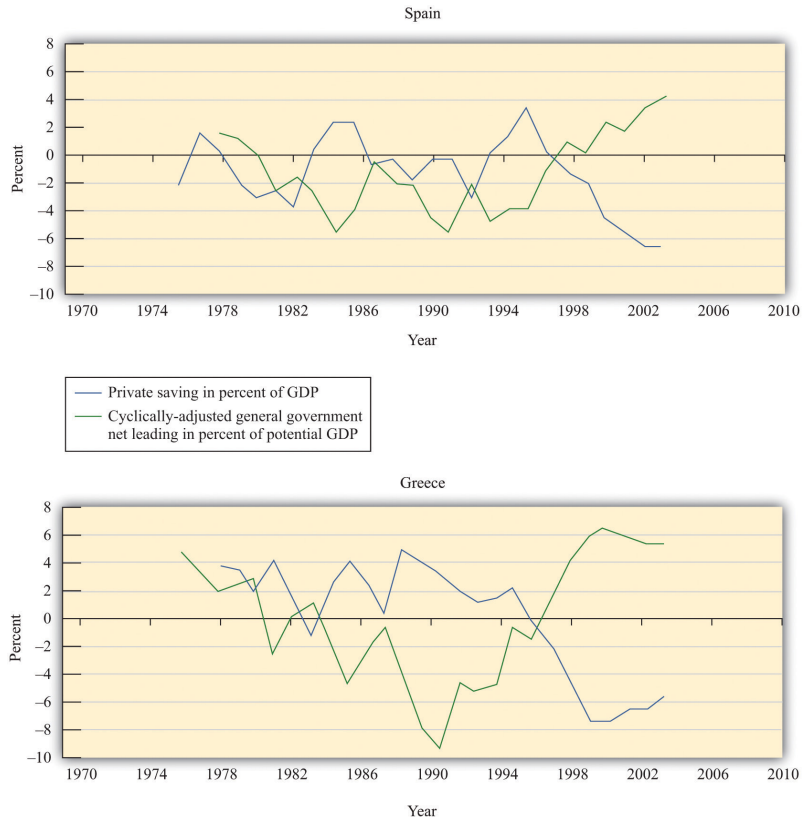


Figure 29.6.4: Government and Private Savings Rates in Spain and Greece. Source: Economic Report of the President, 2010, Tables B-63 and B-72.

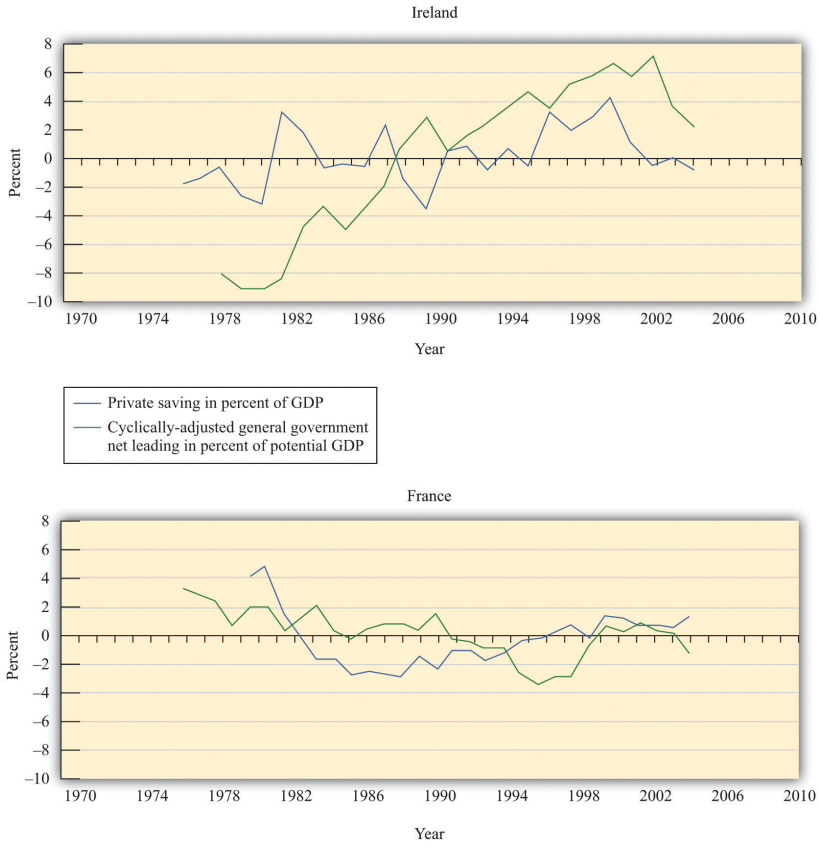


Figure 29.6.5: Government and Private Savings Rates in France and Ireland. Source: Calculations based on Economic Report of the President, Table B-32.

The data from the United States and other countries indicate that this is almost certainly one of those questions where the truth is in the middle. We do not observe households behaving completely in accordance with the Ricardian perspective. As a result, we conclude that deficits do have the real effects on the economy that we discussed at length in this chapter. At the same time, there is evidence suggesting that households pay attention to the government budget constraint. The Ricardian perspective is more than just an academic curiosity: some households, some of the time, adjust their behavior to some extent.

Key Takeaways

1. According to Ricardian theory, a government deficit will be offset by an increase in household saving, leaving real interest rates and the level of economic activity unchanged. The key to the theory is the anticipation of households of future taxes when the government runs a deficit.
2. There is some evidence that interest rates are high when deficits are high, contrary to the prediction of the Ricardian view. But during some periods of large deficits, the household saving rate is high as well. The evidence on Ricardian equivalence is not conclusive.

Exercises

1. If the government cuts taxes, what happens to public saving, private saving, and national saving according to the Ricardian theory?
2. What is the difference between causation and correlation when we examine the relationship between budget deficits and real interest rates?

This page titled 29.6: The Ricardian Perspective is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

29.7: End-of-Chapter Material 29.7: End-of-Chapter MaterialEdit Update titlecancel

In Conclusion

We started this chapter by asking whether the United States should adopt a balanced-budget amendment to the constitution. This question has both political and economic ramifications.

It is not our purpose in this book to answer this question, or others like it, for you. Most interesting questions do not have easy answers. Instead, they come down to assessments of costs and benefits and judgments about which frameworks best describe the world that we live in. Our intent here was to provide *you* with the ability to assess the arguments about a balanced-budget amendment and, more generally, the effects of deficit spending on the economy.

We saw in this chapter that there are certainly both benefits and costs associated with deficit finance. Key benefits include the ability to spread out the payments for large government purchases and the opportunity to use deficits to stimulate economies in recession. The main cost of deficits is that they increase real interest rates, thus crowding out investment and slowing long-term growth.

As we also saw, these effects might be tempered by an increase in household savings in response to government deficits. The evidence suggests that the Ricardian perspective on deficits has partial validity. Changes in government savings are likely to be partially, but not completely, offset by changes in households' saving behavior.

We also noted that a balanced-budget amendment would not absolve government of the difficult choices involved in balancing the budget. It is one thing to pass a law saying that the budget must be balanced. It is quite another to come up with the spending cuts and tax increases that are needed to make it happen.

Meanwhile, time is passing. Go and look again at the size of the debt outstanding reported at the US Treasury (www.treasurydirect.gov/NP/BPDLogin?application=np). How much has it changed since you first checked it? How much has your share of the debt changed?

Key Links

- The Congressional Budget Office: http://www.cbo.gov
- US Treasury calculation of the public debt: http://www.publicdebt.treas.gov
- CIA World Factbook: https://www.cia.gov/library/publications/the-world-factbook
- US national debt clock: www.brillig.com/debt_clock
- Debate on balancing the budget:
  - Concord Coalition: http://www.concordcoalition.org/issues
  - Americans for a Balanced-Budget Amendment: http://www.balanceourbudget.com
  - Center on Budget and Policy Priorities: http://www.cbpp.org/archiveSite/bba.htm

exercises

1. The following table is a table of the same form as Table 29.2.1 "Calculating the Deficit" but with some missing entries. Complete the table. In which years was there a balanced budget?

Year	Government Purchases	Tax Revenues	Transfers	Net Taxes	Deficit
1	60	10	20	−10	
2	80	100	20		
3		120	20	100	−20
4	140	180			0
5			20	140	40

TABLE 29.7.1: CALCULATING THE DEFICIT

1. The following table lists income and the tax rate at different levels of income. In this exercise the tax rate is different at different levels of income. For income below 500, the tax rate is 20 percent. For income in excess of 500, the tax rate is 25 percent. Calculate tax receipts for this case.

Income	Marginal Tax Rate	Tax Receipts
0	0.2	
100	0.2	
500	0.2	
1000	0.25	
2000	0.25	
5000	0.25	

TABLE 29.7.2: TAX RECEIPTS AND INCOME

1. Consider the following table. Suppose that government purchases are 500, and the tax rate is 20 percent. Furthermore, suppose that real gross domestic product (GDP) takes the values indicated in the table. If the initial stock of debt is 1,000, find the level of debt for each of the 5 years in the table.

Year	GDP	Deficit	Debt (Start of Year)	Debt (End of Year)
1	3,000		1,000	
2	2,000			
3	4,000			
4	1,500			
5	2,500			

TABLE 29.7.3: EXERCISE

1. For the example in the preceding table titled “Exercise”, are the deficits and surpluses due to variations in the level of GDP or fiscal policy? Suppose you were told that potential GDP was 4,000. Is there a full employment deficit or surplus when actual GDP is 3,000? Design a fiscal policy so that the budget is in balance when real GDP is equal to potential GDP.
2. Draw a version of Figure 29.7.4 using the data for tax receipts you calculated in the table titled “Tax Receipts and Income”, and assuming government purchases equal 475. At what level of GDP is the budget in balance?
3. The text says that expansionary fiscal policy increases the deficit given the level of GDP. Would an expansionary fiscal policy necessarily increase the deficit if GDP changes as well?
4. Compare Figure 29.7.1 (from 1940 onward) with Figure 29.7.2. Why do the figures look so different from each other?
5. Suppose that investment is very sensitive to real interest rates. What does this mean for the slope of the demand curve in the credit market? Will it make the crowding-out effect large or small?

Economics Detective

1. The price of government debt during the Civil War makes for a fascinating case study. Both the Union and the Confederacy were issuing debt to finance their expenditures. Try to do some research on the value of Civil War debt to answer the following questions.
    1. How much did the Union and the Confederacy rely on deficits rather than taxes to finance the war efforts?
    2. What do you think happened to the value of the Union and Confederacy debt over the course of the war?
    3. Do you think these values were positively or negatively correlated?
    4. A starting point for your research is a website (http://www.tax.org/Museum/1861-1865.htm) that summarizes the way in which the North and the South financed their war efforts.
2. What happened to the budget deficits of European Union member countries during the financial crisis that started in 2008? Were these cyclically adjusted budget deficits?

3. Using the CBO as a source, make a table of the budget deficits for the period 1990 to the present in constant rather than current dollars (that is, obtain figures for real receipts, outlays, and deficits). Describe the behavior of real receipts, real outlays, and the real deficit over this period. Does it differ qualitatively from the description in the text? (If necessary, check the toolkit for instructions on how to convert nominal variables into real variables.)
4. Using the CBO as a source, make a table of the on-budget deficits for the period 1990 to the present. Compare these calculations with those reported in Table 29.2.2 "Recent Experience of Deficits and Surpluses (Billions of Dollars)". Explain the main differences between these tables.
5. Each month, the Congressional Budget Office (CBO) posts its monthly budget review. Look for the most recent monthly budget review. What are the largest outlays and revenues? How large are interest payments on the debt?
6. We saw that the government budget went from surplus to deficit in 2002. Based on the discussion in the text, try to find two different things that happened around this time that might explain this change.
7. This exercise builds on Table 29.5.1 "Budget Deficits around the World, 2005*".
    1. Find the levels of GDP in 2005 for each country listed in Table 29.5.1 "Budget Deficits around the World, 2005*". Using this information, find the ratio of the deficit to GDP for each of the countries.
    2. Which country in the world has the highest ratio of debt to GDP? How do the countries listed in Table 29.5.1 "Budget Deficits around the World, 2005*" compare in terms of the debt-to-GDP ratio?
    3. For the countries listed in Table 29.5.1 "Budget Deficits around the World, 2005*", find the growth rate of real GDP in 2005. Do countries that grow faster have smaller deficits? Hint: The CIA Fact Book (https://www.cia.gov/library/publications/the-world-factbook/index.html) will be useful.

Spreadsheet Exercises

1. Suppose that government purchases are 500 and the tax rate is 20 percent. Create a table to calculate the budget deficit for each level of income from 0 to 1,000, increasing by 50 each time. At what level of income does the budget balance? Compare your results to those in Table 29.3.2 "Deficit and Income".
2. Create a spreadsheet to study the debt as in Table 29.2.5 "Deficit and Debt" using the data from Table 29.2.1 "Calculating the Deficit". But assume that the level of debt outstanding at the start of the first period was 100, not 0. Assume that the interest rate is 2 percent each year. Add a column to Table 29.2.1 "Calculating the Deficit" to indicate the payment of interest on the debt. Calculate the deficit for each year and then the debt outstanding at the start of the next year. Also calculate the primary deficit in your spreadsheet. What happens to these calculations when the interest rate increases to 5 percent?

This page titled 29.7: End-of-Chapter Material is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

30: The Global Financial Crisis

30.1: A World in Crisis
30.2: The Financial Crisis in the United States
30.3: From Financial Crisis to Recession
30.4: The Crisis in Europe and the Rest of the World
30.5: Currency Crises
30.6: End-of-Chapter Material

This page titled 30: The Global Financial Crisis is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

30.1: A World in Crisis 30.1: A World in CrisisEdit Update titlecancel

The following quotation describes a meeting held in Washington, DC, among the G-20 countries.Associated Press, “World Leaders Pledge to Combat Global Crisis,” Minnesota Public Radio News, November 17, 2008, accessed July 25, 2011, http://minnesota.publicradio.org/display/web/2008/11/17/financial_meltdown. The G-20 countries are a group of 20 of the richest countries in the world.

President George W. Bush, who served as host for the G-20 discussions, said it was the seriousness of the current crisis that had convinced him that massive government intervention was warranted.

He said he felt “extraordinary measures” were needed after being told “if you don’t take decisive measures then it’s conceivable that our country could go into depression greater than the Great Depression.”

As we wrote this chapter in 2011, the world economy was slowly emerging from the worst financial crisis since the Great Depression. Economists and others formerly thought that the Great Depression was an interesting piece of economic history and nothing more. After all, they thought, we now understand the economy much better than did the policymakers at that time, so we could never have another Great Depression. But this belief that monetary and fiscal policymakers around the world knew how to ensure economic stability was shattered by financial turmoil that began in 2007, blossomed into a full-fledged global crisis in the fall of 2008, and led to sustained downturns in many economies in the years that followed.

That was the background to the November 2008 meeting of the G-20 countries. The world leaders attending that meeting were attempting to cope with economic problems that they had never even contemplated. The events that led to this meeting were unprecedented since the Great Depression, in part because of the magnitude and worldwide nature of the crisis.

As the quotation from President George W. Bush attests, extraordinary times prompted extraordinary action. The US government passed an “emergency rescue plan” in October 2008 to provide $700 billion in funding to (among other things) buy up assets of troubled banks and firms. This was followed by a large stimulus package, called the American Recovery and Reinvestment Act of 2009, which was passed during the first year of the Obama administration. Other countries brought in similar stimulus packages. Increased government expenditures and cuts in taxes were enacted by governments around the world. Monetary authorities also took extraordinary steps, with many countries rapidly reducing interest rates to very low levels. In addition, the US Federal Reserve and other monetary authorities engaged in other unprecedented policies in an attempt to provide liquidity to the financial system.

Although the roots of the crisis can be traced to 2007 or before, and although the implications of the crisis are still being felt, the full-fledged crisis began in 2008. As shorthand, we therefore refer to all these events as the “crisis of 2008,” and the question we ask in this chapter is as follows:

What happened during the crisis of 2008?

Road Map

In this chapter, we explore the policies enacted by governments to deal with the crisis. First we need a framework to understand these events. We make sense of the events of the past few years by drawing on the tools that we have developed in this book. We aim to do more than just give a narrative account of what happened; we also offer explanations of what happened. Whereas other chapters in this book are largely self-contained, this chapter is designed as a capstone. We therefore make frequent references to topics discussed in other chapters.

The crisis of 2008 was a highly complex event, with many different and imperfectly understood causes. Moreover, some of the details involve highly arcane aspects of financial markets. We are not going to give you a comprehensive account of the crisis. But we will show you how you can use the tools you have learned in this book to make some sense of what happened. We highlight three themes in particular.

1. As emphasized in Chapter 19, markets in the economy and around the world are interconnected. Various connections among markets caused the crisis to spill over across different financial markets, from financial markets into the real economy, and from the United States to economies all around the world. These are sometimes called “contagion problems.”
2. There were coordination failures in addition to contagion problems.
3. Monetary and fiscal policies are interconnected. We will see that responses to the crisis around the globe often required monetary and fiscal authorities to work together.

We start by summarizing events in the United States. In doing so, we use a tool from game theory to study how financial instability might arise. We use this framework to consider both recent events in the United States and events from the Great Depression. We then look specifically at the housing market at the start of the 21st century.

After understanding the experience in the United States, we study how the crisis spread from the United States to other countries. We stress both financial and trade links across countries as ways in which the crisis spread. We look at a few countries in particular, such as the United Kingdom, China, Iceland, and the countries of the European Union. The crisis in the European Union is particularly interesting to economists because the interconnections between the monetary and fiscal authorities are very different to those in other places. Finally, we consider exchange rates and currency crises.

This page titled 30.1: A World in Crisis is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

30.2: The Financial Crisis in the United States 30.2: The Financial Crisis in the United StatesEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What was the role of coordination games in the crisis?
2. What was the monetary policy response to the crisis?
3. What was the fiscal policy response to the crisis?

Starting in 2007 and stretching well into 2008, the United States and other countries experienced financial crises that resembled those of the Great Depression. Through the summer of 2011 (when this chapter was written), unemployment remained high, and real gross domestic product (real GDP) growth was low in the US economy. Some countries in Western Europe, such as Greece, were close to defaulting on their government debt.

One indicator of the seriousness of these events is the dramatic action that policymakers took in response. For example, on October 3, 2008, President George W. Bush signed into law the Emergency Economic Stabilization Act of 2008, which authorized the US Treasury to spend up to $700 billion for emergency economic stabilization. The full text of the bill and related facts are available at "Bill Summary & Status: 110th Congress (2007–2008) H.R.1424," THOMAS: The Library of Congress, accessed September 20, 2011, thomas.loc.gov/cgi-bin/bdquery/z?d110:h.r.01424:. As stated in the bill,

The purposes of this Act are—

1. to immediately provide authority and facilities that the Secretary of the Treasury can use to restore liquidity and stability to the financial system of the United States; and
2. to ensure that such authority and such facilities are used in a manner that—
    1. protects home values, college funds, retirement accounts, and life savings;
    2. preserves homeownership and promotes jobs and economic growth;
    3. maximizes overall returns to the taxpayers of the United States; and
    4. provides public accountability for the exercise of such authority.

This was an extraordinary amount of funding—equivalent to more than $2000 for every man, woman, and child in the United States. Perhaps even more strikingly, the funding was to allow the Treasury to do something it had never done before: to purchase shares (that is, become part owners) of financial institutions, such as banks and insurance companies. The United States, unlike some other countries, has never had many cases of firms being owned by the government. Moreover, in previous decades, the trend around the world has been for less government ownership of business—not more. It would have been almost unthinkable even a few months previously for a Republican president to have put in place mechanisms to permit this extent of government involvement in the private economy.

News accounts at the time made many different claims about the financial crisis, including the following:

- Banks and other financial institutions were failing.
- Housing prices had plummeted.
- So-called subprime mortgage loans had been made to borrowers in the early part of the decade, and the default rate on mortgages was rising because borrowers were no longer able to repay the loans.
- Low interest rates fueled asset bubbles that eventually burst.
- The financial crisis started in the United States but then spread to other countries.
- Stock markets around the world fell substantially.
- The next Great Depression might be around the corner.

Each news item has an element of truth, yet each can also mislead. We first sort through the events of 2008 and the policy responses. Then we look at the current state of the economy and at more recent policy actions.

Coordination Games and Coordination Failures

As discussed in Chapter 22, the United States and other economies experienced severe economic downturns in the early 1930s, together with instability in financial markets. It was little wonder that news accounts in 2008 and 2009 were filled with discussions of the parallels and differences between then and now. When we looked at financial instability during the Great Depression in Chapter 22, we studied a "bank-run game"—a strategic situation where depositors had to decide whether to leave their money in the bank or take it out. The bank-run problem is a leading example of a coordination game—a game with two key characteristics:

1. The game has multiple Nash equilibria.
2. These Nash equilibria can be ranked.

In a Nash equilibrium, everyone pursues their own self-interests given the actions of others. This means that no single individual has an incentive to change his or her behavior, given the choices of others. In a coordination game, there is more than one such equilibrium, and one of the Nash equilibria is better than the others. When the outcome of the coordination game is one of the outcomes that are worse than other possible equilibrium outcomes, then we say a **coordination failure** has occurred.

> **Toolkit: Section 31.18**
>
> Nash equilibrium is explained in more detail in the toolkit.

The possibility of coordination failure suggests two more fundamental questions:

1. What gives rise to these coordination games?
2. What can the government do about them?

Economists know that there are many situations that give rise to coordination games. Bank runs are just one example. In the crisis of 2008, actual bank runs did not occur in the United States, but they did happen in other countries. More generally, the financial instability that arose was similar in nature to a bank run. A recent article by Russell Cooper and Jonathan Willis explores in more detail the significance of coordination problems and beliefs during the recent crisis: "Coordination of Expectations in the Recent Crisis: Private Actions and Policy Responses," *Federal Reserve Bank of Kansas City Quarterly Review*, First Quarter 2010, accessed July 25, 2011, http://www.kansascityfed.org/PUBLICAT/ECONREV/PDF/10q1CooperWillis.pdf. Instead of failures of small neighborhood banks, we saw the failure or the near failure of major financial institutions on Wall Street, many of which had other banks as their clients. As noted by then president of the Federal Reserve Bank of New York, Timothy Geithner, the process of intermediation has gone beyond traditional banks to create a parallel (shadow) financial system in the United States: "The scale of long-term risky and relatively illiquid assets financed by very short-term liabilities made many of the vehicles and institutions in this parallel financial system vulnerable to a classic type of run, but without the protections such as deposit insurance that the banking system has in place to reduce such risks" From an address given by Geithner to the Economic Club of New York Press: "Timothy F. Geithner: Reducing Systemic Risk in a Dynamic Financial System," Bank for International Settlements, June 9, 2008, http://www.bis.org/review/r080612b.pdf?frames=0.

But the use of coordination games does not stop with bank runs. We can think of the decline in housing values as coming from a coordination failure. Even more strikingly, the circular flow of income itself can generate something that looks very like a coordination game. Imagine a situation where the economy is in a recession, with high unemployment and low levels of income. Because income is low, households choose low levels of spending. Because spending is low, firms choose low levels of production, leading to low income. By contrast, when income is high, then households engage in lots of spending. This leads firms to choose high levels of production, leading to high income.

What can governments do in the face of coordination games? One feature of these games is that the outcome of the game depends on the beliefs that people hold. An important aspect of economic policy may therefore be to support optimism in the economy. If people believe the economy is in trouble, this can be a self-fulfilling prophecy. But if they believe the economy is strong, they act in such a way that the economy actually is strong.

Crisis in the United States

There was no single root cause of the crisis of 2008. Economists and others have pointed to all sorts of factors that sowed the seeds of the crisis; we will not go through all these here. What is clear is that the housing market in the United States played a critical early role. As we saw in Chapter 19, events in the housing market were linked to events in the credit market, the labor market, and the foreign exchange market.

We begin with an equation that teaches us how the value of a house is determined. In Chapter 24, we explained that houses are examples of assets and that the value of any asset depends on the income that the asset generates. More specifically, the value of a house this year is given by the value of the services provided by the house plus the price of the house next year:

This equation tells us that three factors determine the value of a house. One is the flow of services that the house provides over the course of the coming year. In the case of a house that is rented out, this flow of services is the rental payment. If you own the home that you live in, you can think of this flow of services as being how much you would be willing to pay each year for the right to live in your house. That value reflects the size of the house, its location, and other amenities. The higher the flow of services from a house, the higher is its current price.

The second factor is the price you would expect to receive were you to choose to sell the house next year. If you expect housing prices to be high in the future, then the house is worth more today. This is true even if you do not actually plan to sell the house next year. One way of seeing this is to recognize that if you choose not to sell the house, its worth to you must be at least as large as that price.

The third factor is the interest rate—remember that the interest factor equals $(1 + the interest rate)$. The flow of services and next year's price both lie in the future, and we know that income in the future is worth less than income today. We use the technique of **discounted present value** to convert the flow of services and the future price into today's terms. As in the formula, we do so by dividing by the interest factor. One implication is that a change in the interest rate affects the current value of a house. In particular, a reduction in interest rates leads to higher housing prices today because a reduction in interest rates tells us that the future has become more relevant to the present.

Although we have written the equation in nominal terms, we could equally work with the real version of the same equation. In that case, the value of the service flow and the future price of the house must be adjusted for inflation, so we would use the real interest factor rather than the nominal interest factor.

Toolkit: Section 31.5

You can review discounted present value in the toolkit.

Now that you understand what determines the current value of a house, imagine you are making a decision about whether or not to buy a house. Unless you have a lot of cash, you will need to take out a mortgage to make this purchase. If interest rates are low, then you are more likely to qualify for a mortgage to buy a house. In the early 2000s, mortgage rates were relatively low, with the consequence that large numbers of households qualified for loans. In addition, many lenders offered special deals with very low initial mortgage rates (which were followed by higher rates a year or so later) to entice borrowers. The low interest rates encouraged people to buy houses. We saw this link between interest rates and spending in Chapter 25.

Lenders are also more willing to give you a mortgage if they think the price of a house is going to increase. Normally, you need a substantial down payment to get a loan. But if your mortgage lender expects housing prices to rise, then the lender will think that it will have the option of taking back the house and selling it for a profit if you cannot repay your mortgage in the future.

Thus, the expectation of rising housing prices in the future increases the current demand for houses and thus the current price of houses. In the early and mid-2000s, rising housing prices were seen in many markets in the United States and elsewhere. The rise in prices was fueled at least in part by expectations, in a manner that is very similar to a coordination game.

However, the optimism that underlies the price increases can at some point be replaced by pessimism, leading instead to a decrease in housing prices. Looking back at our equation for the value of a house, how can we explain the decrease in housing prices in 2007 and 2008? Interest rates did not rise over that time. It also seems unlikely that the service flow from a house decreased dramatically. This suggests that the main factor explaining the collapse of housing prices was a drop in the expected future price of houses. Notice the self-fulfilling nature of expectations: if everyone expects an asset to decrease in value in the future, it decreases in value today.

But what happens when housing prices start to decrease? Suppose you had put down $20,000 and borrowed $200,000 from a bank to buy a $220,000 home. If the price of your house decreases to, say, $150,000, you might just walk away from the house and default on the loan. Of course, default does not mean that the house disappears. Instead, it is taken over by the bank. But the bank does not want the house, so it is likely to try to sell it. When lots of banks find themselves with houses that they do not want, then the supply of houses increases, and the price of houses decreases.

We now see that there is a vicious circle operating:

1. Housing prices decrease.
2. People default on their loans.
3. Banks sell more houses in the market.
4. Housing prices decrease even more.

This again looks a lot like a coordination game. If housing prices are low, there are more mortgage defaults and thus houses put on the market for sale. The increased supply of houses drives down housing prices even further.

The crisis of 2008 may have begun in the housing market, but it did not stop there. It spread beyond housing to all corners of the financial markets. As explained in Chapter 24, a loan from your perspective is an asset from the perspective of the bank. Banks that held mortgage assets did not simply hold on to those assets, but neither did they merely sell them on to other banks. Instead, they bundled them up in various creative ways and then sold these bundled assets to other financial institutions. These financial institutions in turn rebundled the assets for sale to other financial institutions and so forth. The bundling of assets was designed to create more efficient sharing of the risk in financial markets. Fannie Mae (www.fanniemae.com/kb/index?page=home) and Freddie Mac (http://www.freddiemac.com), two government created and supported enterprises, were among those involved in the bundling and reselling of mortgages to facilitate this sharing of risks. These companies are currently in conservatorship. But there were also costs: (1) it became harder to evaluate the riskiness of assets, and (2) the original bank had a reduced incentive to carefully evaluate the loans that it made because it knew the risk would be passed on to others. This incentive problem made the bundles of mortgage loans riskier.

The Policy Response in the United States

The US government did not stand idle as these events were unfolding. They took the following actions: (1) they provided more **deposit insurance**, (2) they decreased interest rates, (3) they facilitated various mergers and acquisitions of financial entities, and (4) they bailed out some financial institutions. Some of these actions were an outgrowth of policies enacted after the Great Depression. The most important of these, deposit insurance, is discussed next.

Guarantee Funds and the Role of Deposit Insurance

In Chapter 22, we explained that, during the Great Depression, much of the disruption to the financial system came through bank runs. But in 2007 and 2008, we did not see bank runs in the United States. This was a striking difference between the crisis of 2008 and the Great Depression. The absence of bank runs is almost certainly because deposit insurance "changes the game." To see how, look at the bank-run coordination game in part (a) of Figure 30.2.1. In particular, look at the outcome if other players run and you do not run. In that case you get zero, so this would be a bad decision. You do better if you choose to participate in the run, obtaining 20. If everybody else chooses to run on the bank, you should do the same thing. In this case, the bank fails. But if everyone else leaves their money in the bank, you should do likewise. In this case, the bank is sound. The fact that there are two possible equilibrium outcomes is what makes this a coordination game.

Deposit insurance, which is run by the Federal Deposit Insurance Corporation (FDIC; http://www.fdic.gov/deposit), insures the bank deposits of individuals (up to a limit). Suppose that deposit insurance provides each depositor who leaves money in the bank a payoff of 110 *even if everyone else runs*. Now the game has the payoffs shown in part (b) of Figure 30.2.1. The strategy of "do not run" is now better than "run" regardless of what other people do. You choose "do not run"—as does everyone else in the game. The outcome is that nobody runs and the banks are stable. Remarkably, this policy costs the government nothing. Since there are no bank runs, the government never has to pay any deposit insurance. By changing the rules of the game, the government has made the bad equilibrium disappear.

(a)

	Others leave their money in the bank	Others take their money out of the bank
You leave your money in the bank (don't run)	110	0
You leave your money in the bank (don't run)	100	20

(b)

	Others leave their money in the bank	Others take their money out of the bank
You leave your money in the bank (don't run)	110	110
You leave your money in the bank (don't run)	100	20

Figure 30.2.1: The Payoffs in a Bank-Run Game with and without Deposit Insurance. You deposit $100 in the bank. Part (a) shows payoffs without deposit insurance. There are two Nash equilibria: if all people leave their money in the bank, then you should do the same, but if all people make runs on the bank, you are better running as well. In Part (b), deposit insurance means that the game has a unique equilibrium.

Decreasing Interest Rates

Deposit insurance may have prevented bank runs, but credit markets still did not function smoothly during the crisis of 2008. So what else was going on in credit markets? During the financial crisis, the Federal Reserve (the Fed) decreased its target interest rate. The way in which it does this and its implications for the aggregate economy are covered in Chapter 25. The Federal Open Market Committee (FOMC) reduced the target federal funds rate from 4.75 percent in September 2007 to 1.0 percent by the end of October 2008 and 0.25 percent by the end of the year. The target rate is indicated in the last column of the Table 30.2.1 "The Federal Funds Rate: Target and Realized Rates".

However, the Fed lost its usual ability to tightly control the actual federal funds rate. We see this in the other columns of Table 30.2.1 "The Federal Funds Rate: Target and Realized Rates". The column labeled "average" is the average federal funds rate over the day. The highest and lowest rates during the day are indicated as well. Prior to September 2008, the average and target rates were very close, but from mid-September onward, the average rate frequently diverged from the target. In addition, the difference between the high and low rates was much higher after the middle of September 2008.

Table 30.2.1: The Federal Funds Rate: Target and Realized Rates

Date	Average	Low	High	Target
October 14, 2008	1.1	0.25	2	1.5
October 7, 2008	2.97	0.01	6.25	2
September 29, 2008	1.56	0.01	3	2
September 15, 2008	2.64	0.01	7	2
July 16, 2008	1.95	0.5	2.5	2

Source: Data summarized from "Federal Funds Chart," Federal Reserve Bank of New York, 2008, http://www.newyorkfed.org/charts/ff.

As we explained in Chapter 25, these low interest rates meant that the Fed had hit the zero lower bound on monetary policy. Because nominal interest rates cannot be less than zero, the Fed was no longer able to stimulate the economy using the normal tools of monetary policy. Because its traditional tools were proving less effective than usual, the Fed turned to other, unusual, policy measures. The Fed created several lending facilities through which it provided funds to financial markets. For example, a commercial paper funding facility was created on October 7, 2008, to promote liquidity in a market that is central to the credit needs of both businesses and households. These are summarized by the Board of Governors, "Information Regarding Recent Federal Reserve Actions," Federal Reserve, accessed September 20, 2011, http://www.federalreserve.gov/newsevents/recentactions.htm. They are also discussed in the press release: "Press Release," Federal Reserve, October 7, 2008, accessed September 20, 2011,

http://www.federalreserve.gov/newsevents/press/monetary/20081007c.htm. The Board of Governors listed these as tools of the Fed in addition to three familiar tools: open-market operations, discount-window lending, and changes in reserve requirements.

Short-term interest rates, such as the overnight rate on interbank loans (the so-called LIBOR [London Interbank Offered Rate]), followed the Federal Funds rate down for much of 2008. (As the name suggests, this is the rate on loans that banks make to each other overnight.) But when the crisis became severe in September and October 2008, the LIBOR rose sharply. "US Dollar LIBOR Rates 2008," accessed September 20, 2011, http://www.global-rates.com/interest-rates/libor/american-dollar/2008.aspx. This rate averaged just about 5 percent for the month of September and 4.64 percent in October.The data are from the British Bankers' Association. Short-term rates increased despite the Fed's attempts to *reduce* interest rates.

Why did these rates not decrease along with the Fed's targeted federal funds rate? One explanation comes from the following equation:

$$loan\ rate \times probability\ of\ loan\ repayment = cost\ of\ funds\ to\ the\ bank.$$

On the left hand side is the loan rate charged by a bank—for example, the interest rate on a car loan, a household improvement loan, or a small business loan. The other term is the likelihood that the loan will actually be repaid. Together these give the expected return to the bank from making a loan. The right side is the cost of funds to the bank. This might be measured as the rate paid to depositors or the rate paid to other banks for loans from one bank to another. When this equation holds, the cost of the input into the loan process, measured as the interest cost on funds to the bank, equals the return on loans made. The bank does not then expect to make any profits or losses on the loan.

In Chapter 25, we argued that interest rates on loans usually follow the federal funds rate quite closely. If the Fed reduces the targeted federal funds rate, this reduces the cost of funds to banks. Banks typically follow by decreasing their lending rates. This close connection between the cost of funds and the loan rate holds true provided there is a stable probability of loan repayment. In normal times, that is (approximately) true, so variations in the federal funds rate lead directly to variations in loan rates.

During the fall of 2008, the link was weakened. Though the Fed reduced its targeted interest rate so that the cost of funds decreased, loan rates did not decrease. The reason was a fall in the perceived probability of loan repayment: banks perceived the risk of default to be much higher. Banks were cautious because they had suffered through the reduction in the value of mortgage-based assets and had seen some financial institutions fail. The state of the economy, with increasing unemployment and decreasing asset prices, led banks to be more prudent. In terms of our equation, the probability of repayment was decreasing at the same time as the cost of funds was decreasing. As a consequence, loan rates did not decrease as rapidly as the bank's costs of funds.

There was also a reduction in the amount of lending. The quantity of loans decreased because banks became more careful about whom to lend to. When you go to a bank to borrow, it makes an evaluation of how likely you are to repay the loan. During the fall of 2008, bank loans became much more difficult to obtain because many customers were viewed as higher credit risks. Even more significantly, the uncertainty of repayment was not limited only to loans from banks to households. Many of the loans in the short-term market are from banks to other banks or to firms. Uncertainty over asset valuations, growing out of the belief that some mortgage securities were overvalued, permeated the market, making lenders less willing to extend credit to other financial institutions.

Another factor in keeping interest rates high was the behavior of investors who held deposits that were *not* covered by federal deposit insurance—particularly deposits in "money market funds." In the early part of October 2008, there were huge outflows from money market funds into insured deposits as investors sought safety. This was a problem for the banking system, because it left the financial system with fewer funds to provide to borrowers. It also led short-term interest rates to rise.

So while the presence of deposit insurance was valuable in reducing the risk faced by individual households, banks still perceived higher lending risks. They therefore looked for ways to limit this risk. One prominent device they used is known as a *credit default swap*. This is a fancy term for a kind of financial insurance contract. The buyer of the contract pays a premium to the seller of the contract to cover bankruptcy risk. For example, suppose an institution owns some risky bonds issued by a bank. To shed this risk, the institution engages in a credit swap with an insurance provider.

These swaps played a big role in the stories of two key players in the financial crisis: American International Group (AIG) and Lehman Brothers. The US government eventually bailed out AIG, but Lehman Brothers went bankrupt. Because Lehman Brothers was an active trader of credit default swaps, their exit severely curtailed the functioning of this market. Without the added protection of these default swaps, lenders directly faced default risk and hence decided to charge higher loan rates. AIG was also a prominent player in the credit swap market.PBS Newshour provides a full report on AIG and credit swaps. They state that "AIG

wrote some $450 billion worth of credit default swap insurance" of the $62 trillion of credit swaps. "Risky Credit Default Swaps Linked to Financial Troubles," PBS Newshour, October 7, 2008, accessed September 20, 2011, www.pbs.org/newshour/bb/business/july-dec08/solmancredit_10-07.html. They sold insurance to cover many defaults, some linked directly to the holding of mortgages. As the mortgage crisis loomed, likely claims against AIG increased, putting them too on the brink of bankruptcy.

So although there were no bank runs during the crisis in the United States, credit markets were still severely disrupted. Given the centrality of the financial sector in the circular flow, disruption in the credit markets led to a downturn in overall economic activity. Put yourself in the place of a builder of new homes. Your customers are finding it hard to qualify for mortgages. As a result, the demand for your product is lower (the demand curve has shifted inward). Meanwhile (since the construction of a new home takes time), you need to borrow from a bank to finance payments to your suppliers of raw materials and to pay your carpenters and other workers. Tight credit markets mean that you find it more expensive to obtain funds: interest rates are higher, and the terms are less generous. Not surprisingly, disruption in the credit markets shows up particularly starkly in the market for new houses.

Facilitating Takeovers of Financial Firms

The problems of AIG, Lehman Brothers, and other financial firms led policymakers to worry about such firms going bankrupt. In some cases, these firms had too many bad assets on their books and were not able to continue in the market. One example is Bear Stearns, which was heavily involved in the trading of assets that were backed by mortgages. In March 2008, it became clear that those assets were highly overvalued. When the prices of these assets decreased, Bear Stearns was close to bankruptcy. With the help of a loan (http://www.federalreserve.gov/newsevents/press/other/other20080627a2.pdf) from the Board of Governors of the Federal Reserve (operating through the Federal Reserve Bank of New York), JPMorgan Chase and Company acquired Bear Stearns.

It is perhaps remarkable that the Fed took such an active role in this acquisition. When a local grocer goes out of business, you simply shift your business to another seller. Nobody expects the government to take a role in rescuing the store. But when we are talking about large financial firms, shifting from one financial intermediary to another may not be as easy. When a large institution fails, it is highly disruptive to the financial system as a whole. The minutes of the March 16, 2008, meeting of the Board of Governors confirm this view:

The evidence available to the Board indicated that Bear Stearns would have difficulty meeting its repayment obligations the next business day. Significant support, such as an acquisition of Bear Stearns or an immediate guarantee of its payment obligations, was necessary to avoid serious disruptions to financial markets.

Thus the Fed thought it was necessary to ensure the takeover of Bear Stearns and hence the continuation of its operations. In fact, prior to this takeover, Bear Stearns was listed among a small set of financial firms as "primary dealers." These are financial intermediaries that are viewed as central to the orderly operation of financial markets and the conduct of monetary policy. The list of primary dealers is available at "Primary Dealers List," Federal Reserve Bank of New York, accessed July 25, 2011, http://www.newyorkfed.org/markets/pridealers_current.html.

AIG received a loan up to $85 billion from the Fed in September 2008. The monetary authority was concerned that a failure of AIG would further destabilize financial markets. As part of this deal, the US government acquired a 79.9 percent equity ownership in AIG. See the announcement by the Board of Governors: "Press Release," Board of Governors of the Federal Reserve System, September 16, 2008, accessed July 25, 2011, http://www.federalreserve.gov/newsevents/press/other/20080916a.htm. As authority for this action, this announcement by the Fed cites Section 13(3) of the Federal Reserve Act, which allows the Fed to provide this type of funding in "unusual and exigent" circumstances: "Section 13: Powers of Federal Reserve Banks," Board of Governors of the Federal Reserve System, December 14, 2010, accessed July 25, 2011, http://www.federalreserve.gov/aboutthefed/section13.htm. AIG was special enough to warrant this government loan because of its role in providing insurance (through credit default swaps) against the default on debt by individual companies. Without this insurance, the debt of these companies becomes riskier, and they find it harder to borrow. AIG was a large enough actor in this market for its departure to have meant severe disruptions in the provision of insurance.

In contrast to these actions for Bear Stearns and AIG, the Fed did nothing to help Lehman Brothers, a 158-year-old financial firm. It went out of business in September 2008. There was no bailout for this company from the Fed or the US Treasury. It simply disappeared from the financial markets.

A $700 Billion Bailout

In October 2008, Congress passed and the president signed legislation called the "Emergency Economic Stabilization Act of 2008" to provide $700 billion in funding available to the Department of the Treasury. The legislation authorized the Treasury (www.house.gov/apps/list/press/financialsvcs_dem/essabill.pdf) to purchase mortgages and other assets of financial institutions (including shares) to create a flow of credit within the financial markets. The Treasury Department then set up a Troubled Asset Relief Program as a vehicle for making asset purchases. In addition to these measures, the legislation called for an increase in FDIC deposit insurance to cover deposits up to a cap of $250,000 instead of the standard cap of $100,000. See "Insured or Not Insured?," Federal Deposit Insurance Corporation, accessed July 25, 2011, http://www.fdic.gov/consumers/consumer/information/fdiciorn.html. The FDIC does not insure money market funds, though these were protected under a temporary Treasury program: "Frequently Asked Questions about Treasury's Temporary Guarantee Program for Money Market Funds," US Department of the Treasury, September 29, 2008, accessed September 20, 2011, http://www.treasury.gov/press-center/press-releases/Pages/hp1163.aspx.

One interesting element of the bailout legislation was the explicit interaction of the Treasury and the Fed. A joint statement issued after the passage of this act indicated that these players in the conduct of fiscal and monetary policy were working together to resolve the crisis. In the United States, the Treasury and the Fed each contributed to the financing of these rescue packages.

In Chapter 26, we pointed out that

$$government\ deficit = change\ in\ government\ debt + change\ in\ money\ supply.$$

In other words, when the government runs a deficit, it must finance this deficit by either issuing more debt or printing money. This equation is consistent with the institutional structure in the United States where the Treasury and the Fed are independent entities. In effect, the Treasury issues debt to finance a deficit, then some of that debt is purchased by the Fed. When the Fed purchases debt, it injects new money into the economy.

Key Takeaways

1. Though there were no bank runs in the United States during the crisis of 2008, the structure of coordination games is useful for thinking about instability of the housing sector, the interactions of banks within the financial system, and the interaction between income and spending.
2. During the crisis, the Fed moved aggressively to decrease interest rates and provide liquidity to the system.
3. The George W. Bush administration created a $700 billion program to purchase or guarantee troubled assets, such as mortgages and shares of financial firms.

Exercises

1. As the probability of default increases, what happens to the lending rate?
2. What is a credit default swap?
3. Why does the sale of bank-owned houses cause the price of houses to decrease?
4. We said that deposit insurance was available in 2008. Was it available during the Great Depression?

This page titled 30.2: The Financial Crisis in the United States is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

30.3: From Financial Crisis to Recession 30.3: From Financial Crisis to RecessionEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. How did the financial crisis spread to the aggregate economy?
2. What was the fiscal policy response?
3. What was the monetary policy response?

So far we have focused on the financial side of the crisis of 2008 because the initial stage of the crisis was within the financial sector. As in the Great Depression, though, the disruptions in the financial sector then spread to the rest of the economy.

From Housing to the Aggregate Economy

The crisis of 2008 saw financial disruptions spread from financial markets to the economy at large. In Chapter 22, we introduced the **aggregate expenditure model** to understand the reduction in economic activity in the early 1930s. That same framework is useful in understanding recent events.

Toolkit: Section 31.30

You can review the aggregate expenditure model in the toolkit.

The aggregate expenditure model takes as its starting point the fact that gross domestic product (GDP) measures both total spending and total production. When planned and actual spending are in balance,

$$real\ GDP = planned\ spending = autonomous\ spending + marginal\ propensity\ to\ spend \times real\ GDP.$$

Autonomous spending is the intercept of the planned spending line. It is the amount of spending that there would be in the economy if income were zero. The equilibrium level of real GDP is as follows:

The framework tells us that a reduction in autonomous spending leads to a decrease in real GDP.

Just as in the Great Depression, the two leading candidates for the decrease in autonomous spending are consumption and investment. Specifically, the crisis in the housing market had two significant implications for the rest of the economy. First, the decrease in housing prices starting in 2008 reduced the wealth of many households. Because households were poorer, they reduced their consumption. Second, the disruptions in the financial system made it difficult for firms to obtain financing, which meant that there was less investment. The aggregate expenditure model teaches us that these reductions in consumption and investment can lead to a reduction in real GDP.

Reductions in autonomous spending are magnified through the **circular flow of income**. As spending decreases, income decreases, leading to further reductions in spending. This is the **multiplier** process; it shows up as the term $1/(1 - marginal\ propensity\ to\ spend)$, which multiplies autonomous spending in the expression for real GDP.

Toolkit: Section 31.27

You can review the circular flow of income and the multiplier in the toolkit.

Stabilization Policy

We have already observed that, in contrast to the Great Depression, policymakers in the crisis of 2008 took several actions to try to address the economic problems. In addition to the measures aimed specifically at dealing with problems in the financial markets, policymakers turned to monetary and fiscal policy in an attempt to counteract the economic downturn.

To begin our discussion of this stabilization policy, it is useful to start with a summary of the state of the economy in the 2006–10 period. By so doing, we are making life somewhat easier for us than it was for policymakers because they did not know in early 2009 what would happen in the aggregate economy during that year. The annual growth rates of the main macroeconomic variables

during the crisis are highlighted in Table 30.3.1 "State of the Economy: Growth Rates from 2006 to 2010". All variables are in percentage terms.

From Table 30.3.1 "State of the Economy: Growth Rates from 2006 to 2010", you can see how US real GDP growth slowed in 2007, stalled in 2008, and turned negative in 2009. The recovery in 2010 had a positive growth rate slightly larger than the decline in 2009. Had these growth rates been identical in absolute value, the economy would have recovered, roughly speaking, to the 2008 level of real GDP. The annual growth rate of real GDP in the last quarter of 2010 was a robust 3.1 percent, but the growth rate in the first quarter of 2011 was only 1.8 percent. Concerns remain over the viability of the current recovery.

The next four columns of Table 30.3.1 "State of the Economy: Growth Rates from 2006 to 2010" show that the declines in real GDP came largely from spending on investment and durables by firms and households. Housing played a particularly significant role. This fits with the theory of consumption smoothing that we discussed in Chapter 27 and Chapter 28. The last column shows the unemployment rate. Although the economy enjoyed positive real GDP growth in 2010, the unemployment rate remained high.A recent BLS publication looked at job creation and job destruction up to 2009 to try to understand the slow recovery of unemployment. “Business Dynamics Statistics Briefing: Historically Large Decline in Job Creation from Startup and Existing Firms in the 2008–2009 Recession,” U.S. Census Bureau’s Business Dynamics Statistics, March 2011, accessed September 20, 2011, http://www.ces.census.gov/docs/bds/plugin-BDS%20March%202011%20single_0322_FINAL.pdf.

Table 30.3.1: State of the Economy: Growth Rates from 2006 to 2010

Year	GDP	Consumption	Household Durables	Investment	Housing	Unemployment Rate (%)
2006	2.7	2.9	4.1	2.7	−7.3	4.4
2007	1.9	2.4	4.2	−3.1	−18.7	5.0
2008	0.0	−0.3	−5.2	−9.5	−24.0	7.4
2009	−2.6	−1.2	−3.7	−22.6	−22.9	10.0
2010	2.9	1.7	7.7	17.1	-3.0	9.6

Source: Bureau of Economic Analysis, Department of Commerce (www.bea.gov/newsreleases/national/gdp/2010/txt/gdp2q10_adv.txt and www.bea.gov/newsreleases/national/gdp/2011/pdf/gdp1q11_2nd.pdf) and Bureau of Labor Statistics (http://www.bls.gov/cps).

Fiscal Policy

One of the priorities of the Obama administration after taking office in January 2009 was to formulate a stimulus package to deal with the looming recession. As is clear from Table 30.3.1 "State of the Economy: Growth Rates from 2006 to 2010", growth in the economy was near zero for the preceding year, and the unemployment rate was much higher than it had been in the previous two years. Although the financial rescue plans of the George W. Bush administration may have stemmed the financial crisis, the aggregate economy was now limping along at best.

The American Recovery and Reinvestment Act of 2009 (ARRA) was signed into law on February 17, 2009. The stimulus package contained approximately $800 billion in spending increases and tax cuts. These numbers are approximate for a couple of reasons: (1) parts of the package depend on the state of the economy in the future, so the exact outlays are not determined in the legislation, and (2) the disbursements were not all within a single year, so the timing of the outlays and thus their discounted present value could not be precisely known at the time of passage.

The package contained a mixture of spending increases and tax cuts. According to a Congressional Budget Office (CBO) study (www.cbo.gov/ftpdocs/106xx/doc10682/Frontmatter.2.2.shtml) from November 2009, federal government purchases of goods and services were to increase by about $90 billion over the 2009–19 period. Transfer payments to households were set to increase by about $100 billion, and transfers to state and local governments were to increase by nearly $260 billion. This last category of outlays was quite visible, taking the form of road projects and other construction in towns across the United States. Interestingly, the federal government was investing in infrastructure, thus building up the public component of the capital stock.

In the same publication, the CBO provided a summary of ARRA’s macroeconomic effects in November 2009. At that point, due to ARRA, the CBO estimated that federal government outlays (not only spending on goods and services) had increased by about $100

billion, and tax collections were lower by about $90 billion. So clearly some but not the entire stimulus went into the US economy within seven months of ARRA's passage. The CBO also produced its own assessment of the effects of ARRA through September 2009. To do so, it had to use an economic model to calculate the effects of the increases in outlays and reductions in taxes. In many ways, the framework for the assessment is quite similar to the analysis of the Kennedy tax cuts we discussed in Chapter 27.

According to the CBO, ARRA meant that real GDP in the United States was between 1.2 percent and 3.2 percent higher than it would otherwise have been, whereas the unemployment rate was between 0.3 and 0.9 percentage points lower. These numbers were obtained by attaching a multiplier to each component of the stimulus package and calculating the change in real GDP from that component. For example, the CBO estimated that the multiplier associated with federal government purchases of goods and services was between 1 and 2.5. The effect of this federal spending on real GDP is simply the product of the spending of the federal government funded under ARRA times the multiplier. The CBO did this calculation for each component of the stimulus package and then added up the effects on real GDP. The range of the effects reflects the range for each multiplier used in their analysis. The CBO also calculated that 640,000 jobs were either created or retained due to ARRA. This calculation underlies their estimate of how much ARRA reduced the unemployment rate in the United States.

Some economists have disputed the effects of ARRA on economic activity, however. John Taylor, a Stanford University economist, argued that the short-term nature of the tax cuts meant that most households simply saved the tax cut, as the theory of consumption smoothing predicts. This argument was supported by evidence of increasing saving rates by households in the United States during the period of the tax cuts. Testimony reproduced in John B. Taylor, "The 2009 Stimulus Package: Two Years Later," Hoover Institution, February 16, 2011, accessed September 20, 2011, http://media.hoover.org/sites/default/files/documents/2009-Stimulus-two-years-later.pdf.

During 2010 and 2011, there were some calls for further stimulus. The unemployment rate in the United States remained high despite the stimulus; it was 9.5 percent in July 2010. The Bureau of Labor Statistics (http://www.bls.gov/news.release/empsit.b.htm) tells us that while job creation had been brisk in May 2010 at 432,000 jobs, the total job destruction in June and July 2010 was 350,000. Further, real GDP growth was only 2.4 percent in the second quarter of 2010, down from 3.7 percent in the first quarter. Together this news put more pressure on policymakers to conduct further attempts at stabilization policy.

But at the same time, policymakers became increasingly concerned about the long-run fiscal health of the government. In effect, they began to worry about the government budget constraint, which we explained in Chapter 29. The attention of policymakers moved away from stimulus and toward "fiscal consolidation." This culminated in a political battle in the summer of 2011 over an increase in the debt ceiling, a limit on the amount of US debt outstanding. Ultimately an agreement was reached to allow an increase in the ceiling, but this agreement was combined with a reduction in government spending of nearly $900 billion over the coming 10 years and an agreement to seek further cuts in spending amounting to another $1.5 trillion.The bill passed by the House of Representatives is contained here: "Text of Budget Control Act Amendment," House of Representatives Committee on Rules, accessed September 20, 2011, http://rules.house.gov/Media/file/PDF_112_1/Floor_Text/DEBT_016_xml.pdf. This agreement was not enough to avert a downgrade of US debt from AAA to AA+ by Standard and Poors. The decision to downgrade the debt is discussed at "Research Update: United States of America Long-Term Rating Lowered to 'AA+' on Political Risks and Rising Debt Burden; Outlook Negative," Standard & Poor's, August 5, 2011, accessed September 20, 2011, www.standardandpoors.com/servlet/BlobServer?blobheadername3 =MDT-Type&blobcol=urldata&blobtable=MungoBlobs&blobheadervalue 2=inline%3B+filename%3DUS_Downgraded_AA%2B.pdf&blobheadername2=Content-Dis position&blobheadervalue1=application%2Fpdf&blobkey=id&blob headername1=content-type&blobwhere=1243942957443&blobh eadervalue3=UTF-8.

Monetary Policy

The current state of monetary policy is well summarized in the Federal Open Market Committee (FOMC) statement of August 10, 2010. Here is an excerpt: "Press Release," Federal Open Market Committee, August 10, 2010, accessed July 26, 2011, http://www.federalreserve.gov/newsevents/press/monetary/20100810a.htm.

Press Release

Release Date: August 10, 2010

For immediate release

Information received since the Federal Open Market Committee met in June indicates that the pace of recovery in output and employment has slowed in recent months.… Nonetheless, the Committee anticipates a gradual return to higher levels of resource utilization in a context of price stability, although the pace of economic recovery is likely to be more modest in the near term than had been anticipated.

Measures of underlying inflation have trended lower in recent quarters and, with substantial resource slack continuing to restrain cost pressures and longer-term inflation expectations stable, inflation is likely to be subdued for some time.

The Committee will maintain the target range for the federal funds rate at 0 to 1/4 and continues to anticipate that economic conditions, including low rates of resource utilization, subdued inflation trends, and stable inflation expectations, are likely to warrant exceptionally low levels of the federal funds rate for an extended period.

To help support the economic recovery in a context of price stability, the Committee will keep constant the Federal Reserve's holdings of securities at their current level by reinvesting principal payments from agency debt and agency mortgage-backed securities in longer-term Treasury securities.…

Voting for the FOMC monetary policy action were: Ben S. Bernanke, Chairman; William C. Dudley, Vice Chairman; James Bullard; Elizabeth A. Duke; Donald L. Kohn; Sandra Pianalto; Eric S. Rosengren; Daniel K. Tarullo; and Kevin M. Warsh.

Voting against the policy was Thomas M. Hoenig, who judges that the economy is recovering modestly, as projected. Accordingly, he believed that continuing to express the expectation of exceptionally low levels of the federal funds rate for an extended period was no longer warranted and limits the Committee's ability to adjust policy when needed.…

We can make several observations about this FOMC statement. First, the FOMC shared the general perception that the recovery is not very robust and is showing signs of slowing. Their response was to maintain the targeted federal funds rate at between 0 and 0.25 percent. The FOMC put the targeted rate into this range in December 2008; in August 2011 the Fed indicated that it would keep rates low for at least another two years. You can find the FOMC statements and minutes of the meetings from December 2008 onward at "Meeting Calendars, Statements, and Minutes (2006–2012)," Federal Reserve, accessed July 26, 2011, http://www.federalreserve.gov/monetarypolicy/fomccalendars.htm.

Second, the FOMC talks about "reinvesting principal payments from agency debt and agency mortgage-backed securities…." This somewhat complicated phrase refers to the fact that the Fed purchased various assets in the attempt to keep financial markets working during the financial crisis. Those programs are summarized at "Credit and Liquidity Programs and the Balance Sheet," Federal Reserve, accessed July 26, 2011, http://www.federalreserve.gov/monetarypolicy/bst_crisisresponse.htm. As reported by the Fed, "[s]ince the beginning of the financial market turmoil in August 2007, the Federal Reserve's balance sheet has grown in size and has changed in composition. Total assets of the Federal Reserve have increased significantly from $869 billion on August 8, 2007, to well over $2 trillion." The Fed maintains an interactive web site that displays and explains its balance sheet items. "Credit and Liquidity Programs and the Balance Sheet," Federal Reserve, accessed July 26, 2011, http://www.federalreserve.gov/monetarypolicy/bst_recenttrends.htm. Observers are waiting for the Fed to reduce its holdings of these assets. The policy statement indicated that the Fed was not yet ready to take those steps.

The final point concerns the position of Thomas Hoenig, the president of the Federal Reserve Bank of Kansas City. Over the year, he took the view that monetary policy was too lax. As the economy recovered, there was, he believed, no longer any need to keep interest rates at such low levels. One of the implicit concerns here is that periods of low interest rates have tended to promote bubbles in assets, such as housing. The FOMC had to weigh this concern against the view that, with a slow economic recovery and no signs of inflation, expansionary monetary policy was still warranted. When the FOMC took the unusual decision to commit to low interest rates for two years, three members of the committee dissented from the decision.

Key Takeaways

1. Disruptions in the financial system led to reductions in consumption and investment, which led to a decrease in real GDP.
2. An $800 billion stimulus package was passed in February 2009 to offset the recessionary effects of the financial crisis.
3. From December 2008 through (at least) the summer of 2011 the target federal funds rate was near zero.

Exercises

1. If the federal funds rate is near zero, what is the real return on a loan in that market? (Hint: if you are not sure about the answer, look up the Fisher equation in the toolkit for more information.)
2. In Table 30.3.1 "State of the Economy: Growth Rates from 2006 to 2010", what items are counted as investment?
3. If the CBO calculates that the multiplier on tax cuts is 1.5 and taxes are cut by $100 billion, how much will GDP change by?

This page titled 30.3: From Financial Crisis to Recession is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

- **Current page** by Anonymous is licensed CC BY-NC-SA 3.0. Original source: https://2012books.lardbucket.org/books/theory-and-applications-of-economics.
- **23.3: Job and Worker Flows** by Anonymous is licensed CC BY-NC-SA 3.0. Original source: https://2012books.lardbucket.org/books/theory-and-applications-of-economics.

30.4: The Crisis in Europe and the Rest of the World 30.4: The Crisis in Europe and the Rest of the WorldEdit Update titlecancel

Learning Objectives

After you have read this section, you should be able to answer the following questions:

1. What are the ways the crisis spread from the United States to the rest of the world?
2. In what ways did the institutional structure of the European Union (EU) hamper Europe's ability to cope with the crisis?

In Chapter 24, we spoke of the day when people in several European countries woke up to a new monetary regime that used different pieces of paper than were used previously. In that chapter we used that experience to help us understand why people hold money. When these countries adopted the euro, they were not expecting to wake up about a decade later to read something like this:

"On the eve of a confidence vote that may bring down Papandreou's government, euro-area finance ministers pushed Greece to pass laws to cut the deficit and sell state assets. They left open whether the country will get the full 12 billion euros ($17.1 billion) promised for July" as part of last year's 110 billion-euro lifeline.

"We forcefully reminded the Greek government that by the end of this month they have to see to it that we are all convinced that all the commitments they made are fulfilled," Luxembourg Prime Minister Jean-Claude Juncker told reporters early today after chairing a crisis meeting in Luxembourg. James G. Neuger and Stephanie Bodoni, "Bailout Bid for Greece Falters as Europe Insists Papandreou Cut Budget Gap," June 20, 2011, accessed July 26, 2011, http://www.bloomberg.com/news/2011-06-20/europe-fails-to-agree-on-greek-aid-payout-pressing-papandreou-to-cut-debt.html; Karen Kissane, "EU puts brakes on loan to Greece," *Sydney Morning Herald*, June 21, 2011, accessed July 26, 2011, http://www.smh.com.au/world/eu-puts-brakes-on-loan-to-greece-20110620-1gbxw.html; James Neuger and Stephanie Bodoni, "Europe Fails to Agree on Greek Aid Payout," June 20, 2011, accessed September 20, 2011, http://www.bloomberg.com/news/2011-06-20/europe-fails-to-agree-on-greek-aid-payout-pressing-papandreou-to-cut-debt.html.

The euro was established by the Maastricht Treaty, but the implications of that treaty went beyond the introduction of new pieces of paper. The nature of fiscal and monetary interactions across the countries within the Economic and Monetary Union (EMU) changed dramatically as well.

On the monetary side, in addition to losing their national currencies, the countries that joined the euro effectively lost their central banks. The Central Bank of Italy, say, which formerly conducted monetary policy in that country, handed over that duty to the European Central Bank (ECB). The same thing happened in other countries. Most significantly, the German Bundesbank, which was one of the most important central banks in the world, also ceded its powers to the ECB. Further, the Maastricht Treaty—and the Stability and Growth Pact that followed a few years later—placed restrictions on fiscal policy by member countries. For a discussion of the history and content of the Stability and Growth Pact, see "Stability and Growth Pact," European Commission Economic and Financial Affairs, accessed September 20, 2011, http://ec.europa.eu/economy_finance/sgp/index_en.htm. Prior to the introduction of the euro, member governments had complete discretion over their fiscal policy. Within the EMU, however, constraints on deficit spending were placed on member countries.

Taken together, these two factors radically changed the conduct of monetary and fiscal policy in the countries of the EMU. Some commentators questioned whether adequate tools for stabilization of aggregate economies were still available. Others wondered whether the constraints on fiscal policy would be violated by member countries, leading to the possibility of a debt crisis for a country within the euro area. In that event, how would the other member countries respond?

The crisis of 2008 provided the first big tests of these questions. Debt problems—not only in Greece but also in Portugal and Ireland—revealed that these concerns were well placed. We start by discussing how the crisis spread from the United States to Europe and then turn to the policy actions within Europe.

Sources of Spillovers

In Chapter 20, we explained how countries are linked through the flows across national borders of goods, services, labor, financial capital, and information. Countries do not exist in isolation, and these linkages imply that problems in one country can be transmitted to others. In the crisis of 2008, we can point to three broad channels of spillover from the United States to the rest of the world:

1. Within financial markets across borders (integrated financial markets)
2. From financial markets into real markets in the United States, followed by real spillovers across countries
3. Contagion effects through market psychology

The first two linkages can be seen in the circular flow of income in Figure 30.4.1. In this version of the circular flow, we highlight the interactions between a single country and the rest of the world. These interactions operate through the flows of goods and services and financial assets. During good times, they are a key part of the workings of the world economy. But during bad times, such as a financial crisis, these same links create channels for the sharing of financial crises.

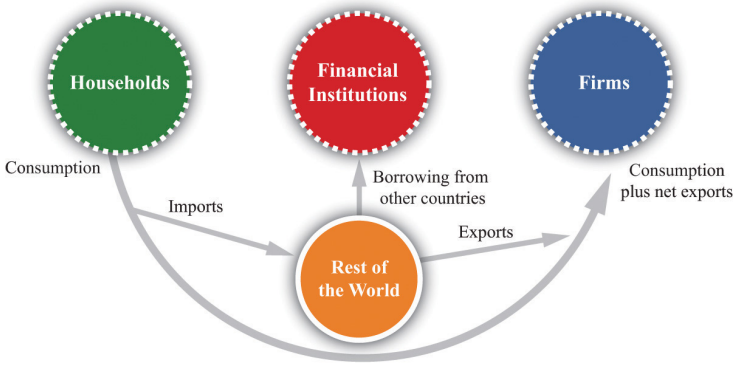


Figure 30.4.1: The Foreign Sector in the Circular Flow. Households purchase goods from other countries; these are called imports. Citizens of other countries purchase our products; these are called exports. A trade deficit requires borrowing from the rest of the world.

There are three international flows in Figure 30.4.1:

1. **Exports.** Households, firms, and the government in the rest of the world purchase goods and services produced in the home country.
2. **Imports.** Households (and also firms and the government) purchase goods and services produced in the rest of the world.
3. **Financial flows.** Financial intermediaries in the home country buy and sell financial assets flows from/to the rest of the world. The net flow can go in either direction; Figure 30.4.1 shows the case where there is a net flow of money into the home country.

International Spillovers in Financial Markets

One channel through which the crisis of 2008 spread was the holding of US financial assets by governments, financial institutions, and banks in other countries. Take, for example, mortgages that were marketed and issued in the United States. These mortgages were usually not ultimately held by the banks that issued them to homeowners. Instead they were bundled together with other mortgages and then resold.

These “mortgage-backed securities” were marketed and sold all over the world, not just in the United States. This means that any risk associated with these assets was shared across investors in different countries. The spread of this risk across world markets also provided a way for the crisis to propagate across countries. When it became clear that these assets were less valuable than investors had previously thought, the reduction in their price reduced the wealth of investors all over the globe. Moreover, the various financial institutions in the United States that were either bought out or went bankrupt were partly owned by investors in other countries. Thus financial links across the world economy provided one avenue for the spread of the crisis.

Second, the financial flows across countries played a significant role in the spread of the crisis. Since the early 1970s, the United States has run current account deficits each year. One consequence of this is that it has been borrowing from abroad to finance these deficits. In other words, foreigners hold substantial amounts of US assets. These assets include US government debt and, in many cases, large amounts of mortgage-backed securities.

One way to see the extent of these financial interactions is to look at the behavior of stock markets around the globe. Figure 30.4.2 shows the values for six indices around the world: the Dow Jones Industrial Average (United States), CAC (France), FTSE (United Kingdom), Hang Seng (China), Nikkei (Japan), and Merval (Argentina). The figure shows that the last six months of 2008 were problematic for stock markets across many countries.

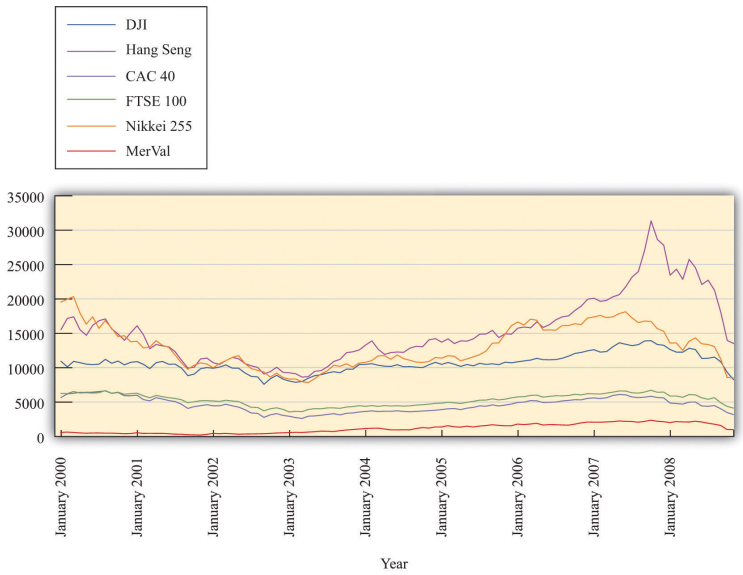


Figure 30.4.2: Stock Markets around the World Crashed Together

Spillovers through the Trade of Goods and Services

Trade is another source of linkage across countries. Because countries sell goods and services to each other, a recession in one country will naturally spread to others. If the major trading partners of a country are in a recession, then there will be a reduced demand for the goods and services produced by that country. So, for example, if the United States enters into a recession as a consequence of financial market distress, then the demand for goods and services produced in other countries will decrease. This reduction in aggregate spending in other countries will then lead to lower economic activity in those countries.

Spillovers through Expectations

The circular flow of income shows two of the three spillovers we have identified: financial flows and trade flows. The third spillover has to do with people's perceptions and expectations about market outcomes. There are two parts to this linkage: (1) expectations matter, and (2) outcomes in one market can have effects in others. The second of these is termed a **contagion effect**.

To the extent that part of the financial distress is due to pessimism, as suggested by the coordination game we discussed in 30.2 Section "The Financial Crisis in the United States", this too is likely to spread across countries. If, day after day, the news from the United States is that the prices of stock and other assets are decreasing, investors in other countries may begin to share this pessimism. This will lead them to sell their assets, leading to decreases in the prices of the assets that they are selling. Decreasing asset prices can feed on themselves through pessimistic expectations. As an example, consider again the September 2008 bankruptcy of Lehman Bros. Landon Thomas Jr., "Examining the Ripple Effect of the Lehman Bankruptcy, *New York Times*, September 15, 2008, accessed July 26, 2011, www.nytimes.com/2008/09/15/business/worldbusiness/15iht-lehman.4.16176487.html.

"Everybody is frozen here after Lehman," said one senior executive from a major financial institution who was paying visits this week to all the major sovereign funds in Asia and the Middle East. His voice was worn from hours spent in conference rooms trying to explain to clients why Lehman failed and who might be next. "Its just fear."

In 30.2 Section "The Financial Crisis in the United States", we gave an equation to explain the price of an asset—specifically, a house. A key part of that equation is that the value of a house today depends in part on the price of the house expected in the future. To emphasize this key point again: if you think that people will pay a lot for a house a year from now, you will be willing to pay a lot for it today. The logic applies to all other assets as well, so there is a link between prices expected for the future and prices today. Think about a stock that you might buy on the New York Stock Exchange. The stock yields a dividend and also has a future price. The higher the price you expect the stock can sell for in the future, the more you are willing to pay for the stock today. Expectations matter.

But where do these expectations come from? During normal times, expectations are disciplined by the usual state of a market. If housing prices have been rising by say 3 percent a year for the past 20 years, most people will predict that over the next year, housing prices will again rise by 3 percent. Most of the time that prediction will be roughly right—but not all the time. Sometimes

markets are subject to unpredictable movements in prices. When asset prices decline rapidly and unexpectedly, this is often referred to as a "bubble bursting."

All this discussion suggests that asset prices can be somewhat fragile—and this is where contagion effects can come into play. If you are trading houses in one location and the prices of houses in other locations are all decreasing quickly, you might get concerned that whatever is hurting housing values in those markets will affect yours as well. If so, you might be tempted to try to sell the houses that you own. Of course, others think the same way. As a consequence, the price of houses in your location decreases as well. This is the contagion effect: the behavior of prices in other markets influences expectations in your market and leads to a price reduction in your market. You and the other market participants who feared a decrease in prices are, in the end, correct. In that sense, contagion effects can be self-fulfilling prophesies.

The Crisis in the EMU

When the crisis hit in the United States, the secretary of the Treasury and the chairman of the Board of Governors of the Federal Reserve System held a joint press conference. By so doing, they made it clear that, in the United States, monetary and fiscal policy were being used jointly to resolve the financial crisis. In Europe, the picture was very different. As we have explained, countries within the EMU have a common central bank but do not have a common fiscal policy. Fiscal policy is decided at the country level, while the ECB is supposed to target the overall European inflation rate (as discussed in Chapter 25) and is not supposed to play any role in bailing out individual governments.

This system may work in normal times. The events of the crisis of 2008 revealed that it did not work so well in abnormal times; this in turn led to some calls for change.

Sarkozy Calls for "Economic Government" for Eurozone

French President Nicolas Sarkozy called Tuesday for "clearly identified economic government" for the eurozone, working alongside the European Central Bank.

"It is not possible for the eurozone to continue without clearly identified economic government" Sarkozy told the European Parliament in Strasbourg.

The European Central Bank, currently the only joint institution overseeing the 15-nation eurozone, "must be independent," but the Frankfurt-based monetary body "should be able to discuss with an economic government," Sarkozy added. See "Sarkozy Calls for 'Economic Government' for Eurozone," *The Economic Times*, October 21, 2008, accessed July 25, 2011, http://articles.economictimes.indiatimes.com/2008-10-21/news/28393734_1_eurozone-french-president-nicolas-sarkozy-economic-government.

President Sarkozy's concern was that there is no centralized entity in the EU that can play the same role as the Treasury in the United States. Member governments devise their own fiscal policies to deal with their own country's problems and do not take account of the effects of their actions on others in the European Union. This matters because the EU countries are so closely linked through trade and capital flows.

Governments within the EU did indeed act unilaterally to preserve their individual banking systems. The French government agreed to a 360 billion euro package of support for the French banking system and made a statement that no banks would collapse. Other countries took similar measures to restore confidence in their banking systems. Such measures sound similar to those taken in the United States, but there is an important difference. For the United States, such spending could be financed by taxes, government borrowing, or monetary expansion. But for, say, France, the equation is different. If the rescue package is not financed by increased taxes, then the French will have to issue more debt. They no longer control their money supply, so they cannot print currency to finance these bailouts.

Moreover, the Stability and Growth Pact, as we explained, places restrictions on the permissible magnitude of deficits by member governments. The reason for these restrictions is that, if many countries in the EMU were to run large deficits, there would be pressure on the ECB to finance some of this spending through additional money creation. In the aftermath of the crisis, many countries violated the fiscal restrictions, and how the monetary and fiscal authorities will ultimately respond to such pressure remains an open question.

One part of the response has been the establishment of additional facilities within Europe to pool resources to provide assistance to member states. Countries within Europe have been fulfilling a similar role to that played by the International Monetary Fund

(IMF). In particular, the crisis in Greece, and related debt problems in Ireland and Portugal, led to the creation in May 2010 of the European Financial Stability Facility (www.efsf.europa.eu/about/index.htm) to provide for the stabilization of countries undergoing financial and debt problems. A June 2011 press release discusses the provision of funds for Ireland and Portugal under this stabilization fund (www.efsf.europa.eu/mediacentre/news/2011/2011-006-eu-and-efsf-funding-plans-to-provide-financial-assistance-for-portugal-and-ireland.htm). The funds for Greece are coming from the EU member states directly.

Within the ECB, the discussion by President Trichet (http://www.ecb.int/press/key/date/2009/html/sp090427.en.html) summarized the perspective and policy choices of the central bank, including the provision of liquidity. Given that the ECB maintains an inflation target, how is this provision of liquidity consistent with that goal? One answer often given is that without this liquidity, the European economies might have fallen into deeper recessions and thus opened up the possibility of deflationary periods, as witnessed in the Great Depression years in the United States and in Japan during the 1990s.

Costs and Benefits of a Common Currency

Sarkozy's discussion of European economic government, and subsequent events in various countries, brought the debate over monetary integration back to the forefront in Europe. After the establishment of the European Monetary System, many European leaders thought the logical next step was a complete monetary union. This dream, embodied in the Maastricht Treaty, was finally realized in January 1999. A concise history of the steps to the Economic and Monetary Union is available at "Economic and Monetary Union (EMU)," European Central Bank, accessed July 20, 2011, http://www.ecb.int/ecb/history/emu/html/index.en.html. During the recent financial turmoil, however, the monetary ties that bind the European countries have been greatly strained. The costs of delegating monetary policy to a common central bank became very visible because individual countries were unable to respond to their own economic situations. In addition, the fiscal constraints included in the Maastricht Treaty hampered the ability of countries to conduct their desired fiscal policy. A recent report highlighted these concerns. See "Crisis Puts European Unity to the Test," *MoneyWeek*, October 10, 2008, accessed July 26, 2011, www.moneyweek.com/news-and-charts/economics/crisis-puts-european-unity-to-the-test-13811.aspx.

Milton Friedman always said that the European Union would not survive a deep recession. Well, that theory is certainly being put to the test now. As the financial crisis radiated across the globe this week, the EU fell into disarray as an ugly bout of tit-for-tat policies helped fuel a rout of European banks.

It began with Ireland's decision on 30 September to guarantee the deposits of its six main banks. This was a chance for European leaders to shore up banking confidence across Europe, says Leo McKinstry in the Daily Express. But instead of rallying behind the decision, German Chancellor Angela Merkel condemned it. Yet that didn't stop Greece from pledging to guarantee its own banks.

The recent crisis has forced a reevaluation of the costs and benefits of the common currency.

Most of the advantages of a common currency are self-apparent. As explained in Chapter 24, money acts as a medium of exchange, facilitating transactions among households and firms. A common currency obviates the need to exchange currencies when buying goods, services, and assets. Secondly, the monetary union removes the uncertainty associated with fluctuations in the exchange rate: within a monetary union, there are, of course, no exchange rate fluctuations at all. Further, unlike in a fixed exchange rate system, there is no need to buy and sell currencies to support the agreed-on exchange rates. Finally, because money also acts as a unit of account, a common currency makes it easier to compare prices across countries. All of these factors encourage countries to benefit from more efficient flows of goods and capital across borders.

There is another gain from a common currency that is more subtle. In some cases, individual countries are unable to do a good job of managing their own monetary policy. In Chapter 26, we explained that governments that run large deficits may decide to pay for these deficits by printing money. We also observed that there are situations where the monetary authority might be tempted to try unexpectedly expansionary policies when inflation is low. Such choices, while tempting, are ultimately damaging to an economy. Yet countries all too frequently indulge in such short-sighted policies. The underlying difficulty is a **commitment problem**. Ahead of time, the monetary authority might like to keep inflation low, but there is pressure to print money; in the end, countries experience high inflation caused by excessive money growth.

In the case of the EMU, this was not an especially pressing concern. The ECB was conducting conservative monetary policy. At the same time, the governments in the euro area ran reasonably sensible fiscal policies for the most part, so there was no pressure on the ECB to finance excessive spending. After the 2008 financial crisis, however, the deficit and debt pictures changed for many countries—particularly Greece, Ireland, and Portugal. The debt situation has now put enormous pressure on European institutions, including the ECB. So far, the ECB has remained on the sidelines by not being a direct contributor to bailout packages.

Commitment problems have arisen often in the monetary affairs of other countries. Argentina adopted a **currency board** in the 1990s because the monetary authorities could not commit to low-inflation policies in the late 1980s and early 1990s. To combat this problem, Argentina effectively adopted the US dollar as its currency. This monetary system meant that the Central Bank of Argentina was not able to increase the money supply independently: in effect, it delegated monetary policy to the United States. The monetary authority in Argentina was able to commit not to print pesos in response to fiscal pressures.

Some European countries, such as Denmark, elected not to make the euro its common currency. However, they adopted fixed exchange rates relative to the euro. Others, like the United Kingdom, did not make the euro its common currency and also elected to have floating exchange rates. Given all the advantages of a common currency, why did some countries reject the idea (and, for that matter, why is there not a single world currency)?

The answer is that there are also costs to adopting a common currency. As we have explained, the EMU entrusted monetary policy to a single central bank that decided monetary policy across a large number of countries. When these countries have different views about appropriate monetary policy, the delegation of monetary policy becomes problematic. Further, the fiscal restrictions imposed on the euro countries further reduced the ability of countries to respond to their own stabilization needs. In recent years, both Germany and France have violated the terms of the Stability and Growth Pact and the future of these fiscal restrictions remains in doubt.

Monetary Policy

Chapter 25 describes in detail the manner in which a central bank can use tools of monetary policy to influence aggregate economic activity and the price level. Monetary policy is a critical tool for stabilizing the macroeconomy. After the introduction of the euro, countries in the common currency area were no longer able to conduct independent monetary policy. The right to conduct monetary policy was ceded to the ECB.

Suppose that all the countries in the EMU were similar in their macroeconomic fortunes, meaning that the state of the macroeconomy in Italy was roughly the same as that in France, Ireland, Portugal, Belgium, and so forth. For example, suppose that when France experiences a period of recession, all the other countries in the union are in recession as well. In this case, the monetary policy that each country would have pursued if it had its own currency would most likely be very similar to the policy pursued by a central bank representing the interests of all the countries together. Each country, acting individually, would choose to cut interest rates to stimulate economic activity. The ECB would have an incentive to stimulate the economies of EMU member countries exactly as those members would have done with their own monetary policies. If countries are similar, in other words, the delegation of monetary policy to a central monetary authority is not that costly.

If countries are very different, it is more costly to move to a common currency. Suppose that Austria is undergoing a boom at the same time that Belgium is in recession. Belgium would like to cut interest rates. Austria would like to increase them. The ECB cannot satisfy both countries and may end up making them both unhappy. The crisis of 2008 did not have an even impact across all the countries in the euro area. Some countries saw major problems in their financial institutions, whereas others were less affected. As a result, different countries in the euro area had different desires in terms of the actions of the ECB.

Monetary policy operates through exchange rates as well as interest rates. By adopting a common currency, countries also give up the ability to stimulate their economies through depreciation or devaluation of their currency. Greece, Portugal, and Ireland have been forced to enact severe austerity measures to bring their debt under control. As a consequence, these countries have seen major recessions. If, say, Portugal was still using the escudo rather than the euro, it could have stimulated the economy by decreasing the value of the currency, thus encouraging net exports. It no longer has this option. It is possible for the real exchange rate to decrease even if the nominal rate is fixed, but this may require deflation in the domestic economy.

Fiscal Policy

The adoption of the single currency in Europe did not directly affect fiscal policy. In principle, it could have been adopted without any reference to fiscal policy. In practice, however, the single currency was accompanied by the fiscal limitations enshrined in the Stability and Growth Pact. In particular, the Stability and Growth Pact said that member countries were not allowed to run a government deficit that exceeded 3 percent of gross domestic product (GDP). The idea was that member countries were permitted to run deficits in periods of low economic activity but were encouraged to avoid large and sustained budget deficits.

As with the monetary agreement, there are costs and benefits to such fiscal restrictions. It is possible that a member country experiencing a period of low economic activity (a recession) would find itself unable to increase its government deficit, even if it

wanted to stimulate economic activity. Chapter 29 explained that there are sometimes gains to running deficits. One cost of the Stability and Growth Pact is that it reduces the ability of countries to use deficits for macroeconomic stabilization.

Fiscal restrictions are common within monetary unions. Within the United States, there are restrictions, largely imposed on the states by themselves, which limit budget deficits at the state level. The idea is that large deficits at the level of a European country or a US state might create an incentive for the central bank to print money and thus bail out the delinquent government. This would occur if the monetary authority lacks the ability to say "no" to a state or a country in financial distress. By limiting deficits in the first place, such bailouts need not occur. This is a gain for all the countries within the EMU.

The Crisis in the United Kingdom

So far we have looked at two large economies: the United States and the euro area. We now turn to the experience of some smaller economies, beginning with the United Kingdom. The United Kingdom is part of the EU but it is not in the euro area. It retained its own currency (the pound sterling) rather than adopting the euro. This meant, of course, that it also retained its own central bank. The Bank of England is known as a very independent monetary authority, and operates under very strict rules of inflation targeting. Yet it, too, responded to the crisis.

The United Kingdom was one of the first countries to face serious implications of the financial crisis when, in September 2007, there was a run on a lending institution called Northern Rock. The Bank of England evidently could have—but chose not to—take action early in the crisis to avoid the run on Northern Rock. Once the run commenced, however, the Bank of England injected liquidity into the system.

In October 2008, the Bank of England was, along with other central banks, cutting interest rates. However, the cuts it enacted were modest relative to the action taken in the United States and other countries. More significantly, the United Kingdom partially nationalized some of its banks over this period under a 400-billion-pound bailout plan. Just as in the US plan, the aim was to provide liquidity directly to these banks and thus open up the market for loans among banks. But, according to contemporary reports, UK banks were still not making new commitments weeks after this bailout plan was enacted.

The Crisis in Iceland

Iceland is a relatively small, very open economy. It has close links to the EU but retains its own currency: the krona. It was particularly hard hit by the financial crisis, in part because Icelandic banks had been borrowing extensively from abroad in the years prior to the financial crisis. According to one estimate, banks held foreign assets and liabilities worth about 10 times Iceland's entire GDP. This is partly based on a BBC article about Iceland: Jon Danielsson, "Why Raising Interest Rates Won't Work," BBC News, October 28, 2008, accessed July 26, 2011, http://news.bbc.co.uk/2/hi/business/7658908.stm.

The sheer size of asset holdings meant that if there was a substantial decrease in asset values, it was simply not possible for the Icelandic central bank or fiscal authorities to bail out domestic banks. Any attempt to bail out the banks would simply have bankrupted the government. You also might wonder why, as a last resort, Iceland could not generate print money to get itself out of trouble, financing a bailout through an **inflation tax**. We explained earlier that this would be a possibility in the United States, for example. The difference is that most of the liabilities of US financial institutions are denominated in US dollars, so inflation would reduce the real value of these liabilities. But much of the debt of Icelandic banks was not denominated in krona; it was denominated in euros, US dollars, or other currencies. Inflation in Iceland would simply lead to a depreciation of the currency and would not reduce the real value of the debt.

Based on estimates from the IMF, the financial and exchange rate problems of Iceland led to a contraction in real GDP of around 3 percent in 2009. In late October 2008, Iceland negotiated a $2.1 billion dollar loan from the IMF (http://www.imf.org/external/np/sec/pr/2008/pr08256.htm) for emergency funding to help stabilize its economy. To put this in perspective, Iceland's GDP is only $12 billion, and the loan was equivalent to almost $7,000 per person. Meanwhile, there was a precipitous decline in the value of the krona: between January and October 2008, the krona lost nearly half of its value.

Iceland's banking system was effectively nationalized in 2008. The government took over three of the biggest banks. During late October, the government tried to peg the krona at about 131 per euro. Their attempt failed, and the government was forced to allow the krona to decrease in response to market forces. There was a report of a trade at 340 krona per euro, far from the government's attempted peg.See Bo Nielsen, "Iceland's Krona Currency Trading Halts as Kaupthing Taken Over," October 9, 2008, accessed July 26, 2011, http://www.bloomberg.com/apps/news?pid=20601085&refer=europe&sid=aiz5QIq94nrw. One way to think about the decline in the value of the krona is through the government budget constraint. Once the government took over the banks, what had been a private liability became a government liability to depositors. One way to meet this obligation is through higher taxes;

another is through the creation of more currency. The rapid depreciation of the krona indicates that market participants were anticipating more inflation in Iceland, so the value of the currency decreased.

Iceland was merely the first country that ran into considerable distress as a result of the crisis of 2008. It was followed a few days later by Ukraine, which agreed to a $16.5 billion loan from the IMF. Countries such as Greece and Spain also faced problems as investors started to worry that their governments might default on their debt.

The Crisis in China

The financial crisis had an impact on China largely through trade linkages. China exports a lot of goods to western economies. As the level of economic activity in these economies slowed, the demand for goods and services produced in China decreased as well. This led to lower real GDP in China. As shown in the circular flow of income (Figure 30.4.1), the reduction in exports by China led to reduced output from Chinese firms, reduced income for Chinese households, and lower spending through the multiplier process.

Even though China owned many US assets, most were not directly linked to mortgage-backed securities. Instead, the Chinese were holding about $900 billion of US Treasury securities. Data on foreign holdings of US government securities is available at “Major Foreign Holders of Treasury Securities, ” US Department of the Treasury, September 16, 2011, accessed September 20, 2011, http://www.ustreas.gov/tic/mfh.txt. Although the value of these securities changed with the financial situation, this simply led to changes in the value of portfolios and did not lead to bankruptcy of financial institutions.

China differs from the United States and Europe because many of the banks operating in China are owned by the government. The top four state-owned banks had about 66 percent of China’s deposit market in 2007. So if the assets of those banks decrease in value, this loss is ultimately reflected in the budget of the government. Whereas the governments of England, the United States, and other countries attacked the crisis of 2008 by partial nationalization—that is, the purchase of bank shares by the government—this was unnecessary in China because the government already had a substantial ownership share in the banks.

Deposit insurance is also rather different in China. In the case of publicly owned banks, the government directly guarantees deposits so banks will not go bankrupt. There is no explicit deposit insurance for private banks, but the lack of explicit deposit insurance does not mean the Chinese government would not bail out a bank that was under attack. The Law of the People’s Republic of China on Commercial Banks Article 64 reads as follows:

When a commercial bank has suffered or will possibly suffer, credit crisis, thereby seriously affecting the interests of the depositors, the banking regulatory authority under the State Council may assume control over the bank.

The purposes of assumption of control are, through taking such measures as are necessary in respect of the commercial bank over which control is assumed, to protect the interests of the depositors and to enable the commercial bank to resume normal business. The debtor-creditor relationship with regard to a commercial bank over which control is assumed shall not change as a result of the assumption of control. “Article 64,” Law of the People’s Republic of China on Commercial Banks, May 10, 1995, accessed July 26, 2011, http://www.china.org.cn/english/DAT/214824.htm.

The Crisis in Argentina

What about the experience in Latin America during the crisis of 2008? Many countries, notably Argentina, Brazil, and Mexico, experienced their own financial and currency crises in recent decades. Those crises were “homegrown” because they were largely caused by domestic economic policies.

But the upheavals of recent years were not created in these countries. The linkages we explained earlier also caused these countries to be affected by the financial events that afflicted the United States and Europe. Figure 30.4.2 shows that the stock market in Argentina had similar volatility and losses to those experienced in other countries. This volatility, along with other financial upheavals, created an interesting response within Argentina: the government announced the nationalization of private pension plans.

What is the connection here? The government announced it was taking over private pensions to protect households who faced added financial risks. Instead of facing the risks of private asset markets, households were now shielded from that risk through a national pension system. Skeptics have argued that this was simply an opportunity for the government of Argentina to obtain some additional resources. Promises of future compensation for the lost pensions were not viewed as credible.

The Crisis in Australia

Finally, not every country in the world was badly hit by the crisis of 2008. Australia, for example, saw a significant stock market decrease but otherwise went through the crisis years with little more than a minor slowdown in economic growth. There are several reasons for this. Australia, like other countries, used both monetary and fiscal policy to stimulate the economy. On the fiscal side, it cut taxes and increased government purchases; on the monetary side, the Reserve Bank of Australia decreased interest rates (although not by as much as many other countries). Australia has historically kept its government debt very close to zero, so there were no concerns about default on Australian debt.

Australia made a very well-publicized cash transfer of about $1,000 to about half the population. Even though much of those transfers were probably saved rather than spent, they are credited with helping to support confidence and limit contagion effects in Australia. Finally, Australia has benefitted from a major resources boom, so demand for net exports was a robust component of aggregate expenditures during the crisis period.

Key Takeaways

1. The United States and the rest of the world are linked through many channels. Key channels that allowed the crisis to spread were financial links due to both holdings of assets across borders and the spread of pessimism across markets. In addition, links across countries due to trade flows meant that as income decreased in some countries, exports and thus real GDP decreased in others.
2. Within the EMU, individual countries were limited in fiscal policy responses due to restrictions on outstanding debt. Further, the ECB follows an inflation target rule and thus is not able to directly intervene to stabilize European economies. In the end, countries did take fiscal actions, and the ECB ultimately provided the needed liquidity to Europe. But this experience highlighted some of the costs of a monetary union.

Exercises

1. During the crisis of 2008, what happened to stock markets across the world?
2. To avoid spillovers from a financial crisis, what would a country have to do?
3. Why do other countries hold US government debt?

This page titled 30.4: The Crisis in Europe and the Rest of the World is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

- **Current page** by Anonymous is licensed CC BY-NC-SA 3.0. Original source: https://2012books.lardbucket.org/books/theory-and-applications-of-economics.
- **23.3: Job and Worker Flows** by Anonymous is licensed CC BY-NC-SA 3.0. Original source: https://2012books.lardbucket.org/books/theory-and-applications-of-economics.

30.5: Currency Crises 30.5: Currency CrisesEdit Update titlecancel

Exercises

After you have read this section, you should be able to answer the following questions:

1. What are the causes of a currency crisis?
2. How are currency crises and financial crises related?

In some countries, the financial crisis of 2008 led to a **currency crisis**. A currency crisis is a sudden and unexpected rapid decrease in the value of a currency. Currency crises are particularly severe in the case of a **fixed exchange rate** because such crises typically force a monetary authority to abandon the fixed rate. In the case of **flexible exchange rates**, a currency crisis occurs when the value of the currency decreases substantially in a short period of time. Such rapid depreciation is not as disruptive as the collapse of a fixed exchange rate, but it can still cause significant turmoil in an economy.

Exchange Rates in the Current Crisis

If you look at exchange rate data for September and October 2008, you can see that the dollar appreciated relative to the euro at that time. In other words, the dollar price of a euro decreased. Over the 10 days ending October 24, 2008, the dollar price of a euro decreased from about $1.35 to $1.25. More generally, several currencies experienced rapid depreciations during the financial crisis. Though there were no runs on these currencies, they nonetheless lost considerable value.

The dollar value of the British pound decreased to $1.62, its lowest level in 5 years, after the October 21, 2008, announcement that the UK economy was on the brink of a recession. There was a drop in value of about 7 percent over the previous week alone. The pound also decreased against the euro.

Currency Crises under Flexible Exchange Rates

A currency crisis can arise from a change in expectations, in ways that are similar to some of our earlier examples. Remember, for instance, how the current value of a house decreases when people expect that its future value will decrease. If you think that the value of Argentine pesos will decline (so each peso will be worth less in dollars), you may respond by selling pesos that you currently own. If everyone in the market shares your beliefs, then everyone will sell, and the value of the peso will decrease now.

Currency Crises under Fixed Exchange Rates

If everyone believes that a monetary authority can and will maintain the exchange rate, then people are happy to hold onto a currency. But if people believe that the fixed exchange rate is not sustainable, then there will be a run on the currency. Consider, for example, Brazil trying to stabilize its currency—the *real*. The monetary authority sets a fixed exchange rate, meaning that it stands ready to exchange Brazilian real for US dollars at a set price. If a fixed exchange rate is set too high, then the Brazilian central bank can maintain this value for a while by buying real with its own stocks of dollars.

But the central bank does not possess unlimited reserves of dollars. If the low demand for the real persists, then eventually the central bank will run out of reserves and thus no longer be able to support the currency. When that happens, the value of the real will have to decrease. A decrease in a fixed exchange rate is called **devaluation**.

In fact, the decrease in the value of the real would occur well before the central bank runs out of reserves. If you believe that the monetary authority will be forced to abandon the fixed rate, you will take your real and exchange them for dollars—and you will want to do this sooner rather than later to ensure you make the exchange before the real decreases in value. When lots of investors do this, the supply curve for real shifts outward. This makes the problem of maintaining the fixed exchange rate even more difficult for the central bank, so the devaluation of the currency will happen even sooner. If everyone does this, then the monetary authority will not have enough dollars on hand and will have to give up the fixed rate. The risk of such currency crises is the biggest potential problem with fixed exchange rates. History has given us many examples of such crises, and shows that they are very disruptive for the economy—and sometimes even for the world as a whole.

You may have noticed that a currency crisis looks a lot like a bank run. In both cases, pessimistic expectations of investors (about the future of a bank in one case and the future value of a currency in the other) lead them to all behave in a way that makes the pessimism self-fulfilling. In the case of a bank run, if all depositors are worried about their deposits and take their money out of the bank, then the bank fails and the depositors' pessimism was warranted. Likewise, if investors believe the devaluation of a currency

is likely, they will all want to sell their currency. This drives down the price and makes the devaluation much more likely to occur. A currency crisis, like a bank run, is an example of a coordination game.

Key Takeaways

1. A currency crisis can occur for several reasons, including being a consequence of a financial crisis or a fiscal crisis, or, in some cases, just driven by expectations like a bank run.
2. A financial crisis can lead to a currency crisis if depositors in one country, seeing the collapse of a financial system, rush to convert a home currency into foreign currencies.

Exercises

1. What is the difference between a fixed exchange rate system and a flexible exchange rate system?
2. What is the difference between a currency crisis and a devaluation?

This page titled 30.5: Currency Crises is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

30.6: End-of-Chapter Material 30.6: End-of-Chapter MaterialEdit Update titlecancel

In Conclusion

Five or six years ago, economists studied a period that they named “the Great Moderation.” In the period after World War II, and even more specifically from the mid 1980s to the mid 2000s, economic performance in the United States, Europe, and many countries had been relatively placid. These countries enjoyed respectable levels of long-run growth, experienced only mild recessions, and enjoyed low and stable inflation. Many observers felt that this performance was in large measure due to the fact that economists and policymakers had learned how to conduct effective monetary and fiscal policies. We learned from the mistakes of the Great Depression and knew how to prevent serious economic downturns. We also learned from the mistakes made in the 1970s and knew how to avoid inflationary policies.

To be sure, other countries still experienced their share of economic problems. Many countries in Latin America experienced currency crises and debt crises in the 1980s. Many countries in Southeast Asia suffered through painful exchange rate crises in the 1990s. Japan suffered a protracted period of low growth. Some countries saw hyperinflation, while others experienced economic decline. Still, for the most part, mature and developed economies experienced very good economic performance. Macroeconomics was becoming less about diagnosing failure and more about explaining success.

The last few years shook that worldview. The crisis of 2008 showed that a major economic catastrophe was not as unthinkable as economists and others hoped. The world experienced the most severe economic downturn since the Great Depression, and there was a period where it seemed possible that the crisis could even be on the same scale as the Great Depression. Countries like the United States and the United Kingdom faced protracted recessions. Countries such as Greece, Portugal, Ireland, and Iceland found themselves mired in debt crises. Spillovers and interconnections—real, financial, and psychological—meant that events like the bankruptcy of Lehman Brothers reverberated throughout the economies of the world.

Because it resurrected old problems, the crisis of 2008 also resurrected old areas of study in macroeconomics. The events in Europe have prompted economists to review the debate over common currencies and the conduct of monetary policy. There has been increased investigation of the size of fiscal policy multipliers. At the same time, macroeconomists are devoting much attention to topics such as the connection between financial markets and the real economy. But this difficult period for the world economy has also been an exciting time for macroeconomists. The study of macroeconomics has become more vital than ever—more alive and more essential.

Key Links

- International Monetary Fund forum on common currency: http://www.imf.org/external/np/exr/ecforums/110800.htm
- Economic and Monetary Union: http://ec.europa.eu/economy_finance/euro/emu/index_en.htm.
- European Central Bank: http://www.ecb.int/home/html/index.en.html
- CBO analysis of Obama package: www.cbo.gov/ftpdocs/106xx/doc10682/Frontmatter.2.2.shtml

exercises

1. Consider the bank run game. If a government is supposed to provide deposit insurance but depositors doubt the word of the government, might there still be a bank run?
2. Comparing the Great Depression to the financial crisis starting in 2008, what were the differences in the response of fiscal and monetary policy between these two episodes?
3. We explained in 30.2 Section "The Financial Crisis in the United States" that an increase in the expected future price of houses leads to an increase in the current price. Draw a supply-and-demand diagram to illustrate this idea.
4. Consider the crisis from the perspective of China. United States imports from China are roughly $300 billion each year. Due to the recession in the United States, imports from China decreased about 10 percent. If the marginal propensity to spend is 0.5 in China, what is the change in Chinese output predicted by the aggregate expenditure model? How much must government spending increase to offset this reduction in exports?
5. In the CBO assessment of ARRA, the multiplier from government purchases was assumed to be larger than the multiplier from tax cuts. How would you explain the differences in these multipliers?
6. If countries within the EMU are supposed to limit their deficits, what must happen to government spending during a recession when tax revenues decrease?

7. (Advanced) We argued that the provision of deposit insurance prevents bank runs. Is there an analogous policy to prevent currency crises?

Economics Detective

1. What has been the ECB's role in the European Financial Stability Facility and in the bailout packages for Greece, Ireland, and Portugal?
2. Find the details of the recent rescue package for Greece. What were the different views of Germany and France about this bailout? How was the IMF involved?
3. What predictions were made about job creation under the Obama administration's stimulus package? What happened to job creation rates in the 2008–10 period?

This page titled 30.6: End-of-Chapter Material is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

CHAPTER OVERVIEW

31: Toolkit

31.1: Introduction
31.2: Individual Demand
31.3: Elasticity
31.4: The Labor Market
31.5: Choices over Time
31.6: New Page
31.7: The Credit Market
31.8: Expected Value
31.9: Correcting for Inflation
31.10: Supply and Demand
31.11: Buyer Surplus and Seller Surplus
31.12: Efficiency and Deadweight Loss
31.13: Production Possibilities Frontier
31.14: Comparative Advantage
31.15: Pricing with Market Power
31.16: Pricing with Market Power
31.17: Comparative Statics
31.18: Production Function
31.19: Nash Equilibrium
31.20: Externalities and Public Goods
31.21: Foreign Exchange Market
31.22: Growth Rates
31.23: Mean and Variance
31.24: Correlation and Causality
31.25: The Credit (Loan) Market (Macro)
31.26: The Fisher Equation: Nominal and Real Interest Rates
31.27: The Aggregate Production Function
31.28: The Circular Flow of Income
31.29: Growth Accounting
31.30: The Solow Growth Model
31.31: The Aggregate Expenditure Model
31.32: Price Adjustment
31.33: Consumption and Saving
31.34: The Government Budget Constraint
31.35: The Life-Cycle Model of Consumption
31.36: Aggregate Supply and Aggregate Demand
31.37: The IS-LM Model

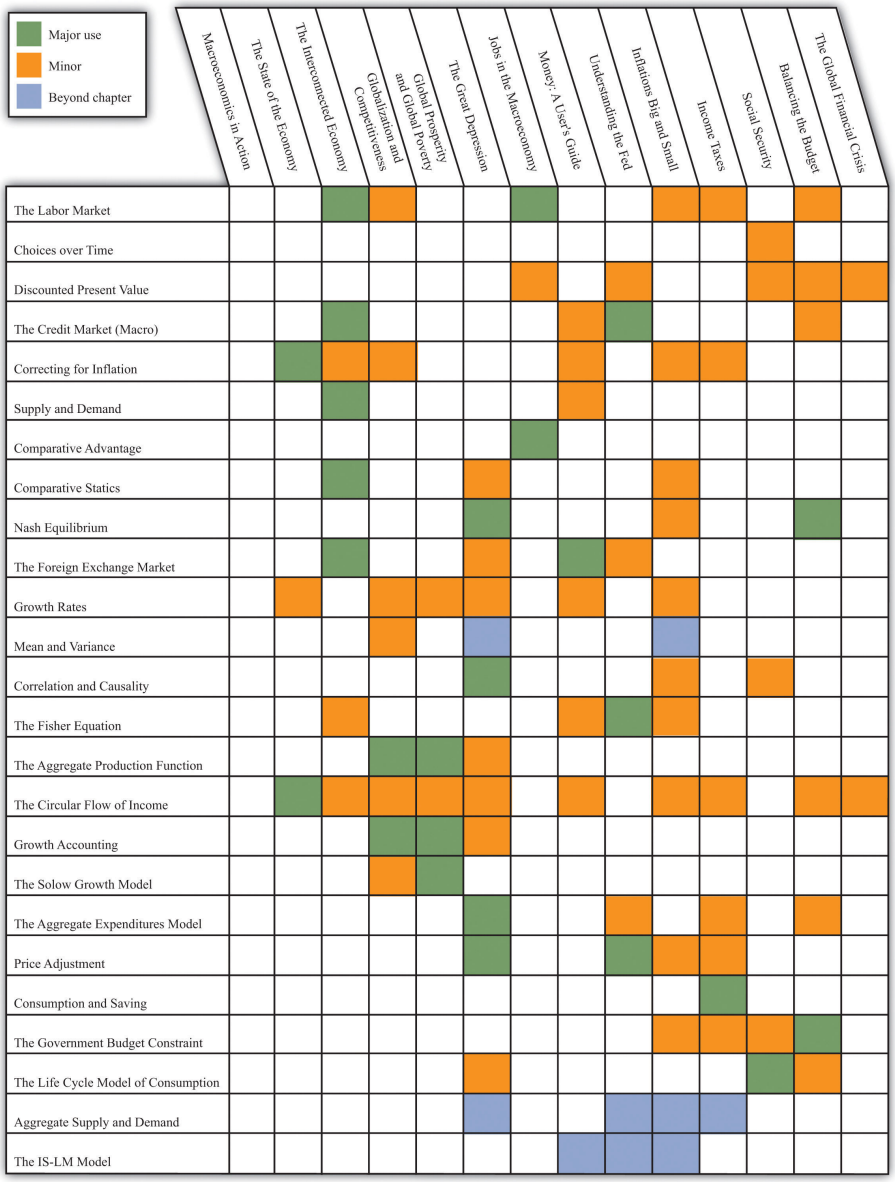
This page titled 31: Toolkit is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.1: Introduction 31.1: IntroductionEdit Update titlecancel

In this chapter, we present the key tools used in the macroeconomics and microeconomics part of this textbook. This toolkit serves two main functions:

1. Because these tools appear in multiple chapters, the toolkit serves as a reference. When using a tool in one chapter, you can refer back to the toolkit to find a more concise description of the tool as well as links to other parts of the book where the tool is used.
2. You can use the toolkit as a study guide. Once you have worked through the material in the chapters, you can review the tools using this toolkit.

The charts below show the main uses of each tool in green, and the secondary uses are in orange.



This page titled 31.1: Introduction is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.2: Individual Demand 31.2: Individual DemandEdit Update titlecancel

Individual demand refers to the demand for a good or a service by an individual (or a household). Individual demand comes from the interaction of an individual's desires with the quantities of goods and services that he or she is able to afford. By desires, we mean the likes and dislikes of an individual. We assume that the individual is able to compare two goods (or collections of goods) and say which is preferred. We assume two things: (1) an individual prefers more to less and (2) likes and dislikes are consistent.

An example is shown in part (a) of Figure 31.2.1. (This example is taken from Chapter 4.) In this example, there are two goods: music downloads ($1 each) and chocolate bars ($5 each). The individual has income of $100. The budget line is the combination of goods and services that this person can afford if he spends all of his income. In this example, it is the solid line connecting 100 downloads and 20 chocolate bars. The horizontal intercept is the number of chocolate bars the individual could buy if all income was spent on chocolate bars and is income divided by the price of a chocolate bar. The vertical intercept is the number of downloads the individual could buy if all income was spent on downloads and is income divided by the price of a download. The budget set is all the combinations of goods and services that the individual can afford, given the prices he faces and his available income. In the diagram, the budget set is the triangle defined by the budget line and the horizontal and vertical axes.

An individual's preferred point is the combination of downloads and chocolate bars that is the best among all of those that are affordable. Because an individual prefers more to less, all income will be spent. This means the preferred point lies on the budget line. The most that an individual would be willing to pay for a certain quantity of a good (say, five chocolate bars) is his valuation for that quantity. The marginal valuation is the most he would be willing to pay to obtain one *extra* unit of the good. The decision rule of the individual is to buy an amount of each good such that

marginal valuation = price.

The individual demand curve for chocolate bars is shown in part (b) of Figure 31.2.1. On the horizontal axis is the quantity of chocolate bars. On the vertical axis is the price. The demand curve is downward sloping: this is the law of demand. As the price of a chocolate bar increases, the individual substitutes away from chocolate bars to other goods. Thus the quantity demanded decreases as the price increases.

In some circumstances, the individual's choice is a zero-one decision: either purchase a single unit of the good or purchase nothing. The unit demand curve shows us the price at which a buyer is willing to buy the good. This price is the same as the buyer's valuation of the good. At any price above the buyer's valuation, the individual will not want to buy the good. At any price below the buyer's valuation, the individual wants to buy the good. If this price is exactly equal to the buyer's valuation, then the buyer is indifferent between purchasing the good or not.

In other words, the individual's decision rule is to purchase the good if the valuation of the good exceeds its price. This is consistent with the earlier condition because the marginal valuation of the first unit is the same as the valuation of that unit. The difference between the valuation and the price is the buyer surplus. (See 31.11 Section "Buyer Surplus and Seller Surplus" for more discussion.)

Key Insights

- The individual demand for a good or a service comes from the interactions of desires with the budget set.
- The individual purchases each good to the point where marginal valuation equals price.
- As the price of a good or a service increases, the quantity demanded will decrease.

More Formally

Let p_d be the price of a download, p_c the price of a chocolate bar, and I the income of an individual. Think of prices and income in terms of dollars. Then the budget set is the combinations of downloads (d) and chocolate bars (c) such that

$$I \geq p_d \times d + p_c \times c$$

The budget line is the combinations of d and c such that

$$I = p_d \times d + p_c \times c$$

In the graph, with downloads on the vertical axis, the equation for the budget line is

You can use this equation to understand how changes in income and prices change the position of the budget line. You can also use this equation to find the vertical and horizontal intercepts of the budget line, along with the slope of $-(p_c/p_d)$.

The individual purchases downloads up to the point where

$$MV_d = p_d$$

(where *MV* represents marginal valuation) and purchases chocolate bars up to the point where

$$MV_c = p_c.$$

Combining these expressions we get

which tells us that (minus) the ratio of marginal valuations equals the slope of the budget line. The ratio of marginal valuations is the rate at which an individual would like to trade one good for the other. The ratio of prices (the slope of the budget line) is the rate at which the market allows an individual to make these trades.

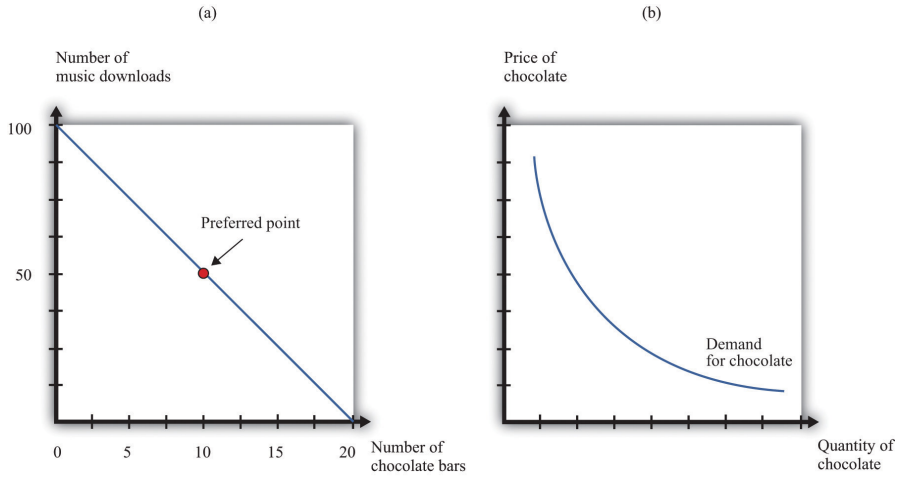


Figure 31.2.1: Individual Demand

The Main Uses of This Tool

- Chapter 4
- Chapter 5
- Chapter 6
- Chapter 8
- Chapter 10
- Chapter 13
- Chapter 14
- Chapter 15
- Chapter 17

This page titled 31.2: Individual Demand is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.3: Elasticity 31.3: ElasticityEdit Update titlecancel

Elasticity measures the proportionate change in one variable relative to the change in another variable. Consider, for example, the response of the quantity demanded to a change in the price. The price elasticity of demand is the percentage change in the quantity demanded divided by the percentage change in the price:

When the price increases (the percentage change in the price is positive), the quantity decreases, meaning that the percentage change in the quantity is negative. In other words, the law of demand tells us that the elasticity of demand is a negative number. For this reason we often use −(elasticity of demand) because we know this will always be a positive number.

- If $-(elasticity of demand) > 1$, demand is relatively elastic.
- If $-(elasticity of demand) < 1$, demand is relatively inelastic.

We can use the idea of the elasticity of demand whether we are thinking about the demand curve faced by a firm or the market demand curve. The definition is the same in either case.

If we are analyzing the demand curve faced by a firm, then we sometimes refer to the elasticity of demand as the own-price elasticity of demand. It tells us how much the quantity demanded changes when the firm changes its price. If we are analyzing a market demand curve, then the price elasticity of demand tells us how the quantity demanded in the market changes when the price changes. Similarly, the price elasticity of supply tells us how the quantity supplied in a market changes when the price changes. The price elasticity of supply is generally positive because the supply curve slopes upward.

The income elasticity of demand is the percentage change in the quantity demanded divided by the percentage change in income. The income elasticity of demand for a good can be positive or negative.

- If the income elasticity of demand is negative, it is an inferior good.
- If the income elasticity of demand is positive, it is a normal good.
- If the income elasticity of demand is greater than one, it is a luxury good.

The cross-price elasticity of demand tells us how the quantity demanded of one good changes when the price of another good changes.

- If the cross-price elasticity of demand is positive, the goods are substitutes.
- If the cross-price elasticity of demand is negative, the goods are complements.

In general, we can use elasticity whenever we want to show how one variable responds to changes in another variable.

Key Insights

- Elasticity measures the responsiveness of one variable to changes in another variable.
- Elasticities are *unitless*: you can measure the underlying variables in any units (for example, dollars or thousands of dollars), and the elasticity will not change.
- Elasticity is not the same as slope. For example, the price elasticity of demand depends on both the slope of the demand curve *and* the place on the demand curve where you are measuring the elasticity.

The Main Uses of This Tool

- Chapter 4
- Chapter 7
- Chapter 8
- Chapter 11
- Chapter 12
- Chapter 13
- Chapter 17

This page titled 31.3: Elasticity is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.4: The Labor Market 31.4: The Labor MarketEdit Update titlecancel

The **labor market** is the market in which labor services are traded. *Individual labor supply* comes from the choices of individuals or households about how to allocate their time. As the **real wage** (the nominal wage divided by the price level) increases, households supply more hours to the market, and more households decide to participate in the labor market. Thus the quantity of labor supplied increases. The labor supply curve of a household is shifted by changes in wealth. A wealthier household supplies less labor at a given real wage.

Labor demand comes from firms. As the real wage increases, the marginal cost of hiring more labor increases, so each firm demands fewer hours of labor input—that is, a firm's labor demand curve is downward sloping. The labor demand curve of a firm is shifted by changes in productivity. If labor becomes more productive, then the labor demand curve of a firm shifts rightward: the quantity of labor demanded is higher at a given real wage.

The labor market equilibrium is shown in Figure 31.4.2. The real wage and the equilibrium quantity of labor traded are determined by the intersection of labor supply and labor demand. At the equilibrium real wage, the quantity of labor supplied equals the quantity of labor demanded.

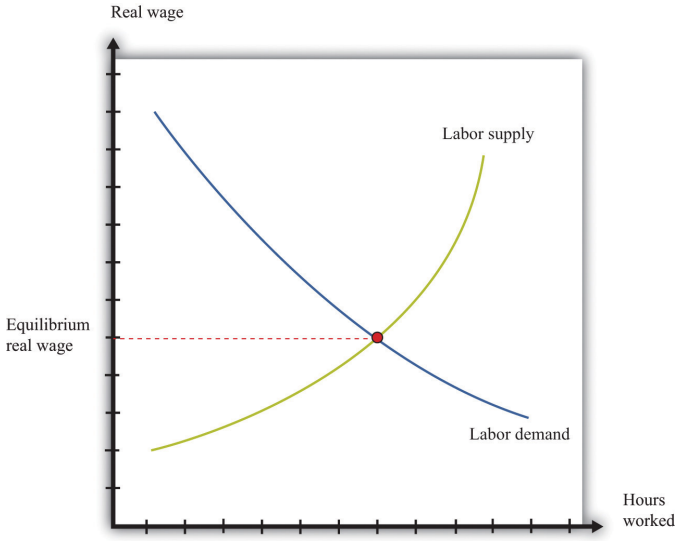


Figure 31.4.2: Labor Market Equilibrium

Key Insights

- Labor supply and labor demand depend on the real wage.
- Labor supply is upward sloping: as the real wage increases, households supply more hours to the market.
- Labor demand is downward sloping: as the real wage increases, firms demand fewer hours of work.
- A market equilibrium is a real wage and a quantity of hours such that the quantity demanded equals the quantity supplied.

The Main Uses of This Tool

- Chapter 19
- Chapter 20
- Chapter 23
- Chapter 26
- Chapter 27
- Chapter 29

This page titled 31.4: The Labor Market is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.5: Choices over Time 31.5: Choices over TimeEdit Update titlecancel

Individuals make decisions that unfold over time. Because individuals choose how to spend income earned over many periods on consumption goods over many periods, they sometimes wish to save or borrow rather than spend all their income in every period.

Figure 31.5.3 shows examples of these choices over a two-year horizon. The individual earns income this year and next. The combinations of consumption that are affordable and that exhaust all of an individual's income are shown on the budget line, which in this case is called an intertemporal budget constraint. The slope of the budget line is equal to $(1 + real\ interest\ rate)$, which is equivalent to the *real interest factor*. The slope is the amount of consumption that can be obtained tomorrow by giving up a unit of consumption today.

The preferred point is also indicated; it is the combination of consumption this year and consumption next year that the individual prefers to all the points on the budget line. The individual in part (a) of Figure 31.5.3 is consuming less this year than she is earning: she is saving. Next year she can use her savings to consume more than her income. The individual in part (b) of Figure 31.5.3 is consuming more this year than he is earning: he is borrowing. Next year, his consumption will be less than his income because he must repay the amount borrowed this year.

When the real interest rate increases, individuals will borrow less and (usually) save more (the effect of interest rate changes on saving is unclear as a matter of theory because **income effects** and **substitution effects** act in opposite directions). Thus **individual loan supply** slopes upward.

Of course, individuals live for many periods and make frequent decisions on consumption and saving. The **lifetime budget constraint** is obtained using the idea of **discounted present value**:

$$discounted\ present\ value\ of\ lifetime\ income = discounted\ present\ value\ of\ lifetime\ consumption.$$

The left side is a measure of all the **disposable income** the individual will receive over his lifetime (disposable means after taking into account taxes paid to the government and transfers received from the government). The right side calculates the value of consumption of all goods and services over an individual's lifetime.

Key Insights

- Over a lifetime, an individual's discounted present value of consumption will equal the discounted present value of income.
- Individuals can borrow or lend to obtain their preferred consumption bundle over their lifetimes.
- The price of borrowing is the real interest rate.

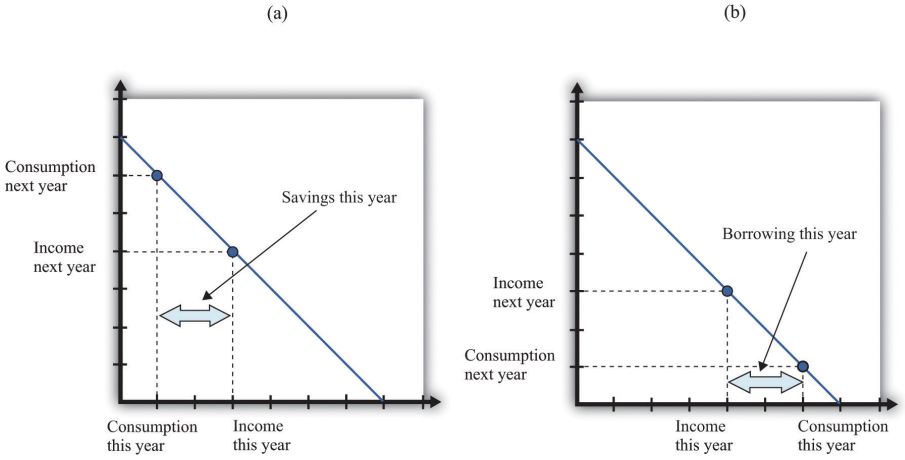


Figure 31.5.3: Choices over Time

The Main Use of This Tool

- Chapter 28

This page titled 31.5: Choices over Time is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.6: New Page 31.6: New PageEdit Update titlecancel

Discounted present value is a technique used to add dollar amounts over time. We need this technique because a dollar today has a different value from a dollar in the future.

The discounted present value this year of $1.00 that you will receive next year is as follows:

If the **nominal interest rate** is 10 percent, then the **nominal interest factor** is 1.1, so $1 next year is worth $1/1.1 = $0.91 this year. As the interest rate increases, the discounted present value decreases.

More generally, we can compute the value of an asset this year from the following formula:

The flow benefit depends on the asset. For a bond, the flow benefit is a coupon payment. For a stock, the flow benefit is a dividend payment. For a fruit tree, the flow benefit is the yield of a crop.

If an asset (such as a bond) yields a payment next year of $10 and has a price next year of $90, then the “flow benefit from asset + price of the asset next year” is $100. The value of the asset this year is then If the nominal interest rate is 20 percent, then the value of the asset is $100/1.2 = 83.33.

We discount nominal flows using a nominal interest factor. We discount real flows (that is, flows already corrected for inflation) using a **real interest factor**, which is equal to (1 + **real interest rate**).

Key Insights

- If the interest rate is positive, then the discounted present value is less than the direct sum of flows.
- If the interest rate increases, the discounted present value will decrease.

More Formally

Denote the dividend on an asset in period *t* as D_t. Define R_t as the *cumulative* effect of interest rates up to period *t*. For example, $R_2 = (1 + r_1)(1 + r_2)$. Then the value of an asset that yields D_t dollars in every year up to year *T* is given by

If the interest rate is constant (equal to *r*), then the one period interest factor is $R = 1 + r$, and $R_t = R^t$.

The discounted present value tool is illustrated in Table 31.2.1 "Discounted Present Value with Different Interest Rates". The number of years (*T*) is set equal to 5. The table gives the value of the dividends in each year and computes the discounted present values for two different interest rates. For this example, the annual interest rates are constant over time.

Table 31.6.3: Discounted Present Value with Different Interest Rates

Year	Dividend ($)	Discounted Present Value with $R = 1.05$ ($)	Discounted Present Value with $R = 1.10$ ($)
1	100	100	100
2	100	95.24	90.91
3	90	81.63	74.38
4	120	103.66	90.16
5	400	329.08	273.20
Discounted present value		709.61	628.65

The Main Uses of This Tool

- Chapter 23
- Chapter 25
- Chapter 28
- Chapter 29
- Chapter 30

This page titled 31.6: New Page is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.7: The Credit Market 31.7: The Credit MarketEdit Update titlecancel

An individual's choices over time determine how much he or she will borrow or lend. In particular, individual loan supply is upward sloping: when the real interest rate increases, a typical household will supply a greater quantity of funds to the credit market. Market loan supply is obtained by adding together the individual loan supplies of everyone in an economy. We use the terms "credit" and "loans" interchangably.

The demand for credit comes from households and firms that are borrowing. Market loan demand is obtained by adding together all the individual demands for loans. When real interest rates increase, borrowing is more expensive, so the quantity of loans demanded decreases. That is, loan demand obeys the law of demand.

Borrowers and lenders interact in the credit market (or loan market), which is illustrated in Figure 31.7.4. Credit market equilibrium occurs at the real interest rate where the quantity of loans supplied equals the quantity of loans demanded. At this equilibrium real interest rate, lenders lend as much as they wish, and borrowers can borrow as much as they wish. All gains from trade through loans are exhausted in equilibrium.

Key Insight

- As the real interest rate increases, more loans are supplied, and fewer loans are demanded.

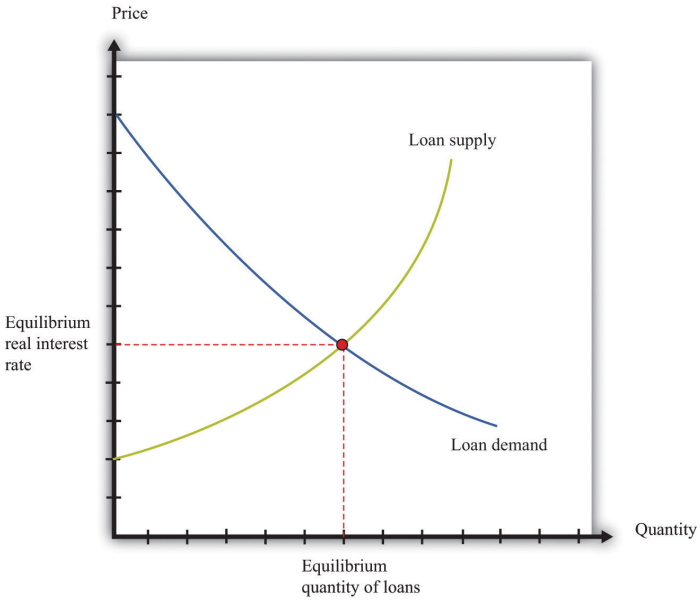


Figure 31.7.4: The Credit Market (or Loan Market)

The Main Uses of This Tool

- Chapter 5
- Chapter 8
- Chapter 10
- Chapter 13

This page titled 31.7: The Credit Market is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.8: Expected Value 31.8: Expected ValueEdit Update titlecancel

Probability is the percentage chance that something will occur. For example, there is a 50 percent chance that a tossed coin will come up heads. We say that the probability of getting the *outcome* "*heads*" is 1/2. There are five things you need to know about probability:

1. The list of possible outcomes must be complete.
2. The list of possible outcomes must not overlap.
3. If an outcome is certain to occur, it has probability 1.
4. If an outcome is certain not to occur, it has probability 0.
5. If we add together the probabilities for all the possible outcomes, the total must equal 1.

The expected value of a situation with financial risk is a measure of how much you would expect to win (or lose) on average if the situation were to be replayed a large number of times. You can calculate expected value as follows:

- For each outcome, multiply the probability of that outcome by the amount you will receive.
- Add together these amounts over all the possible outcomes.

For example, suppose you are offered the following proposal. Roll a six-sided die. If it comes up with 1 or 2, you get $90. If it comes up 3, 4, 5, or 6, you get $30. The expected value is

$$(1/3) \times \$90 + (2/3) \times \$30 = \$50.$$

Most people dislike risk. They prefer a fixed sum of money to a gamble that has the same expected value. *Risk aversion* is a measure of how much people want to avoid risk. In the example we just gave, most people would prefer a sure $50 to the uncertain proposal with the expected value of $50.

Suppose we present an individual with the following gamble:

- With 99 percent probability, you lose nothing.
- With 1 percent probability, you lose $1,000.

The expected value of this gamble is −$10. Now ask the individual how much she would pay to avoid this gamble. Someone who is risk-neutral would be willing to pay only $10. Someone who is risk-averse would be willing to pay more than $10. The more risk-averse an individual, the more the person would be willing to pay.

The fact that risk-averse people will pay to shed risk is the basis of insurance. If people have different attitudes toward risky gambles, then the less risk-averse individual can provide insurance to the more risk-averse individual. There are gains from trade. Insurance is also based on diversification, which is the idea that people can share their risks so it is much less likely that any individual will face a large loss.

Key Insights

- Expected value is the sum of the probability of an event times the gain/loss if that event occurs.
- Risk-averse people will pay to avoid risk. This is the basis of insurance.

More Formally

Consider a gamble where there are three and only three possible outcomes (*x*, *y*, *z*) that occur with probabilities *Pr(x)*, *Pr(y)*, and *Pr(z)*. Think of these outcomes as the number of dollars you get in each case. First, we know that

$$Pr(x) + Pr(y) + Pr(z) = 1.$$

Second, the expected value of this gamble is

$$EV = (Pr(x)^{*}x) + (Pr(y)^{*}y) + (Pr(z)^{*}z).$$

The Main Uses of This Tool

- Chapter 5
- Chapter 9
- Chapter 10
- Chapter 11

- Chapter 16

This page titled 31.8: Expected Value is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.9: Correcting for Inflation 31.9: Correcting for InflationEdit Update titlecancel

If you have some data expressed in nominal terms (for example, in dollars), and you want to convert them to real terms, you should use the following four steps.

1. Select your deflator. In most cases, the Consumer Price Index (CPI) is the best deflator to use. You can find data on the CPI (for the United States) at the Bureau of Labor Statistics website (http://www.bls.gov).
2. Select your base year. Find the value of the index in that base year.
3. For all years (including the base year), divide the value of the index in that year by the value in the base year. The value for the base year is 1.
4. For each year, divide the value in the nominal data series by the number you calculated in step 3. This gives you the value in "base year dollars."

Table 31.2.2 "Correcting Nominal Sales for Inflation" shows an example. We have data on the CPI for three years, as listed in the second column. The price index is created using the year 2000 as a base year, following steps 1–3. Sales measured in millions of dollars are given in the fourth column. To correct for inflation, we divide sales in each year by the value of the price index for that year. The results are shown in the fifth column. Because there was inflation each year (the price index is increasing over time), real sales do not increase as rapidly as nominal sales.

Table 31.9.2: Correcting Nominal Sales for Inflation

Year	CPI	Price Index (2000 Base)	Sales (Millions)	Real Sales (Millions of Year 2000 Dollars)
2000	172.2	1.0	21.0	21.0
2001	177.1	1.03	22.3	21.7
2002	179.9	1.04	22.9	21.9

Source: Bureau of Labor Statistics for the Consumer Price Index

This calculation uses the CPI, which is an example of a price index. To see how a price index like the CPI is constructed, consider Table 31.2.3 "Constructing a Price Index", which shows a very simple economy with three goods: T-shirts, music downloads, and meals. The prices and quantities purchased in the economy in 2012 and 2013 are summarized in the table.

Table 31.9.3: Constructing a Price Index

Year	T-shirts		Music Downloads		Meals		Cost of 2013 Basket	Price Index
	Price ($)	Quantity	Price ($)	Quantity	Price ($)	Quantity	Price ($)	
2012	20	10	1	50	25	6	425	1.00
2013	22	12	0.80	60	26	5	442	1.04

To construct a price index, you must choose a fixed basket of goods. For example, we could use the goods purchased in 2013 (12 T-shirts, 60 downloads, and 5 meals). This fixed basket is then priced in different years. To construct the cost of the 2013 basket at 2013 prices, the product of the price and the quantity purchased for each good in 2013 is added together. The basket costs $442. Then we calculate the cost of the 2013 basket at 2012 prices: that is, we use the prices of each good in 2012 and the quantities purchased in 2013. The sum is $425. The price index is constructed using 2012 as a base year. The value of the price index for 2013 is the cost of the basket in 2013 divided by its cost in the base year (2012).

When the price index is based on a bundle of goods that represents total output in an economy, it is called the **price level**. The CPI and *gross domestic product (GDP) deflator* are examples of measures of the price level (they differ in terms of exactly which goods are included in the bundle). The growth rate of the price level (its percentage change from one year to the next) is called the **inflation rate**.

We also correct interest rates for inflation. The interest rates you typically see quoted are in nominal terms: they tell you how many dollars you will have to repay for each dollar you borrow. This is called a **nominal interest rate**. The **real interest rate** tells you how much you will get next year, in terms of goods and services, if you give up a unit of goods and services this year. To correct interest rates for inflation, we use the **Fisher equation**:

$$real\ interest\ rate \approx nominal\ interest\ rate - inflation\ rate.$$

For more details, see 31.3 Section "The Fisher Equation: Nominal and Real Interest Rates" on the Fisher equation.

Key Insights

- Divide nominal values by the price index to create real values.
- Create the price index by calculating the cost of buying a fixed basket in different years.

The Main Uses of This Tool

- Chapter 18
- Chapter 19
- Chapter 20
- Chapter 24
- Chapter 26
- Chapter 27

This page titled 31.9: Correcting for Inflation is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.10: Supply and Demand 31.10: Supply and DemandEdit Update titlecancel

The supply-and-demand framework is the most fundamental framework in economics. It explains both the price of a good or a service and the quantity produced and purchased.

The **market supply curve** comes from adding together the **individual supply curves** of firms in a particular market. A competitive firm, taking prices as given, will produce at a level such that

$$price = marginal\ cost.$$

Marginal cost usually increases as a firm produces more output. Thus an increase in the price of a product creates an incentive for firms to produce more—that is, the supply curve of a firm is upward sloping. The market supply curve slopes upward as well: if the price increases, all firms in a market will produce more output, and some new firms may also enter the market.

A firm's supply curve shifts if there are changes in input prices or the state of technology. The market supply curve is shifted by changes in input prices and changes in technology that affect a significant number of the firms in a market.

The **market demand curve** comes from adding together the **individual demand curves** of all households in a particular market. Households, taking the prices of all goods and services as given, distribute their income in a manner that makes them as well off as possible. This means that they choose a combination of goods and services preferred to any other combination of goods and services they can afford. They choose each good or service such that

$$price = marginal\ valuation.$$

Marginal valuation usually decreases as a household consumes more of a product. If the price of a good or a service decreases, a household will substitute away from other goods and services and toward the product that has become cheaper—that is, the demand curve of a household is downward sloping. The market demand curve slopes downward as well: if the price decreases, all households will demand more.

The household demand curve shifts if there are changes in income, prices of other goods and services, or tastes. The market demand curve is shifted by changes in these factors that are common across a significant number of households.

A market **equilibrium** is a price and a quantity such that the quantity supplied equals the quantity demanded at the equilibrium price (Figure 31.10.5). Because market supply is upward sloping and market demand is downward sloping, there is a unique equilibrium price. We say we have a **competitive market** if the following are true:

- The product being sold is homogeneous.
- There are many households, each taking the price as given.
- There are many firms, each taking the price as given.

A competitive market is typically characterized by an absence of barriers to entry, so new firms can readily enter the market if it is profitable, and existing firms can easily leave the market if it is not profitable.

Key Insights

- Market supply is upward sloping: as the price increases, all firms will supply more.
- Market demand is downward sloping: as the price increases, all households will demand less.
- A market equilibrium is a price and a quantity such that the quantity demanded equals the quantity supplied.

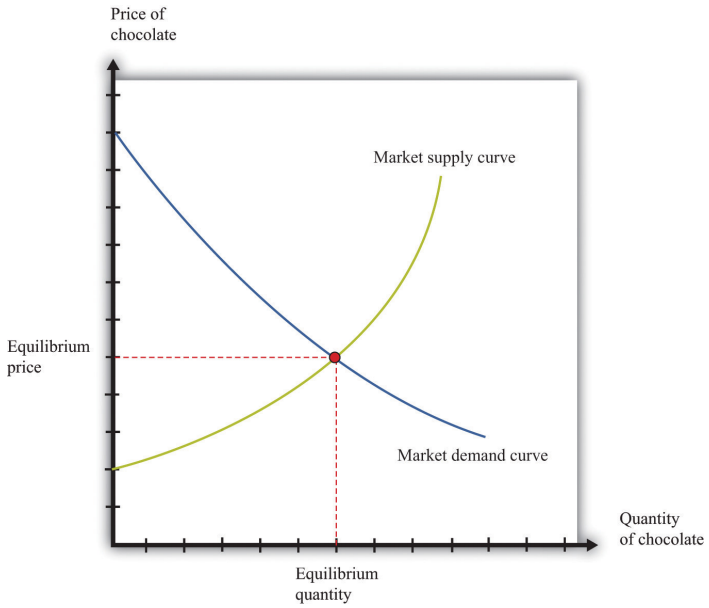


Figure 31.10.5: Market Equilibrium

Figure 31.10.5 shows equilibrium in the market for chocolate bars. The equilibrium price is determined at the intersection of the market supply and market demand curves.

More Formally

If we let *p* denote the price, *qd* the quantity demanded, and *I* the level of income, then the market demand curve is given by

$$qd = a - bp + cI,$$

where *a*, *b*, and *c* are constants. By the law of demand, $b > 0$. For a normal good, the quantity demanded increases with income: $c > 0$.

If we let *qs* denote the quantity supplied and *t* the level of technology, the market supply curve is given by

$$qs = d + ep + ft,$$

where *d*, *e*, and *f* are constants. Because the supply curve slopes upward, $e > 0$. Because the quantity supplied increases when technology improves, $f > 0$.

In equilibrium, the quantity supplied equals the quantity demanded. Set $qs = qd = q^*$ and set $p = p^*$ in both equations. The market clearing price (*p**) and quantity (*q**) are as follows:

and

$$q^* = d + ep^* + ft.$$

The Main Uses of This Tool

- Chapter 19
- Chapter 24

This page titled 31.10: Supply and Demand is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.11: Buyer Surplus and Seller Surplus 31.11: Buyer Surplus and Seller SurplusEdit Update titlecancel

If you buy a good, then you obtain buyer surplus. If you did not expect to obtain any surplus, then you would not choose to buy the good.

- Suppose you buy a single unit of the good. Your surplus is the difference between your valuation of the good and the price you pay. This is a measure of how much you gain from the exchange.
- If you purchase many units of a good, then your surplus is the sum of the surplus you get from each unit. To calculate the surplus from each unit, you subtract the price paid from your marginal valuation of that unit.

If you sell a good, then you obtain seller surplus. If you did not expect to obtain any surplus, you would not sell the good.

- Suppose you sell a single unit of a good. Your surplus is equal to the difference between the price you receive from selling the good and your valuation of the good. This valuation may be a measure of how much you enjoy the good or what you think you could sell it for in some other market.
- If you sell many units of a good, then the surplus you receive is the sum of the surplus for each unit you sell. To calculate the surplus from selling each unit, you take the difference between the price you get for each unit sold and your marginal valuation of that extra unit.

Buyer surplus and seller surplus are created by trade in a competitive market (Figure 31.11.6). The equilibrium price and the equilibrium quantity are determined by the intersection of the supply and demand curves. The area below the demand curve and above the price is the buyer surplus; the area above the supply curve and below the price is the seller surplus. The sum of the buyer surplus and the seller surplus is called total surplus or the gains from trade.

Buyer surplus and seller surplus can also arise from individual bargaining (Figure 31.11.7). When a single unit is traded (the case of unit demand and unit supply), the total surplus is the difference between the buyer's valuation and the seller's valuation. Bargaining determines how they share the gains from trade. The quantity of trades, indicated on the horizontal axis, is either zero or one. The valuations of the buyer and the seller are shown on the vertical axis. In this case, the valuation of the buyer ($3,000) exceeds the valuation of the seller ($2,000), indicating that there are gains from trade equal to $1,000. How these gains are shared between the buyer and seller depends on the price they agree on. In part (a) of Figure 31.11.7, the buyer gets most of the surplus; in part (b) of Figure 31.11.7, the seller gets most of the surplus.

Key Insights

- Buyer surplus and seller surplus are created by trade.
- Buyer surplus is the difference between the marginal value of a good and the price paid.
- Seller surplus is the difference between the price received and the marginal value of a good.

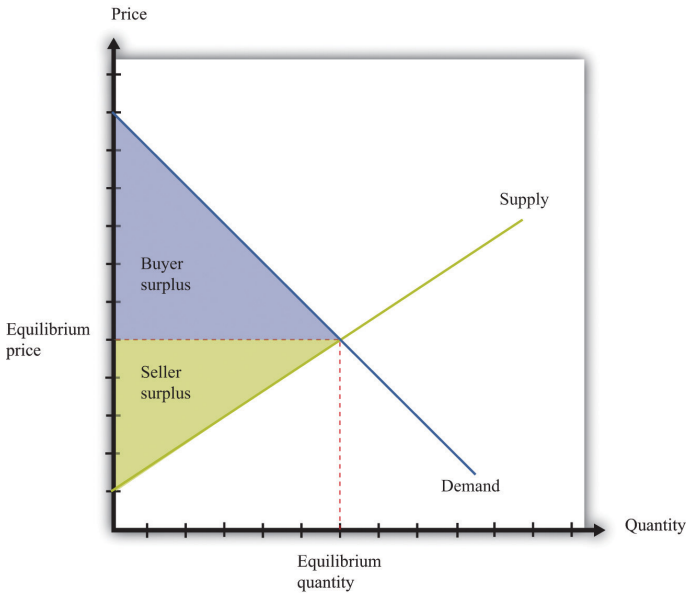


Figure 31.11.6: A Competitive Market

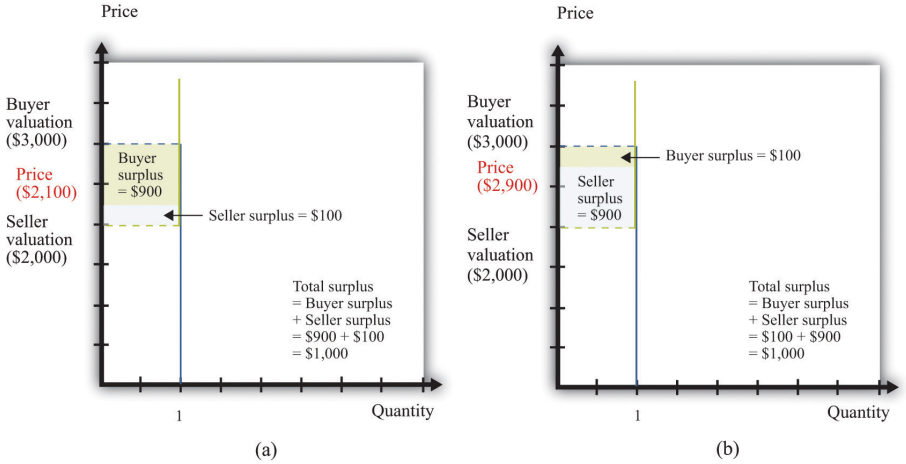


Figure 31.11.7: Individual Bargaining

The Main Uses of This Tool

- Chapter 6
- Chapter 8
- Chapter 11
- Chapter 12
- Chapter 15
- Chapter 17

This page titled 31.11: Buyer Surplus and Seller Surplus is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.12: Efficiency and Deadweight Loss 31.12: Efficiency and Deadweight LossEdit Update titlecancel

The outcome of a competitive market has a very important property. In equilibrium, all gains from trade are realized. This means that there is no additional surplus to obtain from further trades between buyers and sellers. In this situation, we say that the allocation of goods and services in the economy is efficient. However, markets sometimes fail to operate properly and not all gains from trade are exhausted. In this case, some buyer surplus, seller surplus, or both are lost. Economists call this a deadweight loss.

The deadweight loss from a monopoly is illustrated in Figure 31.12.8. The monopolist produces a quantity such that marginal revenue equals marginal cost. The price is determined by the demand curve at this quantity. A monopoly makes a profit equal to total revenue minus total cost. When the total output is less than socially optimal, there is a deadweight loss, which is indicated by the red area in Figure 31.12.8.

Deadweight loss arises in other situations, such as when there are quantity or price restrictions. It also arises when taxes or subsidies are imposed in a market. Tax incidence is the way in which the burden of a tax falls on buyers and sellers—that is, who suffers most of the deadweight loss. In general, the incidence of a tax depends on the elasticities of supply and demand.

A tax creates a difference between the price paid by the buyer and the price received by the seller (Figure 31.12.9). The burden of the tax and the deadweight loss are defined relative to the tax-free competitive equilibrium. The tax burden borne by the buyer is the difference between the price paid under the tax and the price paid in the competitive equilibrium. Similarly, the burden of the seller is the difference between the price in the competitive equilibrium and the price received under the equilibrium with taxes. The burden borne by the buyer is higher—all else being the same—if demand is less elastic. The burden borne by the seller is higher—all else being the same—if supply is less elastic.

The deadweight loss from the tax measures the sum of the buyer's lost surplus and the seller's lost surplus in the equilibrium with the tax. The total amount of the deadweight loss therefore also depends on the elasticities of demand and supply. The smaller these elasticities, the closer the equilibrium quantity traded with a tax will be to the equilibrium quantity traded without a tax, and the smaller is the deadweight loss.

Key Insights

- In a competitive market, all the gains from trade are realized.
- If sellers have market power, some gains from trade are lost because the quantity traded is below the competitive level.
- Other market distortions, such as taxes, subsidies, price floors, or price ceilings, similarly cause the amount to be traded to differ from the competitive level and cause deadweight loss.

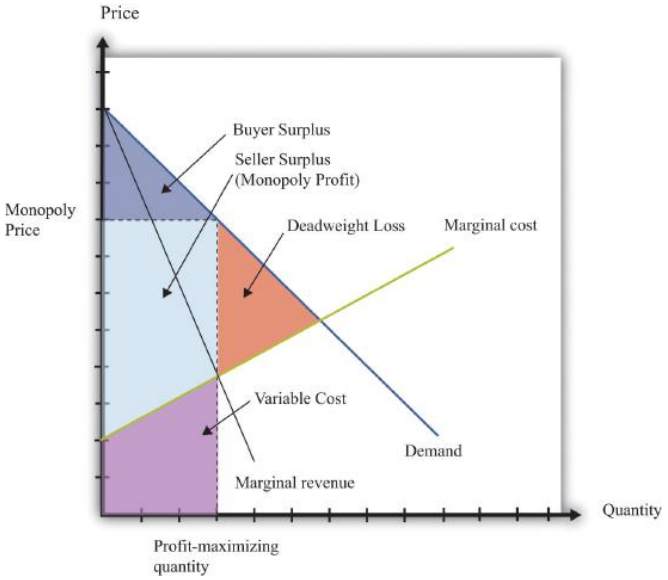


Figure 31.12.8: Deadweight Loss

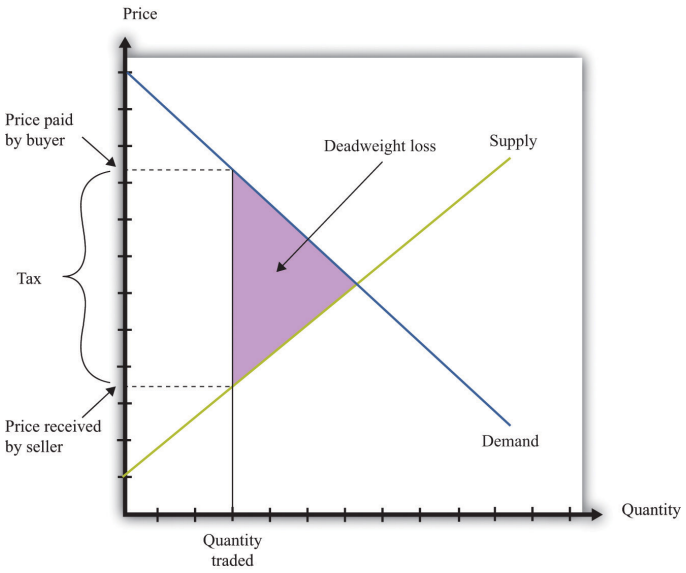


Figure 31.12.9: Tax Burdens

The Main Uses of This Tool

- Chapter 6
- Chapter 8
- Chapter 10
- Chapter 11
- Chapter 12
- Chapter 13
- Chapter 14
- Chapter 15
- Chapter 16
- Chapter 17

This page titled 31.12: Efficiency and Deadweight Loss is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.13: Production Possibilities Frontier 31.13: Production Possibilities FrontierEdit Update titlecancel

The production possibilities frontier shows the combinations of goods and services that an economy can produce if it is efficiently using every available input. A key component in understanding the production possibilities frontier is the term *efficiently*. If an economy is using its inputs in an efficient way, then it is not possible to produce more of one good without producing less of another.

Figure 31.13.10 shows the production possibilities frontier for an economy producing web pages and meals. It is downward sloping: to produce more web pages, the production of meals must decrease. Combinations of web pages and meals given by points inside the production possibilities frontier are possible for the economy to produce but are not efficient: at points inside the production possibilities frontier, it is possible for the economy to produce more of both goods. Points outside the production possibilities frontier are not feasible given the current levels of inputs in the economy and current technology.

The negative slope of the production possibilities frontier reflects opportunity cost. The opportunity cost of producing more meals is that fewer web pages can be created. Likewise, the opportunity cost of creating more web pages means that fewer meals can be produced.

The production possibilities frontier shifts over time. If an economy accumulates more physical capital or has a larger workforce, then it will be able to produce more of all the goods in an economy. Further, it will be able to produce new goods. Another factor shifting the production possibilities frontier outward over time is technology. As an economy creates new ideas (or receives them from other countries) on how to produce goods more cheaply, then it can produce more goods.

Key Insights

- The production possibilities frontier shows the combinations of goods and services that can be produced efficiently in an economy at a point in time.
- The production possibilities frontier is downward sloping: producing more of one good requires producing less of others. The production of a good has an opportunity cost.
- As time passes, the production possibilities frontier shifts outward due to the accumulation of inputs and technological progress.

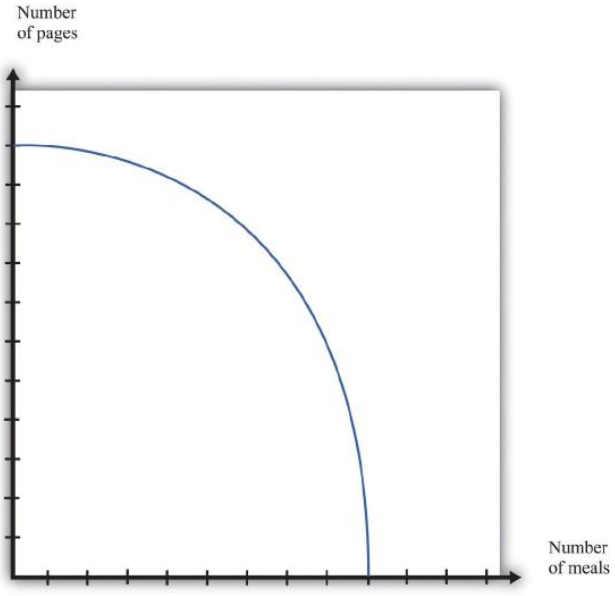


Figure 31.13.10: The Production Possibilities Frontier

The Main Uses of This Tool

- Chapter 6
- Chapter 12

This page titled 31.13: Production Possibilities Frontier is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.14: Comparative Advantage 31.14: Comparative AdvantageEdit Update titlecancel

Comparative advantage explains why individuals and countries trade with each other. Trade is at the heart of modern economies: individuals specialize in production and generalize in consumption. To consume many goods while producing relatively few, individuals must sell what they produce in exchange for the output of others. Countries likewise specialize in certain goods and services and import others. By so doing, they obtain **gains from trade**.

Table 31.2.4 "Hours of Labor Required" shows the productivity of two different countries in the production of two different goods. It shows the number of labor hours required to produce two goods—tomatoes and beer—in two countries: Guatemala and Mexico. From these data, Mexico has an **absolute advantage** in the production of both goods. Workers in Mexico are more productive at producing both tomatoes and beer in comparison to workers in Guatemala.

Table 31.14.4: Hours of Labor Required

	Tomatoes (1 Kilogram)	Beer (1 Liter)
Guatemala	6	3
Mexico	2	2

In Guatemala, the **opportunity cost** of 1 kilogram of tomatoes is 2 liters of beer. To produce an extra kilogram of tomatoes in Guatemala, 6 hours of labor time must be taken away from beer production; 6 hours of labor time is the equivalent of 2 liters of beer. In Mexico, the opportunity cost of 1 kilogram of tomatoes is 1 liter of beer. Thus the opportunity cost of producing tomatoes is lower in Mexico than in Guatemala. This means that Mexico has a comparative advantage in the production of tomatoes. By a similar logic, Guatemala has a comparative advantage in the production of beer.

Guatemala and Mexico can have higher levels of consumption of both beer and tomatoes if they trade rather than produce in isolation; each country should specialize (either partially or completely) in the good in which it has a comparative advantage. It is never efficient to have both countries produce both goods.

Key Insights

- Comparative advantage helps predict the patterns of trade between individuals and/or countries.
- A country has a comparative advantage in the production of a good if the opportunity cost of producing that good is lower in that country.
- Even if one country has an absolute advantage in all goods, it will still gain from trading with another country.
- Although this example is cast in terms of countries, the same logic is also used to explain production patterns between two individuals.

The Main Use of This Tool

- Chapter 23

31.14 Costs of Production

The costs of production for a firm are split into two categories. One type of cost, fixed costs, is independent of a firm's output level. A second type of cost, variable costs, depends on a firm's level of output. Total costs are the sum of the fixed costs and the variable costs.

The change in costs as output changes by a small amount is called marginal cost. It is calculated as follows:

Because fixed costs do not depend on the quantity, if we produce one more unit, then the change in total cost and the change in the variable cost are the same. Marginal cost is positive because variable costs increase with output. Marginal cost is usually increasing in the level of output, reflecting the diminishing marginal product of factors of production.

For example, suppose that total costs are given by

$$total\ costs = 50 + 10 \times quantity.$$

Here the fixed cost is 50, and the variable cost is 10 times the level of output. In this example, marginal cost equals 10. These costs are shown in Table 31.2.5.

Table 31.14.5

Output	Fixed Cost	Variable Cost	Total Cost
0	50	0	50
10	50	100	150
20	50	200	250
50	50	500	550

We sometimes divide fixed costs into two components: entry costs, which are the one-time fixed costs required to open a new business or set up a new plant, and fixed operating costs, which are the fixed costs incurred regularly during the normal operation of a business.

Some costs are sunk costs; once incurred, these costs cannot be recovered. Such costs should be ignored in forward-looking business decisions. Other costs are partially or fully recoverable costs. For example, if a firm purchases an asset that can be resold, then the cost of that asset is recoverable.

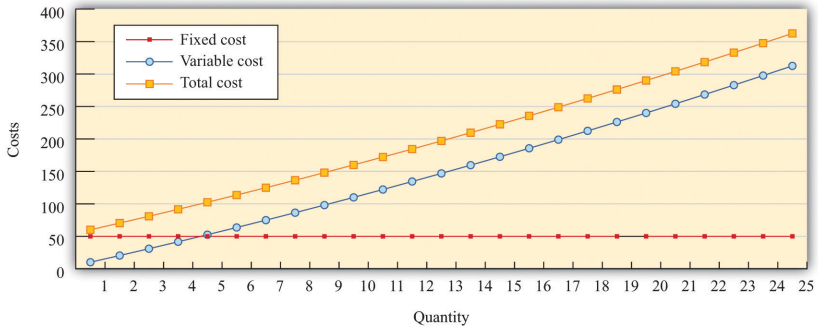


Figure 31.14.11: Cost Measures

Figure 31.14.11 shows these various measures of costs. It is drawn assuming a fixed cost of 50 and variable costs given by

$$variable\ costs = 10 \times quantity + 0.1 \times quantity^2.$$

For this example, marginal cost is positive and increasing.

Key Insights

- Fixed costs are independent of the level of output, whereas variable costs depend on the output level of a firm.
- Pricing decisions depend on marginal costs.
- Decisions to enter and/or exit an industry depend on both fixed and variable costs.

The Main Uses of This Tool

- Chapter 7
- Chapter 8
- Chapter 9
- Chapter 15

This page titled 31.14: Comparative Advantage is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.15: Pricing with Market Power 31.15: Pricing with Market PowerEdit Update titlecancel

The costs of production for a firm are split into two categories. One type of cost, fixed costs, is independent of a firm's output level. A second type of cost, variable costs, depends on a firm's level of output. Total costs are the sum of the fixed costs and the variable costs.

The change in costs as output changes by a small amount is called marginal cost. It is calculated as follows:

Because fixed costs do not depend on the quantity, if we produce one more unit, then the change in total cost and the change in the variable cost are the same. Marginal cost is positive because variable costs increase with output. Marginal cost is usually increasing in the level of output, reflecting the diminishing marginal product of factors of production.

For example, suppose that total costs are given by

$$total\ costs = 50 + 10 \times quantity.$$

Here the fixed cost is 50, and the variable cost is 10 times the level of output. In this example, marginal cost equals 10. These costs are shown in Table 31.2.5.

Table 31.15.5

Output	Fixed Cost	Variable Cost	Total Cost
0	50	0	50
10	50	100	150
20	50	200	250
50	50	500	550

We sometimes divide fixed costs into two components: entry costs, which are the one-time fixed costs required to open a new business or set up a new plant, and fixed operating costs, which are the fixed costs incurred regularly during the normal operation of a business.

Some costs are sunk costs; once incurred, these costs cannot be recovered. Such costs should be ignored in forward-looking business decisions. Other costs are partially or fully recoverable costs. For example, if a firm purchases an asset that can be resold, then the cost of that asset is recoverable.

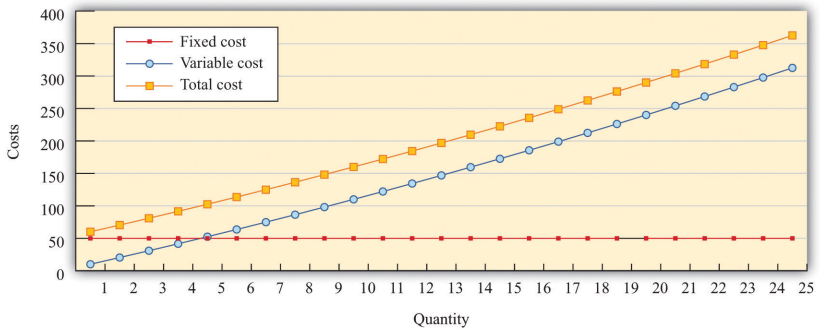


Figure 31.15.11: Cost Measures

Figure 31.15.11 shows these various measures of costs. It is drawn assuming a fixed cost of 50 and variable costs given by

$$variable\ costs = 10 \times quantity + 0.1 \times quantity^2.$$

For this example, marginal cost is positive and increasing.

Key Insights

- Fixed costs are independent of the level of output, whereas variable costs depend on the output level of a firm.
- Pricing decisions depend on marginal costs.

- Decisions to enter and/or exit an industry depend on both fixed and variable costs.

The Main Uses of This Tool

- Chapter 7
- Chapter 8
- Chapter 9
- Chapter 15

This page titled 31.15: Pricing with Market Power is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.16: Pricing with Market Power 31.16: Pricing with Market PowerEdit Update titlecancel

The goal of the managers of a firm is to maximize the firm's profit.

$$profit = revenues - costs.$$

We can think of a firm as choosing either the price to set or the quantity that it sells. Either way, the firm faces a demand curve and chooses a point on that curve that maximizes its profits. In reality, most firms choose the price of the good that they sell. However, it is often simpler to analyze a firm's behavior by looking at the quantity that it chooses.

Profits are maximized (Figure 31.16.12) when the extra revenue from selling one more unit of output (marginal revenue) is equal to the extra cost of producing one more unit (marginal cost). The firm's decision rule is to select a point on the demand curve such that

$$marginal\ revenue = marginal\ cost.$$

We can rearrange this condition to obtain a firm's pricing rule:

$$price = markup \times marginal\ cost.$$

Figure 31.16.12 illustrates this pricing decision. The markup depends on the price elasticity of demand. When demand is relatively inelastic, firms have a lot of market power and set a high markup. This is *not* a "plug-and-play" formula because both the markup and marginal cost depend, in general, on the price that a firm chooses. However, it does provide a useful description of a firm's decision.

Key Insights

- When marginal cost is higher, a firm sets a higher price.
- When demand is more inelastic (so a firm has more market power), the markup is higher, so a firm sets a higher price.
- When demand is perfectly elastic, the markup is 1, and the firm sets its price equal to marginal cost. This is the case of a competitive market.
- Any price you see has two components: the marginal cost and the markup. When a price changes, one or both of these must have changed.

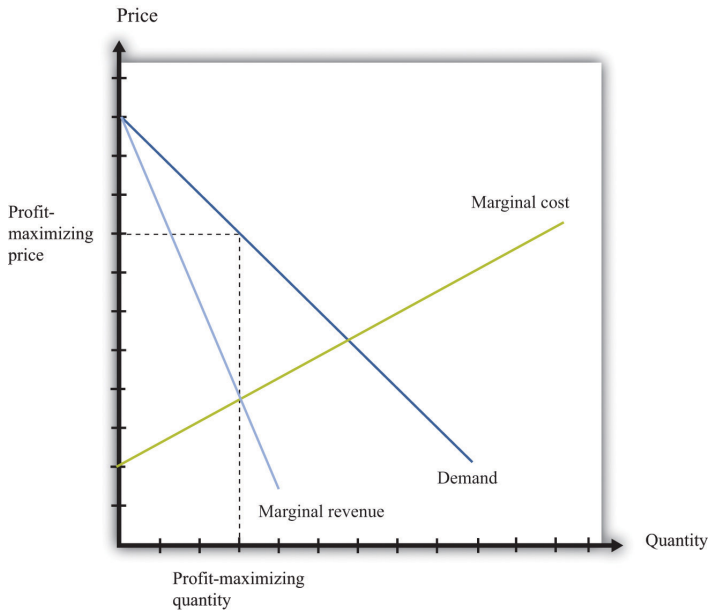


Figure 31.16.12: Markup Pricing

More Formally

We can derive the markup pricing formula as follows, where π = profit, R = revenues, C = costs, MR = marginal revenue, MC = marginal cost, P = price, Q = output, ε = $(\Delta Q/Q)/(\Delta P/P)$ = elasticity of demand, and μ = markup.

First we note that

The firm sets marginal revenue equal to marginal cost:

Rearranging, we obtain

$$P = \mu \times MC,$$

where the markup is given by

The Main Uses of This Tool

- Chapter 7
- Chapter 8
- Chapter 9
- Chapter 15
- Chapter 17

This page titled 31.16: Pricing with Market Power is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.17: Comparative Statics 31.17: Comparative StaticsEdit Update titlecancel

Comparative statics is a tool used to predict the effects of **exogenous** variables on market outcomes. Exogenous variables shift either the **market demand curve** (for example, news about the health effects of consuming a product) or the **market supply curve** (for example, weather effects on a crop). By market outcomes, we mean the **equilibrium** price and the equilibrium quantity in a market. Comparative statics is a comparison of the market equilibrium before and after a change in an exogenous variable.

A comparative statics exercise consists of a sequence of five steps:

1. Begin at an equilibrium point where the quantity supplied equals the quantity demanded.
2. Based on a description of the event, determine whether the change in the exogenous variable shifts the market supply curve or the market demand curve.
3. Determine the direction of this shift.
4. After shifting the curve, find the new equilibrium point.
5. Compare the new and old equilibrium points to predict how the exogenous event affects the market.

Figure 31.17.13 and Figure 31.17.14 show comparative statics in action in the market for Curtis Granderson replica shirts and the market for beer. In Figure 31.17.13, the market demand curve has shifted leftward. The consequence is that the equilibrium price and the equilibrium quantity both decrease. The demand curve shifts along a fixed supply curve. In Figure 31.17.14, the market supply curve has shifted leftward. The consequence is that the equilibrium price increases and the equilibrium quantity decreases. The supply curve shifts along a fixed demand curve.

Key Insights

- Comparative statics is used to determine the market outcome when the market supply and demand curves are shifting.
- Comparative statics is a comparison of equilibrium points.
- If the market demand curve shifts, then the new and old equilibrium points lie on a fixed market supply curve.
- If the market supply curve shifts, then the new and old equilibrium points lie on a fixed market demand curve.

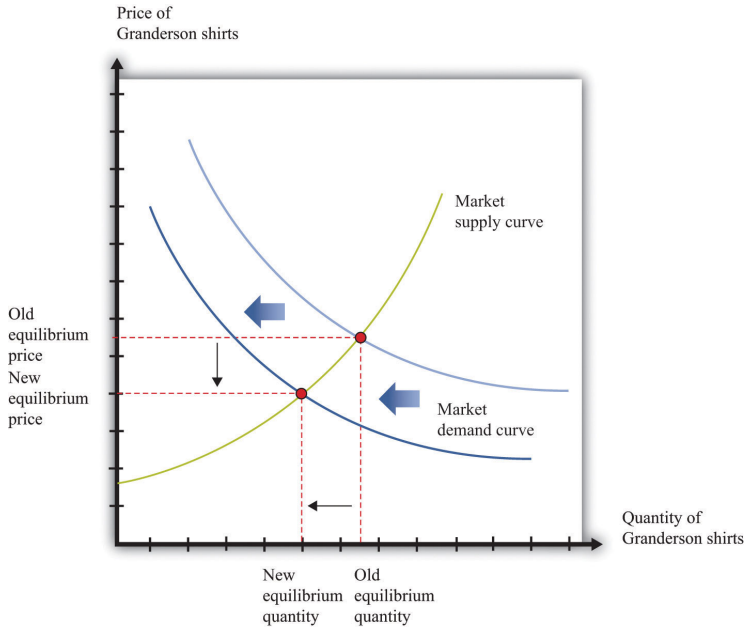


Figure 31.17.13: A Shift in the Demand Curve

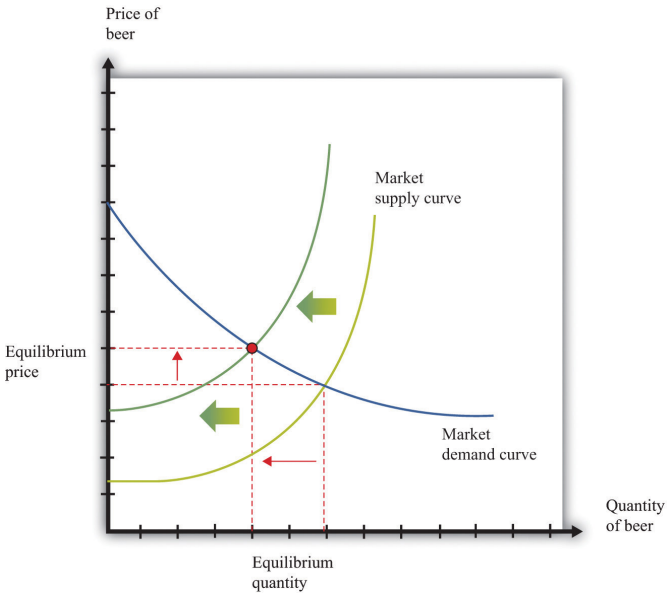


Figure 31.17.14: A Shift in the Supply Curve

The Main Uses of This Tool

- Chapter 19
- Chapter 22
- Chapter 26

This page titled 31.17: Comparative Statics is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.18: Production Function 31.18: Production FunctionEdit Update titlecancel

The production function characterizes the output of a firm given the inputs it uses. The link between inputs and output is shown Figure 31.18.15. The production function combines a firm's physical capital stock, labor, raw materials (or intermediate inputs), and technology to produce output. Technology is the knowledge (the "blueprints") that the firm possesses, together with managerial skills.

Production functions generally have two important properties:

1. Positive marginal product of an input
2. Diminishing marginal product of an input

By input, we mean any of the factors of production, such as physical capital, labor, or raw materials. The marginal product of an input is the extra output obtained if extra input is used. In this conceptual exercise, all other inputs are held fixed so that we change only one input at a time.

The first property asserts that additional output will be obtained from additional units of an input. Adding another machine, another worker, some more fuel, and so on, increases the output of a firm. A positive marginal product does not necessarily mean that the extra output is profitable: it might be that the cost of the extra input is high relative to the value of the additional output obtained.

The second property explains how the marginal product of an input changes as we increase the amount of that input, keeping the quantities of other inputs fixed. An additional unit of an input will (usually) increase output more when there is a small (rather than a large) amount of that input being used. For example, the extra output obtained from adding the first machine is greater than the additional output obtained from adding the 50th machine.

A simple production function relating output to labor input is shown in Figure 31.18.16. This figure illustrates the two properties of positive and diminishing marginal product of labor. As more labor is added, output increases: there is a positive marginal product of labor (that is, the slope of the relationship is positive). But the extra output obtained from adding labor is greater when the labor input is low: there is diminishing marginal product of labor. From the graph, the slope of the production function (which is the marginal product of labor) is greater at low levels of the labor input.

Key Insights

- The production function shows the output produced by a firm given its inputs.
- The production function displays two important properties: positive marginal product and diminishing marginal product.

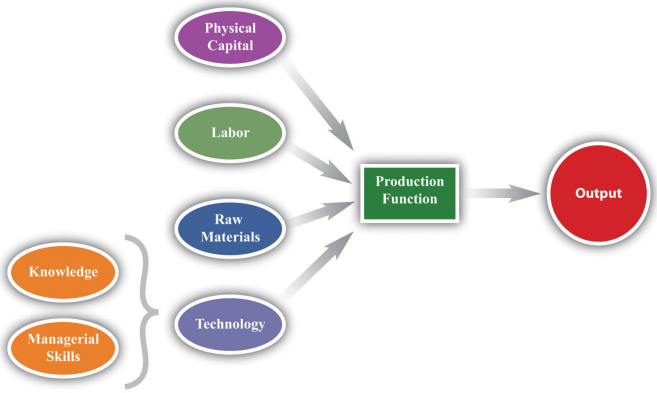


Figure 31.18.15: The Production Function

https://socialsci.libretexts.org/@go/page/47681

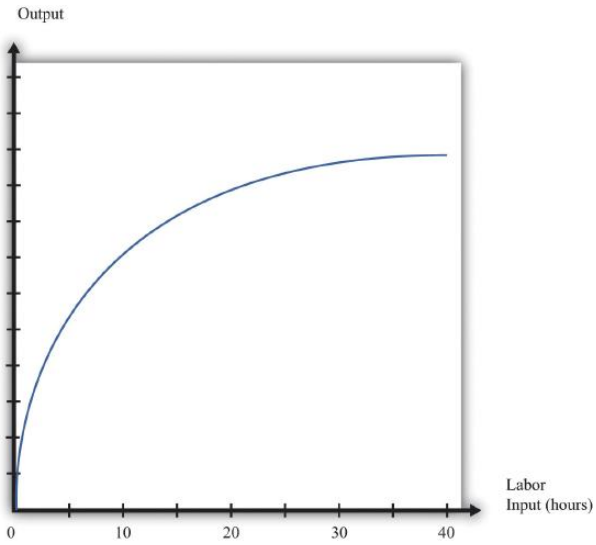


Figure 31.18.16: Labor Input in the Production Function

The Main Uses of This Tool

- Chapter 6
- Chapter 9
- Chapter 16

This page titled 31.18: Production Function is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.19: Nash Equilibrium 31.19: Nash EquilibriumEdit Update titlecancel

A **Nash equilibrium** is used to predict the outcome of a game. By a game, we mean the interaction of a few individuals, called *players*. Each player chooses an *action* and receives a *payoff* that depends on the actions chosen by everyone in the game.

A Nash equilibrium is an action for each player that satisfies two conditions:

1. The action yields the highest payoff for that player given her predictions about the other players' actions.
2. The player's predictions of others' actions are correct.

Thus a Nash equilibrium has two dimensions. Players make decisions that are in their own self-interests, and players make accurate predictions about the actions of others.

Consider the games in Table 31.2.6 "Prisoners' Dilemma", Table 31.2.7 "Dictator Game", Table 31.2.8 "Ultimatum Game", and Table 31.2.9 "Coordination Game". The numbers in the tables give the payoff to each player from the actions that can be taken, with the payoff of the row player listed first.

	Left	Right
Up	5, 5	0, 10
Down	10, 0	2, 2

Table 31.19.6: Prisoners' Dilemma

Number of dollars (x)	$100 - x$, x

Table 31.19.7: Dictator Game

	Accept	Reject
Number of dollars (x)	$100 - x$, x	0, 0

Table 31.19.8: Ultimatum Game

Table 31.19.9: Coordination Game

	Left	Right
Up	5, 5	0, 1
Down	1, 0	4, 4

- **Prisoners' dilemma.** The row player chooses between the action labeled *Up* and the one labeled *Down*. The column player chooses between the action labeled *Left* and the one labeled *Right*. For example, if row chooses *Up* and column chooses *Right*, then the row player has a payoff of 0, and the column player has a payoff of 10. If the row player predicts that the column player will choose *Left*, then the row player should choose *Down* (that is, down for the row player is her *best response* to left by the column player). From the column player's perspective, if he predicts that the row player will choose *Up*, then the column player should choose *Right*. The Nash equilibrium occurs when the row player chooses *Down* and the column player chooses *Right*. Our two conditions for a Nash equilibrium of making optimal choices and predictions being right both hold.
- **Social dilemma.** This is a version of the prisoners' dilemma in which there are a large number of players, all of whom face the same payoffs.
- **Dictator game.** The row player is called the dictator. She is given $100 and is asked to choose how many dollars (x) to give to the column player. Then the game ends. Because the column player does not move in this game, the dictator game is simple to analyze: if the dictator is interested in maximizing her payoff, she should offer nothing ($x = 0$).
- **Ultimatum game.** This is like the dictator game except there is a second stage. In the first stage, the row player is given $100 and told to choose how much to give to the column player. In the second stage, the column player accepts or rejects the offer. If the column player rejects the offer, neither player receives any money. The best choice of the row player is then to offer a penny (the smallest amount of money there is). The best choice of the column player is to accept. This is the Nash equilibrium.

- **Coordination game.** The coordination game has two Nash equilibria. If the column player plays *Left*, then the row player plays *Up*; if the row player plays *Up*, then the column player plays *Left*. This is an equilibrium. But *Down/Right* is also a Nash equilibrium. Both players prefer *Up/Left*, but it is possible to get stuck in a bad equilibrium.

Key Insights

- A Nash equilibrium is used to predict the outcome of games.
- In real life, payoffs may be more complicated than these games suggest. Players may be motivated by fairness or spite.

More Formally

We describe a game with three players (1, 2, 3), but the idea generalizes straightforwardly to situations with any number of players. Each player chooses a strategy (s_1, s_2, s_3). Suppose $\sigma_1(s_1, s_2, s_3)$ is the payoff to player 1 if (s_1, s_2, s_3) is the list of strategies chosen by the players (and similarly for players 2 and 3). We put an asterisk (*) to denote the best strategy chosen by a player. Then a list of strategies (s^*_1, s^*_2, s^*_3) is a Nash equilibrium if the following statements are true:

$$\sigma_1\left(s_1^*, s_2^*, s_3^*\right) \geq \sigma_1\left(s_1, s_2^*, s_3^*\right) \sigma_2\left(s_1^*, s_2^*, s_3^*\right) \geq \sigma_2\left(s_1^*, s_2, s_3^*\right) \sigma_3\left(s_1^*, s_2^*, s_3^*\right) \geq \sigma_3\left(s_1^*, s_2^*, s_3\right)$$

In words, the first condition says that, given that players 2 and 3 are choosing their best strategies (s^*_2, s^*_3), then player 1 can do no better than to choose strategy s^*_1. If a similar condition holds for every player, then we have a Nash equilibrium.

The Main Uses of This Tool

- Chapter 22
- Chapter 26
- Chapter 30

This page titled 31.19: Nash Equilibrium is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.20: Externalities and Public Goods 31.20: Externalities and Public GoodsEdit Update titlecancel

Some economic transactions have effects on individuals not directly involved in that transaction. When this happens, we say there is an externality present. An externality is generated by a decision maker who disregards the effects of his actions on others. In the case of a positive externality, the individual's actions increase the welfare of others (for example, research and development by firms). In the case of a negative externality, an individual's actions decrease the welfare of others (for example, pollution).

Economic outcomes are not efficient when externalities are present. So the government may be able to improve on the private outcome. The possible remedies are as follows:

- Subsidies (in the case of positive externalities) and taxes (in the case of negative externalities)
- The creation of markets by the government

If people are altruistic, then they may instead take into account others' welfare and may internalize some of the effects of their actions.

We typically see externalities associated with nonexcludable goods (or resources)—goods for which it is impossible to selectively deny access. In other words, it is not possible to let some people consume the good while preventing others from consuming it. An excludable good (or resource) is one to which we can selectively allow or deny access. If a good is nonexcludable or partially excludable, there are positive externalities associated with its production and negative externalities associated with its consumption.

We say that a good is a rival if one person's consumption of the good prevents others from consuming the good. Most of the goods we deal with in economics are rival goods. A good is nonrival if one person can consume the good without preventing others from consuming the same good. Knowledge is a nonrival good.

If a good is both nonexcludable and nonrival, it is a public good.

Key Insights

- When externalities are present, the outcome is inefficient.
- The market will typically not provide public goods.

The Main Uses of This Tool

- Chapter 12
- Chapter 14
- Chapter 16
- Chapter 17

This page titled 31.20: Externalities and Public Goods is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.21: Foreign Exchange Market 31.21: Foreign Exchange MarketEdit Update titlecancel

A **foreign exchange market** is where one currency is traded for another. There is a demand for each currency and a supply of each currency. In these markets, one currency is bought using another. The price of one currency in terms of another (for example, how many dollars it costs to buy one Mexican peso) is called the **exchange rate**.

Foreign currencies are demanded by domestic households, firms, and governments that wish to purchase goods, services, or financial assets denominated in the currency of another economy. For example, if a US auto importer wants to buy a German car, the importer must buy euros. The law of demand holds: as the price of a foreign currency increases, the quantity of that currency demanded will decrease.

Foreign currencies are supplied by foreign households, firms, and governments that wish to purchase goods, services, or financial assets denominated in the domestic currency. For example, if a Canadian bank wants to buy a US government bond, the bank must sell Canadian dollars. As the price of a foreign currency increases, the quantity supplied of that currency increases.

Exchange rates are determined just like other prices—by the interaction of supply and demand. At the equilibrium exchange rate, the supply and demand for a currency are equal. Shifts in the supply or the demand for a currency lead to changes in the exchange rate. Because one currency is exchanged for another in a foreign exchange market, the demand for one currency entails the supply of another. Thus the dollar market for euros (where the price is dollars per euro and the quantity is euros) is the mirror image of the euro market for dollars (where the price is euros per dollar and the quantity is dollars).

To be concrete, consider the demand for and the supply of euros. The supply of euros comes from the following:

- European households and firms that wish to buy goods and services from countries that do not have the euro as their currency
- European investors who wish to buy assets (government debt, stocks, bonds, etc.) that are denominated in currencies other than the euro

The demand for euros comes from the following:

- Households and firms in noneuro countries that wish to buy goods and services from Europe
- Investors in noneuro countries that wish to buy assets (government debt, stocks, bonds, etc.) that are denominated in euros

Figure 31.21.17 shows the dollar market for euros. On the horizontal axis is the quantity of euros traded. On the vertical axis is the price in terms of dollars. The intersection of the supply and demand curves determines the equilibrium exchange rate.

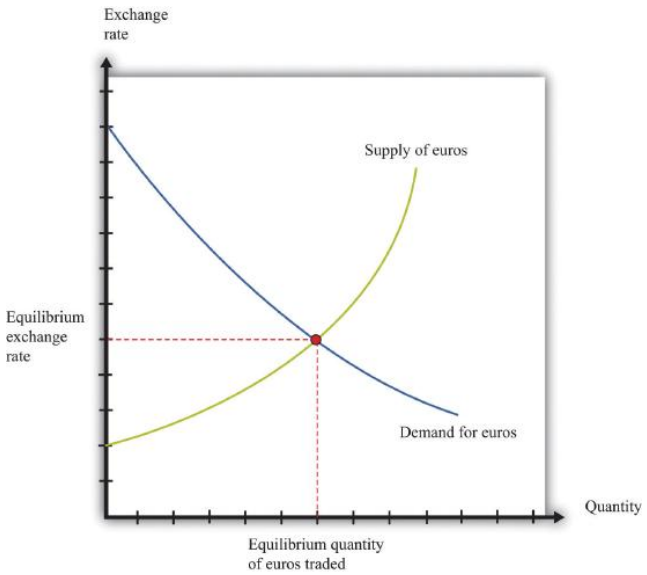


Figure 31.21.17: The Foreign Exchange Market

The foreign exchange market can be used as a basis for comparative statics exercises. We can study how changes in an economy affect the exchange rate. For example, suppose there is an increase in the level of economic activity in the United States. This leads

to an increase in the demand for European goods and services. To make these purchases, US households and firms will demand more euros. This causes an outward shift in the demand curve and an increase in the dollar price of euros.

When the dollar price of a euro increases, we say that the dollar has depreciated relative to the euro. From the perspective of the euro, the depreciation of the dollar represents an appreciation of the euro.

Key Insight

- As the exchange rate increases (so a currency becomes more valuable), a greater quantity of the currency is supplied to the market and a smaller quantity is demanded.

The Main Uses of This Tool

- Chapter 19
- Chapter 22
- Chapter 24
- Chapter 25

This page titled 31.21: Foreign Exchange Market is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.22: Growth Rates 31.22: Growth RatesEdit Update titlecancel

If some variable x (for example, the number of gallons of gasoline sold in a week) changes from x_1 to x_2, then we can define the change in that variable as $\Delta x = x_2 - x_1$. But there are difficulties with this simple definition. The number that we calculate will change, depending on the units in which we measure x. If we measure in millions of gallons, x will be a much smaller number than if we measure in gallons. If we measured x in liters rather than gallons (as it is measured in most countries), it would be a bigger number. So the number we calculate depends on the units we choose. To avoid these problems, we look at percentage changes and express the change as a fraction of the individual value. In what follows, we use the notation $\%\Delta x$ to mean the percentage change in x and define it as follows: $\%\Delta x = (x_2 - x_1)/x_1$. A percentage change equal to 0.1 means that gasoline consumption increased by 10 percent. Why? Because 10 percent means 10 "per hundred," so 10 percent = 10/100 = 0.1.

Very often in economics, we are interested in changes that take place over time. Thus we might want to compare gross domestic product (GDP) between 2012 and 2013. Suppose we know that GDP in the United States in 2012 was $14 trillion and that GDP in 2013 was $14.7 trillion. Using the letter Y to denote GDP measured in trillions, we write Y_{2012} = 14.0 and Y_{2013} = 14.7. If we want to talk about GDP at different points in time without specifying a particular year, we use the notation Y_t. We express the change in a variable over time in the form of a growth rate, which is just an example of a percentage change. Thus the growth rate of GDP in 2013 is calculated as follows:

$$\%\Delta Y_{2013} = (Y_{2013} - Y_{2012})/Y_{2012} = (14.7 - 14)/14 = 0.05$$

The growth rate equals 5 percent. In general, we write $\%\Delta Y_{t+1} = (Y_{t+1} - Y_t)/Y_t$. Occasionally, we use the *gross growth rate*, which simply equals 1 + the growth rate. So, for example, the gross growth rate of GDP equals Y_{2013}/Y_{2012}, or 1.05.

There are some useful rules that describe the behavior of percentage changes and growth rates.

The Product Rule. Suppose we have three variables, x, y, and z, and suppose

$$x = yz.$$

Then

$$\%\Delta x = \%\Delta y + \%\Delta z$$

In other words, the growth rate of a product of two variables equals the sum of the growth rates of the individual variables.

The Quotient Rule. Now suppose we rearrange our original equation by dividing both sides by z to obtain

If we take the product rule and subtract %Δz from both sides, we get the following:

$$\%\Delta x = \%\Delta y - \%\Delta z$$

The Power Rule. There is one more rule of growth rates that we make use of in some advanced topics, such as growth accounting. Suppose that

$$y = x^a$$

.

Then

$$\%\Delta y = a(\%\Delta x)$$

.

For example, if $y = x^2$, then the growth rate of y is twice the growth rate of x. If then the growth rate of y is half the growth rate of x (remembering that a square root is the same as a power of ½).

More Formally

Growth rates compound over time: if the growth rate of a variable is constant, then the change in the variable increases over time. For example, suppose GDP in 2020 is 20.0, and it grows at 10 percent per year. Then in 2021, GDP is 22.0 (an increase of 2.0), but in 2022, GDP is 24.2 (an increase of 2.2). If this compounding takes place every instant, then we say that we have *exponential growth*. Formally, we write exponential growth using the number e = 2.71828.... If the value of Y at time 0 equals Y_0 and if Y grows at the constant rate g (where g is an "annualized" or per year growth rate), then at time t (measured in years),

$$Y_t = e^{gt} Y_0$$

.

A version of this formula can also be used to calculate the average growth rate of a variable if we know its value at two different times. We can write the formula as

$$e^{gt} = Y_t / Y_0$$

,

which also means

$$gt = ln(Y_t / Y_0$$

,

where ln() is the natural logarithm. You do not need to know exactly what this means; you can simply calculate a logarithm using a scientific calculator or a spreadsheet. Dividing by *t* we get the average growth rate

$$g = ln(Y_t / Y_0)/t$$

.

For example, suppose GDP in 2020 is 20.0 and GDP in 2030 is 28.0. Then Y_{2030}/Y_{2020} = 28/20 = 1.4. Using a calculator, we can find ln(1.4) = 0.3364. Dividing by 10 (since the two dates are 10 years apart), we get an average growth rate of 0.034, or 3.4 percent per year.

The Main Uses of This Tool

- Chapter 18
- Chapter 20
- Chapter 21
- Chapter 22
- Chapter 24
- Chapter 26

This page titled 31.22: Growth Rates is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.23: Mean and Variance 31.23: Mean and VarianceEdit Update titlecancel

To start our presentation of descriptive statistics, we construct a data set using a spreadsheet program. The idea is to simulate the flipping of a two-sided coin. Although you might think it would be easier just to flip a coin, doing this on a spreadsheet gives you a full range of tools embedded in that program. To generate the data set, we drew 10 random numbers using the spreadsheet program. In the program we used, the function was called RAND and this generated the choice of a number between zero and one. Those choices are listed in the second column of Table 31.2.10.

The third column creates the two events of heads and tails that we normally associate with a coin flip. To generate this last column, we adopted a rule: if the random number was less than 0.5, we termed this a "tail" and assigned a 0 to the draw; otherwise we termed it a "head" and assigned a 1 to the draw. The choice of 0.5 as the cutoff for heads reflects the fact that we are considering the flips of a fair coin in which each side has the same probability: 0.5.

Table 31.23.10

Draw	Random Number	Heads (1) or Tails (0)
1	0.94	1
2	0.84	1
3	0.26	0
4	0.04	0
5	0.01	0
6	0.57	1
7	0.74	1
8	0.81	1
9	0.64	1
10	0.25	0

Keep in mind that the realization of the random number in draw *i* is independent of the realizations of the random numbers in both past and future draws. Whether a coin comes up heads or tails on any particular flip does not depend on other outcomes.

There are many ways to summarize the information contained in a sample of data. Even before you start to compute some complicated statistics, having a way to present the data is important. One possibility is a bar graph in which the fraction of observations of each outcome is easily shown. Alternatively, a pie chart is often used to display this fraction. Both the pie chart and the bar diagram are commonly found in spreadsheet programs.

Economists and statisticians often want to describe data in terms of numbers rather than figures. We use the data from the table to define and illustrate two statistics that are commonly used in economics discussions. The first is the mean (or average) and is a measure of *central tendency*. Before you read any further, ask, "What do you think the average ought to be from the coin flipping exercise?" It is natural to say 0.5, since half the time the outcome will be a head and thus have a value of zero, whereas the remainder of the time the outcome will be a tail and thus have a value of one.

Whether or not that guess holds can be checked by looking at Table 31.2.10 and calculating the mean of the outcome. We let k_i be the outcome of draw *i*. For example, from the table, $k_1 = 1$ and $k_5 = 0$. Then the formula for the mean if there are N draws is $\mu = \Sigma_i k_j / N$. Here $\Sigma_i k_i$ means the sum of the k_i outcomes. In words, the mean, denoted by μ, is calculated by adding together the draws and dividing by the number of draws (N). In the table, $N = 10$, and the sum of the draws of random numbers is about 51.0. Thus the mean of the 10 draws is about 0.51.

We can also calculate the mean of the heads/tails column, which is 0.6 since heads came up 6 times in our experiment. This calculation of the mean differs from the mean of the draws since the numbers in the two columns differ with the third column being a very discrete way to represent the information in the second column.

A second commonly used statistic is a measure of dispersion of the data called the variance. The variance, denoted σ^2, is calculated as $\sigma^2 = \Sigma_j (k_i - \mu)^2 / (N)$. From this formula, if all the draws were the same (thus equal to the mean), then the variance would be zero. As the draws spread out from the mean (both above and below), the variance increases. Since some observations are above the mean and others below, we square the difference between a single observation (k_i) and the mean (μ) when calculating the variance. This means that values above and below the mean both contribute a positive amount to the variance. Squaring also means that values that are a long way away from the mean have a big effect on the variance.

For the data given in the table, the mean of the 10 draws was given as $\mu = 0.51$. So to calculate the variance, we would subtract the mean from each draw, square the difference, and then add together the squared differences. This yields a variance of 0.118 for this draw. A closely related concept is that of the standard deviation, which is the square root of the variance. For our example, the standard deviation is 0.34. The standard deviation is greater than the variance since the variance is less than 1.

The Main Uses of This Tool

- Chapter 20
- Chapter 22
- Chapter 26

This page titled 31.23: Mean and Variance is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.24: Correlation and Causality 31.24: Correlation and CausalityEdit Update titlecancel

Correlation is a statistical measure describing how two variables move together. In contrast, causality (or causation) goes deeper into the relationship between two variables by looking for cause and effect.

Correlation is a statistical property that summarizes the way in which two variables move either over time or across people (firms, governments, etc.). The concept of correlation is quite natural to us, as we often take note of how two variables interrelate. If you think back to high school, you probably have a sense of how your classmates did in terms of two measures of performance: grade point average (GPA) and the results on a standardized college entrance exam (SAT or ACT). It is likely that classmates with high GPAs also had high scores on the SAT or ACT exam. In this instance, we would say that the GPA and SAT/ACT scores were positively correlated: looking across your classmates, when a person's GPA is higher than average, that person's SAT or ACT score is likely to be higher than average as well.

As another example, consider the relationship between a household's income and its expenditures on housing. If you conducted a survey across households, it is likely that you would find that richer households spend more on most goods and services, including housing. In this case, we would conclude that income and expenditures on housing are positively correlated.

When economists look at data for a whole economy, they often focus on a measure of how much is produced, which we call real gross domestic product (real GDP), and the fraction of workers without jobs, called the unemployment rate. Over long periods of time, when GDP is above average (the economy is doing well), the unemployment rate is below average. In this case, GDP and the unemployment rate are negatively correlated, as they tend to move in opposite directions.

The fact that one variable is correlated with another does not inform us about whether one variable *causes* the other. Imagine yourself on an airplane in a relaxed mood, reading or listening to music. Suddenly, the pilot comes on the public address system and requests that you buckle your seat belts. Usually, such a request is followed by turbulence. This is a correlation: the announcement by the pilot is positively correlated with air turbulence. The correlation is of course not perfect because sometimes you hit some bumps without warning, and sometimes the pilot's announcement is not followed by turbulence.

But—obviously—this does not mean that we could solve the turbulence problem by turning off the public address system. The pilot's announcement does not *cause* the turbulence. The turbulence is there whether the pilot announces it or not. In fact, the causality runs the other way. The turbulence causes the pilot's announcement.

We noted earlier that real GDP and unemployment are negatively correlated. When real GDP is below average, as it is during a recession, the unemployment rate is typically above average. But what is the causality here? If unemployment caused recessions, we might be tempted to adopt a policy that makes unemployment illegal. For example, the government could fine firms if they lay off workers. This is not a good policy because we do not think that low unemployment causes high real GDP. Neither do we necessarily think that high real GDP causes low unemployment. Instead, based on economic theory, there are other influences that affect both real GDP and unemployment.

More Formally

Suppose you have N observations of two variables, x and y, where x_i and y_i are the values of these variables in observation $i = 1, 2,\ldots, N$. The mean of x, denoted μ_x, is the sum over the values of x in the sample is divided by N; the scenario applies for y.

and

We can also calculate the variance and standard deviations of x and y. The calculation for the variance of x, denoted is as follows:

The standard deviation of x is the square root of

With these ingredients, the correlation of (x,y), denoted $\text{corr}(x,y)$, is given by

The Main Uses of This Tool

- Chapter 22
- Chapter 26
- Chapter 29

This page titled 31.24: Correlation and Causality is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.25: The Credit (Loan) Market (Macro) 31.25: The Credit (Loan) Market (Macro)Edit Update titlecancel

Consider a simple example of a loan. Imagine you go to your bank to inquire about a loan of $1,000, to be repaid in one year's time. A loan is a contract that specifies three things:

1. The amount being borrowed (in this example, $1,000)
2. The date(s) at which repayment must be made (in this example, one year from now)
3. The amount that must be repaid

What determines the amount of the repayment? The lender—the bank—is a supplier of credit, and the borrower—you—is a demander of credit. We use the terms *credit* and *loans* interchangeably. The higher the repayment amount, the more attractive this loan contract will look to the bank. Conversely, the lower the repayment amount, the more attractive this contract is to you.

If there are lots of banks that are willing to supply such loans, and lots of people like you who demand such loans, then we can draw supply and demand curves in the credit (loan) market. The equilibrium price of this loan is the interest rate at which supply equals demand.

In macroeconomics, we look at not only individual markets like this but also the credit (loan) market for an entire economy. This market brings together suppliers of loans, such as households that are saving, and demanders of loans, such as businesses and households that need to borrow. The **real interest rate** is the "price" that brings demand and supply into balance.

The supply of loans in the domestic loans market comes from three different sources:

1. The private saving of households and firms
2. The saving of governments (in the case of a government surplus)
3. The saving of foreigners (when there is a flow of capital into the domestic economy)

Households will generally respond to an increase in the real interest rate by reducing current consumption relative to future consumption. Households that are saving will save more; households that are borrowing will borrow less. Higher interest rates also encourage foreigners to send funds to the domestic economy. Government saving or borrowing is little affected by interest rates.

The demand for loans comes from three different sources:

1. The borrowing of households and firms to finance purchases, such as housing, durable goods, and investment goods
2. The borrowing of governments (in the case of a government deficit)
3. The borrowing of foreigners (when there is a flow of capital from the domestic economy)

As the real interest rate increases, investment and durable goods spending decrease. For firms, a high interest rate represents a high cost of funding investment expenditures. This is an application of discounted present value and is evident if a firm borrows to purchase capital. It is also true if it uses internal funds (retained earnings) to finance investment because the firm could always put those funds into an interest-bearing asset instead. For households, higher interest rates likewise make it more costly to borrow to purchase housing and durable goods. The demand for credit decreases as the interest rate rises. When it is expensive to borrow, households and firms will borrow less.

Equilibrium in the market for loans is shown in Figure 31.25.18. On the horizontal axis is the total quantity of loans in equilibrium. The demand curve for loans is downward sloping, whereas the supply curve has a positive slope. Loan market equilibrium occurs at the real interest rate where the quantity of loans supplied equals the quantity of loans demanded. At this equilibrium real interest rate, lenders lend as much as they wish, and borrowers can borrow as much as they wish. Equilibrium in the aggregate credit market is what ensures the balance of flows into and out of the financial sector in the circular flow diagram.

Key Insights

- As the real interest rate increases, more loans are supplied, and fewer loans are demanded.
- Adjustment of the real interest rate ensures that, in the circular flow diagram, the flows into the financial sector equal the flows from the sector.

The Main Uses of This Tool

- Chapter 19
- Chapter 24

- Chapter 25
- Chapter 26
- Chapter 29

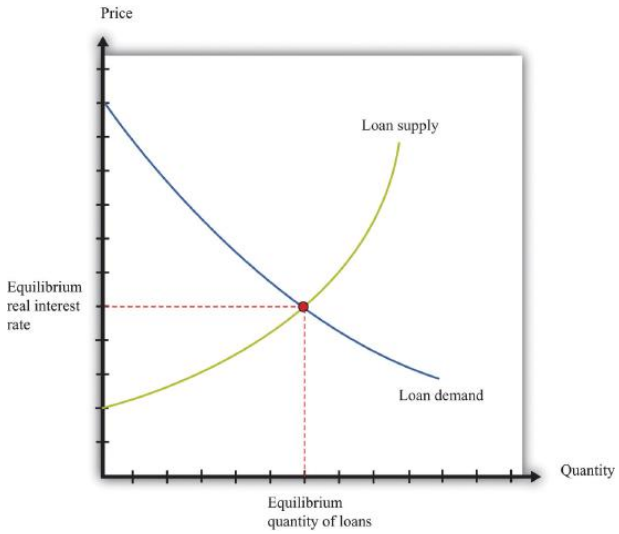


Figure 31.25.18: The Credit Market

This page titled 31.25: The Credit (Loan) Market (Macro) is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.26: The Fisher Equation: Nominal and Real Interest Rates 31.26: The Fisher Equation: Nominal and Real Interest RatesEdit Update titlecancel

When you borrow or lend, you normally do so in dollar terms. If you take out a loan, the loan is denominated in dollars, and your promised payments are denominated in dollars. These dollar flows must be corrected for inflation to calculate the repayment in real terms. A similar point holds if you are a lender: you need to calculate the interest you earn on saving by correcting for inflation.

The **Fisher equation** provides the link between nominal and real interest rates. To convert from nominal interest rates to real interest rates, we use the following formula:

$$real\ interest\ rate \approx nominal\ interest\ rate - inflation\ rate.$$

To find the real interest rate, we take the nominal interest rate and subtract the inflation rate. For example, if a loan has a 12 percent interest rate and the inflation rate is 8 percent, then the real return on that loan is 4 percent.

In calculating the real interest rate, we used the actual inflation rate. This is appropriate when you wish to understand the real interest rate actually paid under a loan contract. But at the time a loan agreement is made, the inflation rate that will occur in the future is not known with certainty. Instead, the borrower and lender use their *expectations* of future inflation to determine the interest rate on a loan. From that perspective, we use the following formula:

$$contracted\ nominal\ interest\ rate \approx real\ interest\ rate + expected\ inflation\ rate.$$

We use the term *contracted nominal interest rate* to make clear that this is the rate set at the time of a loan agreement, not the realized real interest rate.

Key Insight

- To correct a nominal interest rate for inflation, subtract the inflation rate from the nominal interest rate.

More Formally

Imagine two individuals write a loan contract to borrow *P* dollars at a nominal interest rate of *i*. This means that next year the amount to be repaid will be $P \times (1 + i)$. This is a standard loan contract with a nominal interest rate of *i*.

Now imagine that the individuals decided to write a loan contract to guarantee a constant real return (in terms of goods not dollars) denoted *r*. So the contract provides *P* this year in return for being repaid (enough dollars to buy) (1 + *r*) units of real gross domestic product (real GDP) next year. To repay this loan, the borrower gives the lender enough money to buy (1 + *r*) units of real GDP for each unit of real GDP that is lent. So if the inflation rate is π, then the price level has risen to $P \times (1 + \pi)$, so the repayment in dollars for a loan of *P* dollars would be $P(1 + r) \times (1 + \pi)$.

Here (1 + π) is one plus the inflation rate. The inflation rate π_{t+1} is defined—as usual—as the percentage change in the price level from period *t* to period *t* + 1.

$$\Pi_{t+1} = (P_{t+1} - P_t) / P_t$$

.

If a period is one year, then the price level next year is equal to the price this year multiplied by (1 + π):

$$P_{t+1} = (1 + \Pi_t) \times P_t$$

.

The Fisher equation says that these two contracts should be equivalent:

$$(1 + i) = (1 + r) \times (1 + \pi).$$

As an approximation, this equation implies

$$i \approx r + \pi.$$

To see this, multiply out the right-hand side and subtract 1 from each side to obtain

$$i = r + \pi + r\Pi.$$

If r and π are small numbers, then $r\Pi$ is a very small number and can safely be ignored. For example, if $r = 0.02$ and $\Pi = 0.03$, then $r\Pi = 0.0006$, and our approximation is about 99 percent accurate.

The Main Uses of This Tool

- Chapter 19
- Chapter 24
- Chapter 25
- Chapter 26

This page titled 31.26: The Fisher Equation: Nominal and Real Interest Rates is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.27: The Aggregate Production Function 31.27: The Aggregate Production FunctionEdit Update titlecancel

The aggregate production function describes how total real gross domestic product (real GDP) in an economy depends on available inputs. Aggregate output (real GDP) depends on the following:

- Physical capital—machines, production facilities, and so forth that are used in production
- Labor—the number of hours that are worked in the entire economy
- Human capital—skills and education embodied in the workforce of the economy
- Knowledge—basic scientific knowledge, and blueprints that describe the available production processes
- Social infrastructure—the general business, legal and cultural environment
- The amount of natural resources available in an economy
- Anything else that we have not yet included

We group the inputs other than labor, physical, and human capital together, and call them technology.

The aggregate production function has several key properties. First, output increases when there are increases in physical capital, labor, and natural resources. In other words, the marginal products of these inputs are all positive.

Second, the increase in output from adding more inputs is lower when we have more of a factor. This is called diminishing marginal product. That is,

- The more capital we have, the less additional output we obtain from additional capital.
- The more labor we have, the less additional output we obtain from additional labor.
- The more natural resources we have, the less additional output we obtain from additional resources.

In addition, increases in output can also come from increases in human capital, knowledge, and social infrastructure. In contrast to capital and labor, we do not assume that there are diminishing returns to human capital and technology. One reason is that we do not have a natural or an obvious measure for human capital, knowledge, or social infrastructure, whereas we do for labor and capital (hours of work and hours of capital usage).

Figure 31.27.19 shows the relationship between output and capital, holding fixed the level of other inputs. This figure shows two properties of the aggregate production function. As capital input is increased, output increases as well. But the change in output obtained by increasing the capital stock is lower when the capital stock is higher: this is the diminishing marginal product of capital.

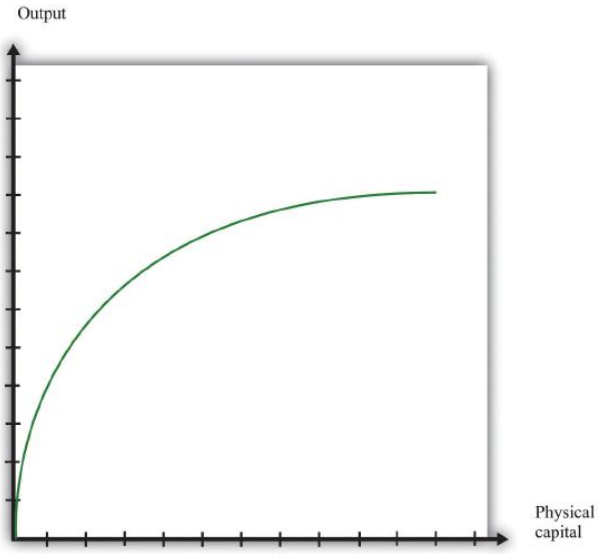


Figure 31.27.19

In many applications, we want to understand how the aggregate production function responds to variations in the technology or other inputs. This is illustrated in Figure 31.27.20. An increase in, say, technology means that for a given level of the capital stock,

more output is produced: the production function shifts upward as technology increases. Further, as technology increases, the production function is steeper: the increase in technology increases the marginal product of capital.

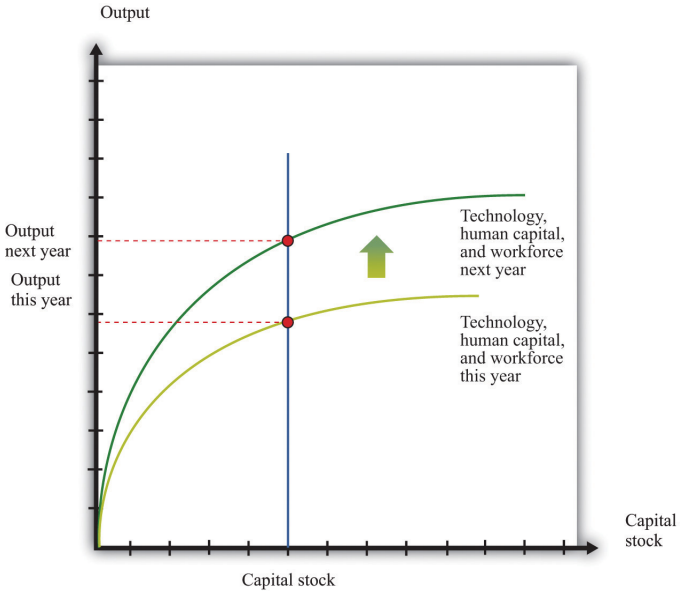


Figure 31.27.2

Key Insight

- The aggregate production function allows us to determine the output of an economy given inputs of capital, labor, human capital, and technology.

More Formally

Specific Forms for the Production Function

We can write the production function in mathematical form. We use Y to represent real GDP, K to represent the physical capital stock, L to represent labor, H to represent human capital, and A to represent technology (including natural resources). If we want to speak about production completely generally, then we can write $Y = F(K, L, H, A)$. Here $F()$ means "some function of."

A lot of the time, economists work with a production function that has a specific mathematical form, yet is still reasonably simple:

$$Y = A \times K^{a} \times (L \times H)^{(1-a)},$$

where a is just a number. This is called a **Cobb-Douglas** production function. It turns out that this production function does a remarkably good job of summarizing aggregate production in the economy. In fact, we also know that we can describe production in actual economies quite well if we suppose that $a = 1/3$.

The Main Uses of This Tool

- Chapter 20
- Chapter 21
- Chapter 22

This page titled 31.27: The Aggregate Production Function is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.28: The Circular Flow of Income 31.28: The Circular Flow of IncomeEdit Update titlecancel

The circular flow of income describes the flows of money among the five main sectors of an economy. As individuals and firms buy and sell goods and services, money flows among the different sectors of an economy. The circular flow of income describes these flows of dollars (pesos, euros, or whatever). From a simple version of the circular flow, we learn that—as a matter of accounting—

$$gross\ domestic\ product(GDP) = income = production = spending.$$

This relationship lies at the heart of macroeconomic analysis.

There are two sides to every transaction. Corresponding to the flows of money in the circular flow, there are flows of goods and services among these sectors. For example, the wage income received by consumers is in return for labor services that flow from households to firms. The consumption spending of households is in return for the goods and services that flow from firms to households.

A complete version of the circular flow is presented in Figure 31.28.21. (Chapter 18 contains a discussion of a simpler version of the circular flow with only two sectors: households and firms.)

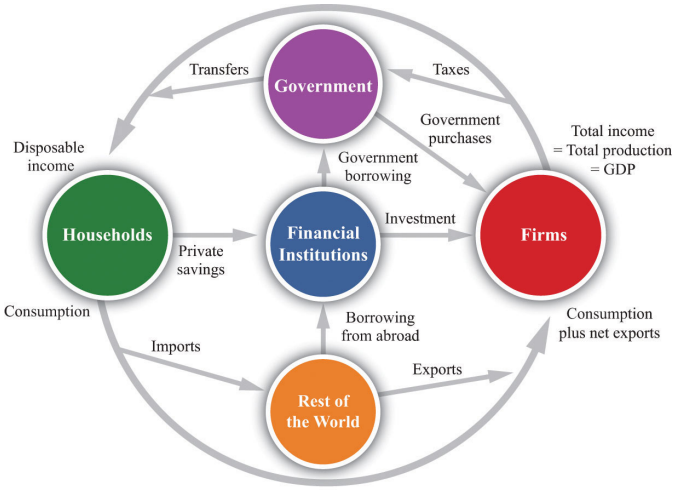


Figure 31.28.21

The complete circular flow has five sectors: a household sector, a firm sector, a government sector, a foreign sector, and a financial sector. Different chapters of the book emphasize different pieces of the circular flow, and Figure 31.28.21 shows us how everything fits together. In the following subsections, we look at the flows into and from each sector in turn. In each case, the balance of the flows into and from each sector underlies a useful economic relationship.

The Firm Sector

Figure 31.28.21 includes the component of the circular flow associated with the flows into and from the firm sector of an economy. We know that the total flow of dollars from the firm sector measures the total value of production in an economy. The total flow of dollars into the firm sector equals total expenditures on GDP. We therefore know that

$$production = consumption + investment + government\ purchases + net\ exports.$$

This equation is called the **national income identity** and is the most fundamental relationship in the national accounts.

By consumption we mean total consumption expenditures by households on final goods and services. Investment refers to the purchase of goods and services that, in one way or another, help to produce more output in the future. Government purchases include all purchases of goods and services by the government. Net exports, which equal exports minus imports, measure the expenditure flows associated with the rest of the world.

The Household Sector

The household sector summarizes the behavior of private individuals in their roles as consumers/savers and suppliers of labor. The balance of flows into and from this sector is the basis of the household budget constraint. Households receive income from firms, in the form of wages and in the form of dividends resulting from their ownership of firms. The income that households have available to them after all taxes have been paid to the government and all transfers received is called disposable income. Households spend some of their disposable income and save the rest. In other words,

$$disposable\ income = consumption + household\ savings.$$

This is the **household budget constraint**. In Figure 31.28.21, this equation corresponds to the fact that the flows into and from the household sector must balance.

The Government Sector

The government sector summarizes the actions of all levels of government in an economy. Governments tax their citizens, pay transfers to them, and purchase goods from the firm sector of the economy. Governments also borrow from or lend to the financial sector. The amount that the government collects in taxes need not equal the amount that it pays out for government purchases and transfers. If the government spends more than it gathers in taxes, then it must borrow from the financial markets to make up the shortfall.

The circular flow figure shows two flows into the government sector and two flows out. Since the flows into and from the government sector must balance, we know that

$$government\ purchases + transfers = tax\ revenues + government\ borrowing.$$

Government borrowing is sometimes referred to as the government budget deficit. This equation is the **government budget constraint**.

Some of the flows in the circular flow can go in either direction. When the government is running a deficit, there is a flow of dollars to the government sector from the financial markets. Alternatively, the government may run a surplus, meaning that its revenues from taxation are greater than its spending on purchases and transfers. In this case, the government is saving rather than borrowing, and there is a flow of dollars to the financial markets from the government sector.

The Foreign Sector

The circular flow includes a country's dealings with the rest of the world. These flows include exports, imports, and borrowing from other countries. Exports are goods and services produced in one country and purchased by households, firms, and governments of another country. Imports are goods and services purchased by households, firms, and governments in one country but produced in another country. Net exports are exports minus imports. When net exports are positive, a country is running a trade surplus: exports exceed imports. When net exports are negative, a country is running a trade deficit: imports exceed exports. The third flow between countries is borrowing and lending. Governments, individuals, and firms in one country may borrow from or lend to another country.

Net exports and borrowing are linked. If a country runs a trade deficit, it borrows from other countries to finance that deficit. If we look at the flows into and from the foreign sector, we see that

$$borrowing\ from\ other\ countries + exports = imports.$$

Subtracting exports from both sides, we obtain

$$borrowing\ from\ other\ countries = imports - exports = trade\ deficit.$$

Whenever our economy runs a trade deficit, we are borrowing from other countries. If our economy runs a trade surplus, then we are lending to other countries.

This analysis has omitted one detail. When we lend to other countries, we acquire their assets, so each year we get income from those assets. When we borrow from other countries, they acquire our assets, so we pay them income on those assets. Those income flows are added to the trade surplus/deficit to give the *current account* of the economy. It is the current account that must be matched by borrowing from or lending to other countries. A positive current account means that net exports plus net income flows from the rest of the world are positive. In this case, our economy is lending to the rest of the world and acquiring more assets.

The Financial Sector

The **financial sector** of an economy summarizes the behavior of banks and other financial institutions. The balance of flows into and from the financial sector tell us that investment is financed by national savings and borrowing from abroad. The financial sector is at the heart of the circular flow. The figure shows four flows into and from the financial sector.

1. Households divide their after-tax income between consumption and savings. Thus any income that they receive today but wish to put aside for the future is sent to the financial markets. The household sector as a whole saves so, on net, there is a flow of dollars from the household sector into the financial markets.
2. The flow of money from the financial sector into the firm sector provides the funds that are available to firms for investment purposes.
3. The flow of dollars between the financial sector and the government sector reflects the borrowing (or lending) of governments. The flow can go in either direction. When government expenditures exceed government revenues, the government must borrow from the private sector, and there is a flow of dollars from the financial sector to the government. This is the case of a government deficit. When the government's revenues are greater than its expenditures, by contrast, there is a government surplus and a flow of dollars into the financial sector.
4. The flow of dollars between the financial sector and the foreign sector can also go in either direction. An economy with positive net exports is lending to other countries: there is a flow of money from an economy. An economy with negative net exports (a trade deficit) is borrowing from other countries.

The **national savings** of the economy is the savings carried out by the private and government sectors taken together. When the government is running a deficit, some of the savings of households and firms must be used to fund that deficit, so there is less left over to finance investment. National savings is then equal to private savings minus the government deficit—that is, private savings minus government borrowing:

$$national\ savings = private\ savings - government\ borrowing.$$

If the government is running a surplus, then

$$national\ savings = private\ savings + government\ surplus.$$

National savings is therefore the amount that an economy as a whole saves. It is equal to what is left over after we subtract consumption and government spending from GDP. To see this, notice that

$$\begin{gathered} private\ savings - government\ borrowing \\ = income - taxes + transfers - consumption - (government\ purchases + transfers - taxes) \\ = income - consumption - government\ purchases. \end{gathered}$$

This is the domestic money that is available for investment.

If we are borrowing from other countries, there is another source of funds for investment. The flows into and from the financial sector must balance, so

$$investment = national\ savings + borrowing\ from\ other\ countries.$$

Conversely, if we are lending to other countries, then our national savings is divided between investment and lending to other countries:

$$national\ savings = investment + lending\ to\ other\ countries.$$

The Main Uses of This Tool

- Chapter 18
- Chapter 19
- Chapter 21
- Chapter 22
- Chapter 24
- Chapter 26
- Chapter 27
- Chapter 29
- Chapter 30

This page titled 31.28: The Circular Flow of Income is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.29: Growth Accounting 31.29: Growth AccountingEdit Update titlecancel

Growth accounting is a tool that tells us how changes in real gross domestic product (real GDP) in an economy are due to changes in available capital, labor, human capital, and technology. Economists have shown that, under reasonably general circumstances, the change in output in an economy can be written as follows:

$$output\ growth\ rate = a \times capital\ stock\ growth\ rate + [(1-a) \times labor\ hours\ growth\ rate] + [(1-a) \times human\ capital\ growth\ rate] + technology\ growth\ rate.$$

In this equation, *a* is just a number. For example, if *a* = 1/3, the growth in output is as follows:

$$output\ growth\ rate = (1/3 \times capital\ stock\ growth\ rate) + (2/3 \times labor\ hours\ growth\ rate) + (2/3 \times human\ capital\ growth\ rate) + technology\ growth\ rate.$$

Growth rates can be positive or negative, so we can use this equation to analyze decreases in GDP as well as increases. This expression for the growth rate of output, by the way, is obtained by applying the rules of growth rates (discussed in Section 31.2.21 "Growth Rates") to the Cobb-Douglas aggregate production function (discussed in 31.3 Section "The Aggregate Production Function").

What can we measure in this expression? We can measure the growth in output, the growth in the capital stock, and the growth in labor hours. Human capital is more difficult to measure, but we can use information on schooling, literacy rates, and so forth. We cannot, however, measure the growth rate of technology. So we use the growth accounting equation to infer the growth in technology from the things we can measure. Rearranging the growth accounting equation,

$$technology\ growth\ rate = output\ growth\ rate - (a \times capital\ stock\ growth\ rate) - [(1-a) \times labor\ hours\ growth\ rate] - [(1-a) \times human\ capital\ growth\ rate].$$

So if we know the number *a*, we are done—we can use measures of the growth in output, labor, capital stock, and human capital to solve for the technology growth rate. In fact, we do have a way of measuring *a*. The technical details are not important here, but a good measure of (1 − *a*) is simply the total payments to labor in the economy (that is, the total of wages and other compensation) as a fraction of overall GDP. For most economies, *a* is in the range of about 1/3 to 1/2.

Key Insight

- The growth accounting tool allows us to determine the contributions of the various factors of economic growth.

The Main Uses of This Tool

- Chapter 20
- Chapter 21
- Chapter 22

This page titled 31.29: Growth Accounting is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.30: The Solow Growth Model 31.30: The Solow Growth ModelEdit Update titlecancel

The analysis in Chapter 21 is (implicitly) based on a theory of economic growth known as the *Solow growth model*. Here we present two formal versions of the mathematics of the model. The first takes as its focus the capital accumulation equation and explains how the capital stock evolves in the economy. This version ignores the role of human capital and ignores the long-run growth path of the economy. The second follows the exposition of the chapter and is based around the derivation of the balanced growth path. They are, however, simply two different ways of approaching the same problem.

Presentation 1

There are three components of this presentation of the model: technology, capital accumulation, and saving. The first component of the Solow growth model is the specification of technology and comes from the aggregate production function. We express output per worker (y) as a function of capital per worker (k) and technology (A). A mathematical expression of this relationship is

$$y = Af(k),$$

where $f(k)$ means that output per worker depends on capital per worker. As in our presentation of production functions, output increases with technology. We assume that $f()$ has the properties that more capital leads to more output per capita at a diminishing rate. As an example, suppose

$$y = Ak^{1/3}.$$

In this case the marginal product of capital is positive but diminishing.

The second component is capital accumulation. If we let k_t be the amount of capital per capita at the start of year t, then we know that

$$k_{t+1} = k_t(1 - \delta) + i_t.$$

This expression shows how the capital stock changes over time. Here δ is the rate of physical depreciation so that between year t and year t +1, δk_t units of capital are lost from depreciation. But during year t, there is investment (i_t) that yields new capital in the following year.

The final component of the Solow growth model is saving. In a closed economy, saving is the same as investment. Thus we link i_t in the accumulation equation to saving. Assume that saving per capita (s_t) is given by

$$s_t = s \times y_t.$$

Here s is a constant between zero and one, so only a fraction of total output is saved.

Using the fact that savings equals investment, along with the per capita production function, we can relate investment to the level of capital:

\[i_{t]=sAf(k_{t}). \nonumber \]

We can then write the equation for the evolution of the capital stock as follows:

$$k_{t+1} = k_l(1 - \delta) + sAf\left(k_t\right).$$

Once we have specified the function $f()$, we can follow the evolution of the capital stock over time. Generally, the path of the capital stock over time has two important properties:

1. **Steady state.** There is a particular level of the capital stock such that if the economy accumulates that amount of capital, it stays at that level of capital. We call this the steady state level of capital, denoted k^*.
2. **Stability.** The economy will tend toward the per capita capital stock k^*.

To be more specific, the steady state level of capital solves the following equation:

$$k^* = k^*(1 - \delta) + sAf\left(k^*\right).$$

At the steady state, the amount of capital lost by depreciation is exactly offset by saving. This means that at the steady state, net investment is exactly zero. The property of stability means that if the current capital stock is below k^*, the economy will

accumulate capital so that $k_{t+1} > k_t$. And if the current capital stock is above *k**, the economy will decumulate capital so that \ (*k_{t+1}<k_{t}*\).

If two countries share the same technology (*A*) and the same production function [*f*(k)], then over time these two countries will eventually have the same stock of capital per worker. If there are differences in the technology or the production function, then there is no reason for the two countries to converge to the same level of capital stock per worker.

Presentation 2

In this presentation, we explain the balanced-growth path of the economy and prove some of the claims made in the text. The model takes as given (exogenous) the investment rate; the depreciation rate; and the growth rates of the workforce, human capital, and technology. The endogenous variables are output and physical capital stock.

The notation for the presentation is given in Table 31.2.11 "Notation in the Solow Growth Model": We use the notation g_x to represent the growth rate of a variable *x*; that is,

There are two key ingredients to the model: the aggregate production function and the equation for capital accumulation.

Table 31.30.11: Notation in the Solow Growth Model

Variable	Symbol
Real gross domestic product	*Y*
Capital stock	*K*
Human capital	*H*
Workforce	*L*
Technology	*A*
Investment rate	*i*
Depreciation rate	δ

The Production Function

The production function we use is the *Cobb-Douglas production function*:

$$Y = K^a(HL)^{1-a}A.$$

Growth Accounting

If we apply the rules of growth rates to Equation 31.2.?????, we get the following expression:

$$g_Y = ag_k + (1-a)(g_L + g_H) + g_A.$$

Balanced Growth

The condition for balanced growth is that $g_Y = g_K$. When we impose this condition on our equation for the growth rate of output (Equation 31.2.???????), we get

where the superscript “BG” indicates that we are considering the values of variables when the economy is on a balanced growth path. This equation simplifies to

Equation 31.3

The growth in output on a balanced-growth path depends on the growth rates of the workforce, human capital, and technology.

Using this, we can rewrite Equation 31.2.????????? as follows:

Equation 31.4

The actual growth rate in output is an average of the balanced-growth rate of output and the growth rate of the capital stock.

Capital Accumulation

The second piece of our model is the capital accumulation equation. The growth rate of the capital stock is given by

Equation 31.5

Divide the numerator and denominator of the first term by *Y*, remembering that $i = I/Y$.

Equation 31.6

The growth rate of the capital stock depends positively on the investment rate and negatively on the depreciation rate. It also depends negatively on the current capital-output ratio.

The Balanced-Growth Capital-Output Ratio

Now rearrange Equation 31.6 to give the ratio of capital to gross domestic product (GDP), given the depreciation rate, the investment rate, and the growth rate of the capital stock:

When the economy is on a balanced growth path, g_K = , so

We can also substitute in our balanced-growth expression for (Equation 31.3) to get an expression for the balanced-growth capital output ratio in terms of exogenous variables.

Convergence

The proof that economies will converge to the balanced-growth ratio of capital to GDP is relatively straightforward. We want to show that if *K*/*Y* < then capital grows faster than output. If capital is growing faster than output, $g_K - g_Y > 0$. First, go back to Equation 31.4:

Subtract both sides from the growth rate of capital:

Now compare the general expression for ratio of capital to GDP with its balanced growth value:

and

If *K*/*Y* < then it must be the case that $g_K >$, which implies (from the previous equation) that $g_K > g_Y$.

Output per Worker Growth

If we want to examine the growth in output per worker rather than total output, we take the per-worker production function (Equation 31.2) and apply the rules of growth rates to that equation.

\[(1 − *a*)g_{Y} = a[g_{K}-g_{Y}] + (1 − *a*)[g_{L}+g_{H}] + g_{A} = *a*[g_{K}-g_{Y}] + (1 − *a*)[g_{L}+g_{H}] + g_{A}.

We then we divide by (1 − *a*) to get

and subtract g_L from each side to obtain

Finally, we note that $g_Y - g_L = gY/L$:

With balanced growth, the first term is equal to zero, so

Endogenous Investment Rate

In this analysis, we made the assumption from the Solow model that the investment rate is constant. The essential arguments that we have made still apply if the investment rate is higher when the marginal product of capital is higher. The argument for convergence becomes stronger because a low value of *K*/*Y* implies a higher marginal product of capital and thus a higher investment rate. This increases the growth rate of capital and causes an economy to converge more quickly to its balanced-growth path.

Endogenous Growth

Take the production function

$$Y = K^a(HL)^{1-a}A$$

Now assume *A* is constant and so

The Main Uses of This Tool

- Chapter 20
- Chapter 21

This page titled 31.30: The Solow Growth Model is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.31: The Aggregate Expenditure Model 31.31: The Aggregate Expenditure ModelEdit Update titlecancel

The aggregate expenditure model relates the components of spending (consumption, investment, government purchases, and net exports) to the level of economic activity. In the short run, taking the price level as fixed, the level of spending predicted by the aggregate expenditure model determines the level of economic activity in an economy.

An insight from the circular flow is that real gross domestic product (real GDP) measures three things: the production of firms, the income earned by households, and total spending on firms' output. The aggregate expenditure model focuses on the relationships between production (GDP) and planned spending:

$$GDP = planned\ spending = consumption + investment + government\ purchases + net\ exports.$$

Planned spending depends on the level of income/production in an economy, for the following reasons:

- If households have higher income, they will increase their spending. (This is captured by the consumption function.)
- Firms are likely to decide that higher levels of production—particularly if they are expected to persist—mean that they should build up their capital stock and should thus increase their investment.
- Higher income means that domestic consumers are likely to spend more on imported goods. Since net exports equal exports minus imports, higher imports means lower net exports.

The negative net export link is not large enough to overcome the other positive links, so we conclude that when income increases, so also does planned expenditure. We illustrate this in Figure 31.31.22 where we suppose for simplicity that there is a linear relationship between spending and GDP. The equation of the line is as follows:

$$spending = autonomous\ spending + marginal\ propensity\ to\ spend \times real\ GDP.$$

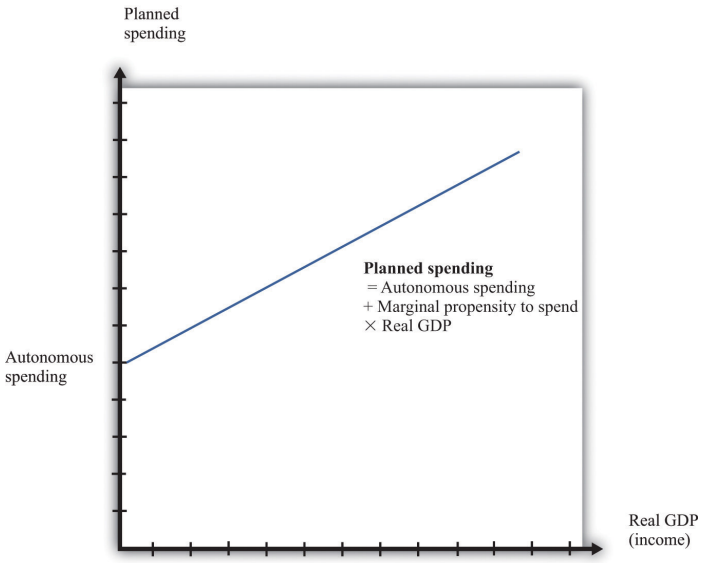


Figure 31.31.22: Planned Spending in the Aggregate Expenditure Model

The intercept in Figure 31.31.22 is called **autonomous spending**. It represents the amount of spending that there would be in an economy if income (GDP) were zero. We expect that this will be positive for two reasons: (1) if a household finds its income is zero, it will still want to consume something, so it will either draw on its existing wealth (past savings) or borrow against future income; and (2) the government would spend money even if GDP were zero.

The slope of the line in Figure 31.31.22 is given by the **marginal propensity to spend**. For the reasons that we have just explained, we expect that this is positive: increases in income lead to increased spending. However, we expect the marginal propensity to spend to be less than one.

The aggregate expenditure model is based on the two equations we have just discussed. We can solve the model either graphically or using algebra. The graphical approach relies on Figure 31.31.23. On the horizontal axis is the level of real GDP. On the vertical axis is the level of spending as well as the level of GDP. There are two lines shown. The first is the 45° line, which equates real

GDP on the horizontal axis with real GDP on the vertical axis. The second line is the planned spending line. The intersection of the spending line with the 45° line gives the equilibrium level of output.

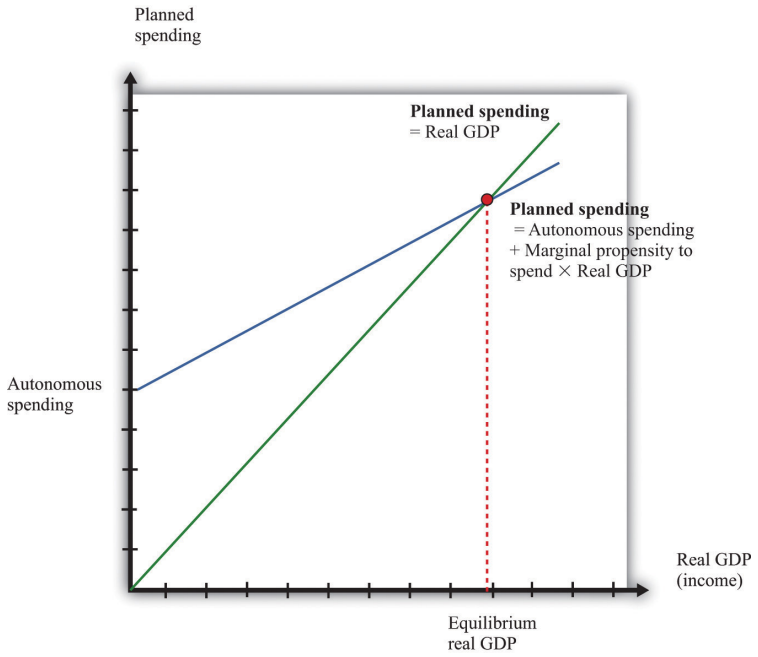


Figure 31.31.23

More Formally

We can also solve the model algebraically. Let us use Y to denote the level of real GDP and E to denote planned expenditure. We represent the marginal propensity to spend by β. The two equations of the model are as follows:

$$Y = E$$

and

$$E = E_0 + \beta \times Y$$

Here, E_0 is autonomous expenditure. We can solve the two equations to find the values of E and Y that are consistent with both equations. Substituting for E in the first equation, we find that

The equilibrium level of output is the product of two terms. The first term—$(1/(1-\beta))$—is called the multiplier. If, as seems reasonable, β lies between zero and one, the multiplier is greater than one. The second term is the autonomous level of spending.

Here is an example. Suppose that

$C = 100 + 0.6Y$,

$I = 400$,

$G = 300$,

and

$NX = 200 - 0.1Y$,

where C is consumption, I is investment, G is government purchases, and NX is net exports. First group the components of spending as follows:

$$C + I + G + NX = (100 + 400 + 300 + 200) + (0.6Y - 0.1Y)$$

Adding together the first group of terms, we find autonomous spending:

$$E_0 = 100 + 400 + 300 + 200 = 1,000.$$

Adding the coefficients on the income terms, we find the marginal propensity to spend:

$$\beta = 0.6 - 0.1 = 0.5.$$

Using β = 0.5, we calculate the multiplier:

We then calculate real GDP:

$$Y = 2 \times 1,000 = 2,000.$$

The Main Uses of This Tool

- Chapter 22
- Chapter 25
- Chapter 27
- Chapter 30

This page titled 31.31: The Aggregate Expenditure Model is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.32: Price Adjustment 31.32: Price AdjustmentEdit Update titlecancel

The price adjustment equation summarizes, at the level of an entire economy, all the decisions about prices that are made by managers throughout the economy. The **price adjustment equation** is as follows:

$$inflation\ rate = autonomous\ inflation - inflation\ sensitivity \times output\ gap.$$

The equation tells us that there are two reasons for rising prices. The first is because the **output gap** is negative. The output gap is the difference between potential output and actual output:

$$output\ gap = potential\ real\ gross\ domestic\ product(real\ GDP) - actual\ real\ GDP.$$

A positive gap means that the economy is in recession—below potential output. If the economy is in a boom, then the output gap is negative.

The second reason for rising prices is that **autonomous inflation** is positive. Autonomous inflation refers to the inflation rate that prevails in an economy when an economy is at potential output (so the output gap is zero). Looking at the second term of the price adjustment equation, we see that when real GDP is greater than potential output, the output gap is negative, so there is upward pressure on prices in the economy. The inflation rate will exceed autonomous inflation. By contrast, when real GDP is less than potential output, the output gap is negative, so there is downward pressure on prices. The inflation rate will be below the autonomous inflation rate. The “inflation sensitivity” tells us how responsive the inflation rate is to the output gap.

The output gap matters because, as GDP increases relative to potential output, labor and other inputs become scarcer. Firms are likely to see rising costs and increase their prices as a consequence. Even leaving this aside—that is, even when an economy is at potential output—firms are likely to increase their prices somewhat. For example, firms may anticipate that their suppliers or their competitors are likely to increase prices in the future. A natural response is to increase prices, so autonomous inflation is positive. Figure 31.32.24 shows the price adjustment equation graphically.

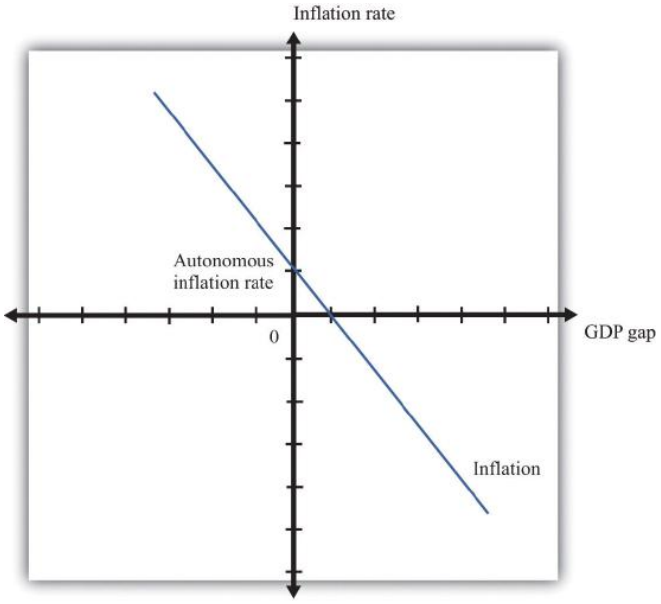


Figure 31.32.24: Price Adjustment

The Main Uses of This Tool

- Chapter 22
- Chapter 25
- Chapter 26
- Chapter 27

This page titled 31.32: Price Adjustment is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.33: Consumption and Saving 31.33: Consumption and SavingEdit Update titlecancel

The consumption function is a relationship between current disposable income and current consumption. It is intended as a simple description of household behavior that captures the idea of consumption smoothing. We typically suppose the consumption function is upward-sloping but has a slope less than one. So as disposable income increases, consumption also increases but not as much. More specifically, we frequently assume that consumption is related to disposable income through the following relationship:

$$consumption = autonomous\ consumption + marginal\ propensity\ to\ consume \times disposable\ income.$$

A consumption function of this form implies that individuals divide additional income between consumption and saving.

- We assume autonomous consumption is positive. Households consume something even if their income is zero. If a household has accumulated a lot of wealth in the past or if a household expects its future income to be larger, autonomous consumption will be larger. It captures both the past and the future.
- We assume that the marginal propensity to consume is positive. The marginal propensity to consume captures the present; it tells us how changes in current income lead to changes in current consumption. Consumption increases as current income increases, and the larger the marginal propensity to consume, the more sensitive current spending is to current disposable income. The smaller the marginal propensity to consume, the stronger is the consumption-smoothing effect.
- We also assume that the marginal propensity to consume is less than one. This says that not all additional income is consumed. When a household receives more income, it consumes some and saves some.

Figure 31.33.25 shows this relationship.

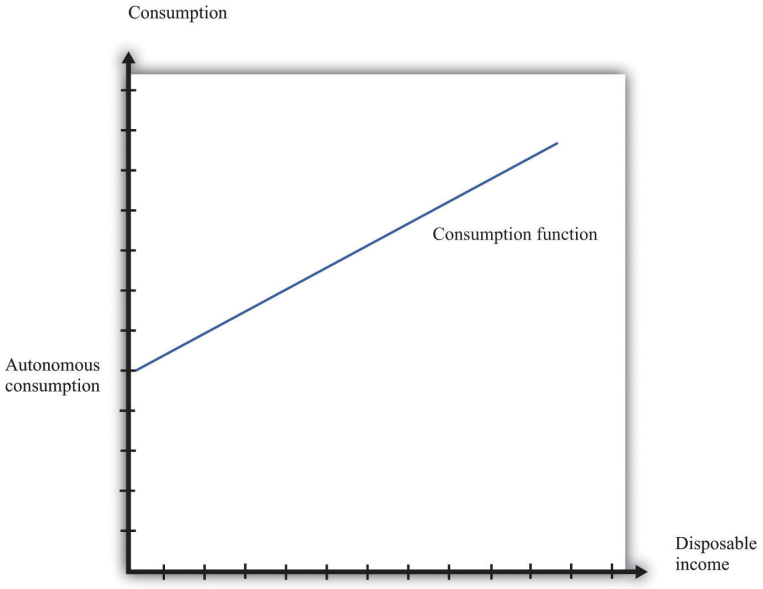


Figure 31.33.25: The Consumption Function

More Formally

In symbols, we write the consumption function as a relationship between consumption (C) and disposable income (Y^d):

\[C = a + bY^{d} \nonumber \]

where a and b are constants. Here a represents autonomous consumption and b is the marginal propensity to consume. We assume three things about a and b:

1. $a > 0$
2. $b > 0$
3. $b < 1$

The first assumption means that even if disposable income is zero ($Y^d = 0$), consumption will still be positive. The second assumption means that the marginal propensity to consume is positive. By the third assumption, the marginal propensity to consume is less that one. With $0 < b < 1$, part of an extra dollar of disposable income is spent.

What happens to the remainder of the increase in disposable income? Since consumption plus saving is equal to disposable income, the increase in disposable income not consumed is saved. More generally, this link between consumption and saving (S) means that our model of consumption implies a model of saving as well.

Using

$$Y^d = C + S$$

and

$$C = a + bY^d$$

we can solve for S:

$$S = Y^d - C = -a + (1 - b)Y^d.$$

So $-a$ is the level of autonomous saving and $(1 - b)$ is the marginal propensity to save.

We can also graph the savings function. The savings function has a negative intercept because when income is zero, the household will dissave. The savings function has a positive slope because the marginal propensity to save is positive.

Economists also often look at the average propensity to consume (APC), which measures how much income goes to consumption on average. It is calculated as follows:

$$APC = C/Y^d.$$

When disposable income increases, consumption also increases but by a smaller amount. This means that when disposable income increases, people consume a smaller fraction of their income: the average propensity to consume decreases. Using our notation, we are saying that using $C = a + bY^d$, so we can write

$$APC = a/Y^d + b.$$

An increase in disposable income reduces the first term, which also reduces the APC.

The Main Use of This Tool

- Chapter 27

This page titled 31.33: Consumption and Saving is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.34: The Government Budget Constraint 31.34: The Government Budget ConstraintEdit Update titlecancel

Like households, governments are subject to budget constraints. These can be viewed in two ways, either within a single year or across many years.

The Single-Year Government Budget Constraint

In any given year, money flows into the government sector, primarily from the taxes that it imposes on individuals and corporations. We call these government revenues. Money flows out in the form of outlays: government purchases of goods and services and government transfers. The circular flow of income tells us that any difference between government purchases and transfers and government revenues represents a government deficit.

Often, we find it useful to group taxes and transfers together as "net taxes" and separate out government purchases, as in the last line of our definition.

When outflows are less than inflows, then we say a government is running a surplus. In other words, a negative government deficit is the same as a positive government surplus, and a negative government surplus is the same as a positive government deficit.

$$government\ surplus = -government\ deficit.$$

When a government runs a deficit, it must borrow from the financial markets. When a government runs a surplus, these funds flow into the financial markets and are available for firms to borrow. A government surplus is sometimes called government saving.

Intertemporal Government Budget Constraint

Tax and spending decisions at different dates are linked. Although governments can borrow or lend in a given year, a government's total spending over time must be matched with revenues. When a government runs a deficit, it typically borrows to finance it. It borrows by issuing more government debt (government bonds).

To express the intertemporal budget constraint, we introduce a measure of the deficit called the **primary deficit**. The primary deficit is the difference between government outlays, *excluding interest payments on the debt*, and government revenues. The **primary surplus** is minus the primary deficit and is the difference between government revenues and government outlays, excluding interest payments on the debt.

The **intertemporal budget constraint** says that if a government has some existing debt, it must run surpluses in the future so that it can ultimately pay off that debt. Specifically, it is the requirement that

$$current\ debt\ outstanding = discounted\ present\ value\ of\ future\ primary\ surpluses.$$

This condition means that the debt outstanding today must be offset by primary budget surpluses in the future. Because we are adding together flows in the future, we have to use the tool of discounted present value. If, for example, the current stock of debt is zero, then the intertemporal budget constraint says that the discounted present value of future primary surpluses must equal zero.

The stock of debt is linked directly to the government budget deficit. As we noted earlier, when a government runs a budget deficit, it finances the deficit by issuing new debt. The deficit is a *flow* that is matched by a *change in the stock* of government debt:

$$change\ in\ government\ debt(in\ given\ year) = deficit(in\ given\ year).$$

The stock of debt in a given year is equal to the deficit over the previous year plus the stock of debt from the start of the previous year. If there is a government surplus, then the change in the debt is a negative number, so the debt decreases. The total government debt is simply the accumulation of all the previous years' deficits.

When a government borrows, it must pay interest on its debt. These interest payments are counted as part of the deficit (they are included in transfers). If a government wants to balance the budget, then government spending must actually be less than the amount government receives in the form of net taxes (excluding interest).

This presentation of the tool neglects one detail. There is another way in which a government can fund its deficit. As well as issuing government debt, it can print money. More precisely, then, every year,

$$change\ in\ government\ debt = deficit - change\ in\ money\ supply.$$

Written this way, the equation tells us that the part of the deficit that is not financed by printing money results in an increase in the government debt.

More Formally

We often denote government purchases of goods and services by *G* and net tax revenues (tax revenues minus transfers) by *T*. The equation for tax revenues is as follows:

$$T = \tau \times Y,$$

where τ is the tax rate on income and *Y* is real gross domestic product (real GDP). The deficit is given as follows:

$$government\ deficit = G - T = G - \tau \times Y.$$

From this equation, the deficit depends on the following:

- Fiscal policy through the choices of *G* and τ
- The level of economic activity (*Y*)

The Main Uses of This Tool

- Chapter 26
- Chapter 27
- Chapter 28
- Chapter 29
- Chapter 30

This page titled 31.34: The Government Budget Constraint is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.35: The Life-Cycle Model of Consumption 31.35: The Life-Cycle Model of ConsumptionEdit Update titlecancel

The life-cycle model of consumption looks at the lifetime consumption and saving decisions of an individual. The choices made about consumption and saving depend on income earned over an individual's entire lifetime. The model has two key components: the lifetime budget constraint and individual choice given that constraint.

Consider the consumption/saving decision of an individual who expects to work for a known number of years and be retired for a known number of years thereafter. Suppose his disposable income is the same in every working year, and he will also receive an annual retirement income—again the same in every year. According to the life-cycle model of consumption, the individual first calculates the discounted present value (DPV) of lifetime income:

$$DPV\ of\ lifetime\ income = DPV\ of\ income\ from\ working + DPV\ of\ retirement\ income.$$

(If the real interest rate is zero, then the DPV calculation simply involves adding income flows across years.)

We assume the individual wants to consume at the same level in each period of life. This is called consumption smoothing. In the special case of a zero real interest rate, we have the following:

More Formally

Suppose an individual expects to work for a total of *N* years and to be retired for *R* years. Suppose his disposable income is equal to Y^d in every year, and he receives annual retirement income of *Z*. Then lifetime income, assuming a zero real interest rate, is given as follows:

$$lifetime\ income = NY^d + RZ.$$

If we suppose that he wants to have perfectly smooth consumption, equal to *C* in each year, then his total lifetime consumption will be

$$C \times (N + R).$$

The lifetime budget constraint says that lifetime consumption equals lifetime income:

$$C \times (N + R) = NY^d + RZ.$$

To obtain his consumption, we simply divide this equation by the number of years he is going to live (*N* + *R*):

Provided that income during working years is greater than income in retirement years, the individual will save during his working years and dissave during retirement.

If the real interest rate is not equal to zero, then the basic idea is the same—an individual smooths consumption based on a lifetime budget constraint—but the calculations are more complicated. Specifically, the lifetime budget constraint must be written in terms of the discounted present values of income and consumption.

The Main Uses of This Tool

- Chapter 22
- Chapter 28
- Chapter 29

This page titled 31.35: The Life-Cycle Model of Consumption is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.36: Aggregate Supply and Aggregate Demand 31.36: Aggregate Supply and Aggregate DemandEdit Update titlecancel

The aggregate supply and aggregate demand (ASAD) model is presented here. To understand the ASAD model, we need to explain both aggregate demand and aggregate supply and then the determination of prices and output.

The aggregate demand curve tells us the level of expenditure in an economy for a given price level. It has a negative slope: the demand for real gross domestic product (real GDP) decreases when the price level increases. The downward sloping aggregate demand curve does *not* follow from the microeconomic "law of demand." As the price level increases, all prices in an economy increase together. The substitution of expensive goods for cheap goods, which underlies the law of demand, does not occur in the aggregate economy.

Instead, the downward sloping demand curve comes from other forces. First, as prices rise, the real value of nominal wealth falls, and this leads to a fall in household spending. Second, as prices rise today relative to future prices, households are induced to postpone consumption. Finally, a higher price level can lead to a higher interest rate through the response of monetary policy. All these factors together imply that higher prices lead to lower overall demand for real GDP.

Aggregate supply is equal to potential output at all prices. Potential output is determined by the available technology, physical capital, and labor force and is unaffected by the price level. Thus the aggregate supply curve is vertical. In contrast to a firm's supply curve, as the price level increases, all prices in an economy increase. This includes the prices of inputs, such as labor, into the production process. Since no relative prices change when the price level increases, firms are not induced to change the quantity they supply. Thus aggregate supply is vertical.

The determination of prices and output depends on the horizon: the long run or the short run. In the long run, real GDP equals potential GDP, and real GDP also equals aggregate expenditure. This means that, in the long run, the price level must be at the point where aggregate demand and aggregate supply meet. This is shown in Figure 31.36.26.

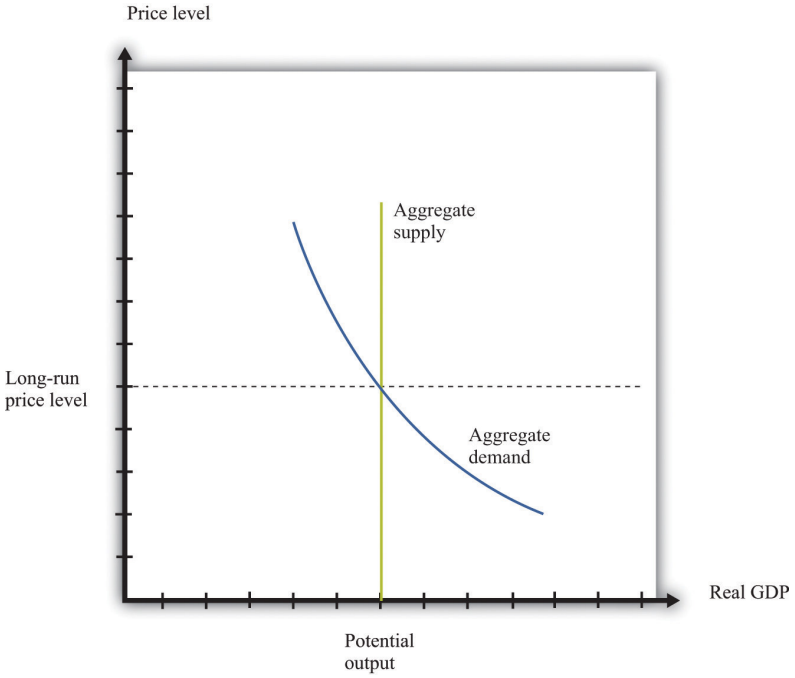


Figure 31.36.26: Aggregate Supply and Aggregate Demand in the Long Run

In the short run, output is determined by aggregate demand at the existing price level. Prices need not be at their long-run equilibrium levels. If they are not, then output will not equal potential output. This is shown in Figure 31.36.27.

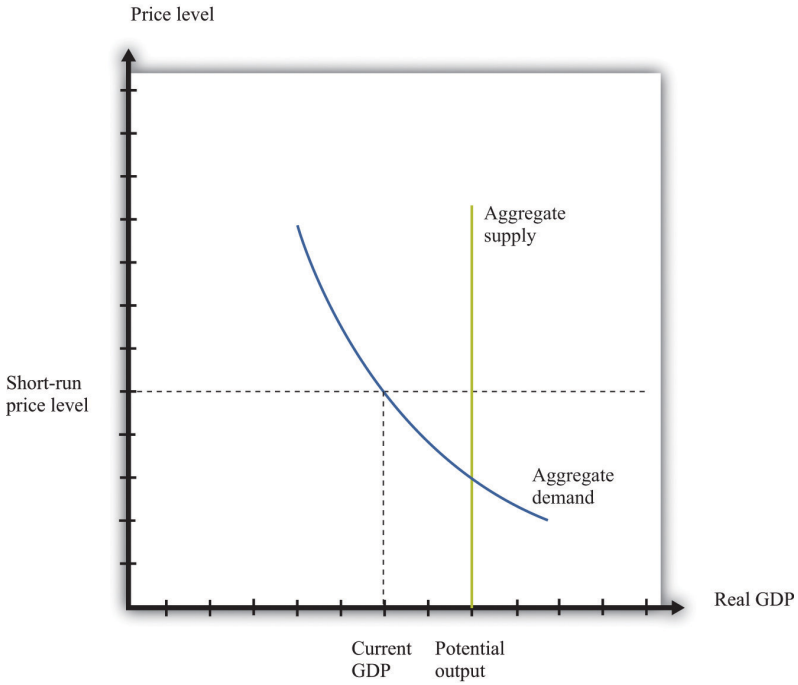


Figure 31.36.27: Aggregate Supply and Aggregate Demand in the Short Run

The short-run price level is indicated on the vertical axis. The level of output is determined by aggregate demand at that price level. As prices are greater than the long-run equilibrium level of prices, output is below potential output. The price level adjusts over time to its long-run level, according to the price-adjustment equation.

The Main Uses of This Tool

We do not explicitly use this tool in our chapter presentations. However, the tool can be used to support the discussions in the following chapters.

- Chapter 22
- Chapter 25
- Chapter 26
- Chapter 27

This page titled 31.36: Aggregate Supply and Aggregate Demand is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

31.37: The IS-LM Model 31.37: The IS-LM ModelEdit Update titlecancel

The IS-LM model provides another way of looking at the determination of the level of short-run real gross domestic product (real GDP) in the economy. Like the aggregate expenditure model, it takes the price level as fixed. But whereas that model takes the interest rate as exogenous—specifically, a change in the interest rate results in a change in autonomous spending—the IS-LM model treats the interest rate as an endogenous variable.

The basis of the IS-LM model is an analysis of the money market and an analysis of the goods market, which together determine the equilibrium levels of interest rates and output in the economy, given prices. The model finds combinations of interest rates and output (GDP) such that the money market is in equilibrium. This creates the LM curve. The model also finds combinations of interest rates and output such that the goods market is in equilibrium. This creates the IS curve. The equilibrium is the interest rate and output combination that is on both the IS and the LM curves.

LM Curve

The LM curve represents the combinations of the interest rate and income such that money supply and money demand are equal. The demand for money comes from households, firms, and governments that use money as a means of exchange and a store of value. The law of demand holds: as the interest rate increases, the quantity of money demanded decreases because the interest rate represents an opportunity cost of holding money. When interest rates are higher, in other words, money is less effective as a store of value.

Money demand increases when output rises because money also serves as a medium of exchange. When output is larger, people have more income and so want to hold more money for their transactions.

The supply of money is chosen by the monetary authority and is independent of the interest rate. Thus it is drawn as a vertical line. The equilibrium in the money market is shown in Figure 31.37.28. When the money supply is chosen by the monetary authority, the interest rate is the price that brings the market into equilibrium. Sometimes, in some countries, central banks target the money supply. Alternatively, central banks may choose to target the interest rate. (This was the case we considered in Chapter 25.) Figure 31.37.28 applies in either case: if the monetary authority targets the interest rate, then the money market tells us what the level of the money supply must be.

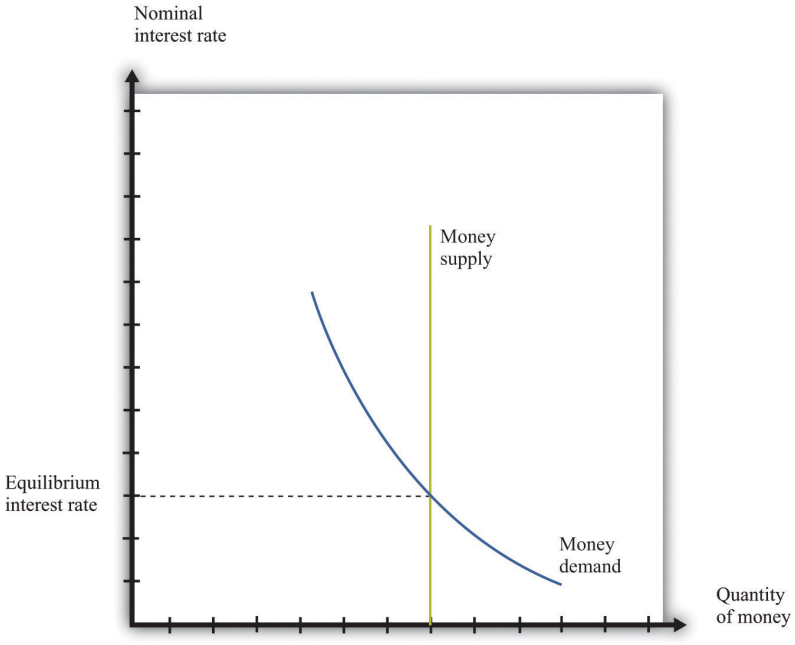


Figure 31.37.28: Money Market Equilibrium

To trace out the LM curve, we look at what happens to the interest rate when the level of output in the economy changes and the supply of money is held fixed. Figure 31.37.29 shows the money market equilibrium at two different levels of real GDP. At the higher level of income, money demand is shifted to the right; the interest rate increases to ensure that money demand equals money

supply. Thus the LM curve is upward sloping: higher real GDP is associated with higher interest rates. At each point along the LM curve, money supply equals money demand.

We have not yet been specific about whether we are talking about nominal interest rates or real interest rates. In fact, it is the nominal interest rate that represents the opportunity cost of holding money. When we draw the LM curve, however, we put the real interest rate on the axis, as shown in Figure 31.37.30. The simplest way to think about this is to suppose that we are considering an economy where the inflation rate is zero. In this case, by the Fisher equation, the nominal and real interest rates are the same. In a more complete analysis, we can incorporate inflation by noting that changes in the inflation rate will shift the LM curve. Changes in the money supply also shift the LM curve.

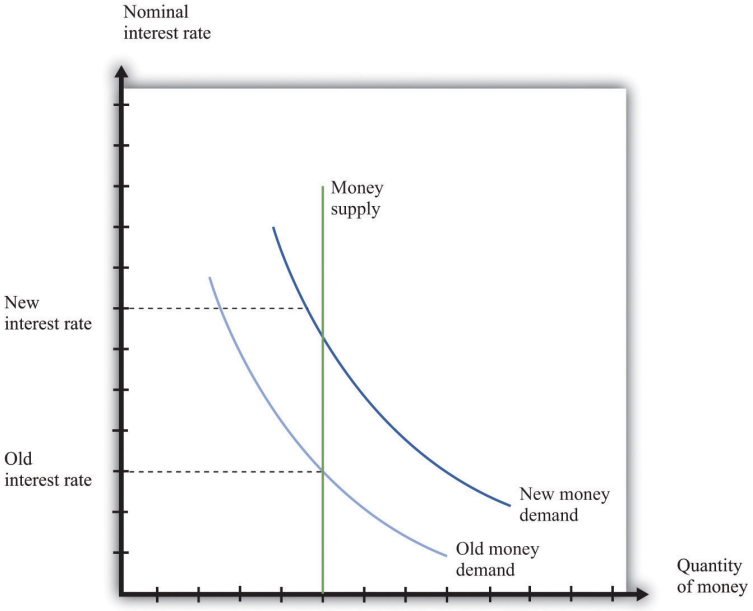


Figure 31.37.29: A Change in Income

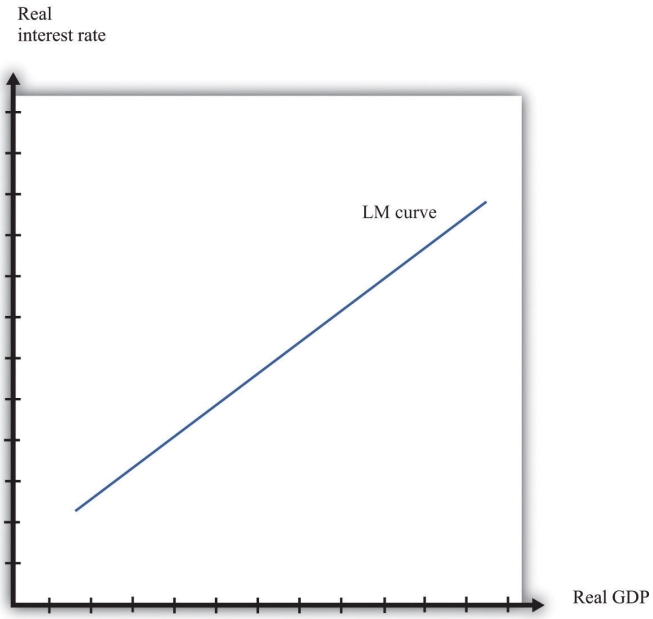


Figure 31.37.30: The LM Curve

IS Curve

The IS curve relates the level of real GDP and the real interest rate. It incorporates both the dependence of spending on the real interest rate and the fact that, in the short run, real GDP equals spending. The IS curve is shown in Figure 31.37.29. We label the

horizontal axis “real GDP” since, in the short run, real GDP is determined by aggregate spending. The IS curve is downward sloping: as the real interest rate increases, the level of spending decreases.

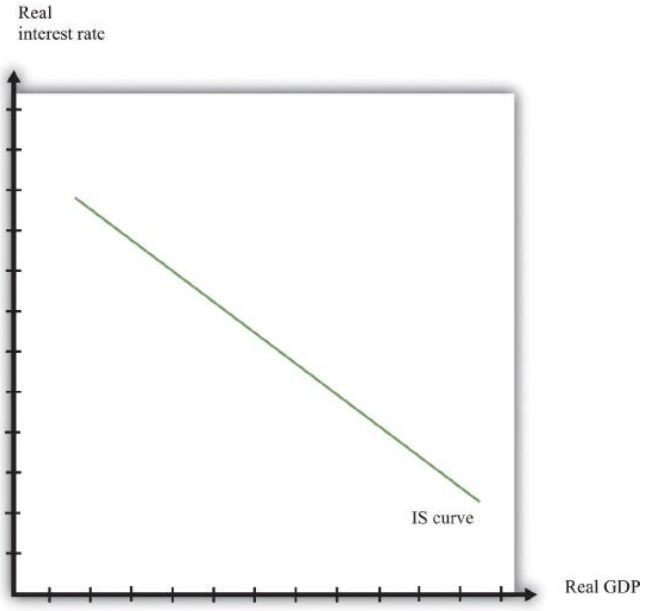


Figure 31.37.31: The IS Curve

In fact, we derived the IS curve in Chapter 25. The dependence of spending on real interest rates comes partly from investment. As the real interest rate increases, spending by firms on new capital and spending by households on new housing decreases. Consumption also depends on the real interest rate: spending by households on durable goods decreases as the real interest rate increases.

The connection between spending and real GDP comes from the aggregate expenditure model. Given a particular level of the interest rate, the aggregate expenditure model determines the level of real GDP. Now suppose the interest rate increases. This reduces those components of spending that depend on the interest rate. In the aggregate expenditure framework, this is a reduction in autonomous spending. The equilibrium level of output decreases. Thus the IS curve slopes downwards: higher interest rates are associated with lower real GDP.

Equilibrium

Combining the discussion of the LM and the IS curves will generate equilibrium levels of interest rates and output. Note that both relationships are combinations of interest rates and output. Solving these two equations jointly determines the equilibrium. This is shown graphically in Figure 31.37.32. This just combines the LM curve from Figure 31.37.30 and the IS curve from Figure 31.37.31. The crossing of these two curves is the combination of the interest rate and real GDP, denoted (r^*,Y^*), such that both the money market and the goods market are in equilibrium.

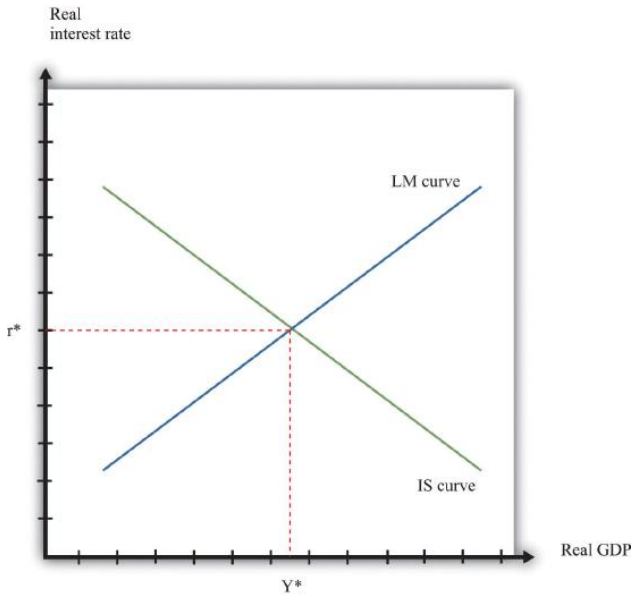


Figure 31.37.32. Equilibrium in the IS-LM Model.

Comparative Statics

Comparative statics results for this model illustrate how changes in exogenous factors influence the equilibrium levels of interest rates and output. For this model, there are two key exogenous factors: the level of autonomous spending (excluding any spending affected by interest rates) and the real money supply. We can study how changes in these factors influence the equilibrium levels of output and interest rates both graphically and algebraically.

Variations in the level of autonomous spending will lead to a shift in the IS curve, as shown in Figure 31.37.33. If autonomous spending increases, then the IS curve shifts out. The output level of the economy will increase. Interest rates rise as we move along the LM curve, ensuring money market equilibrium. One source of variations in autonomous spending is fiscal policy. Autonomous spending includes government spending (G). Thus an increase in G leads to an increase in output and interest rates as shown in Figure 31.37.33.

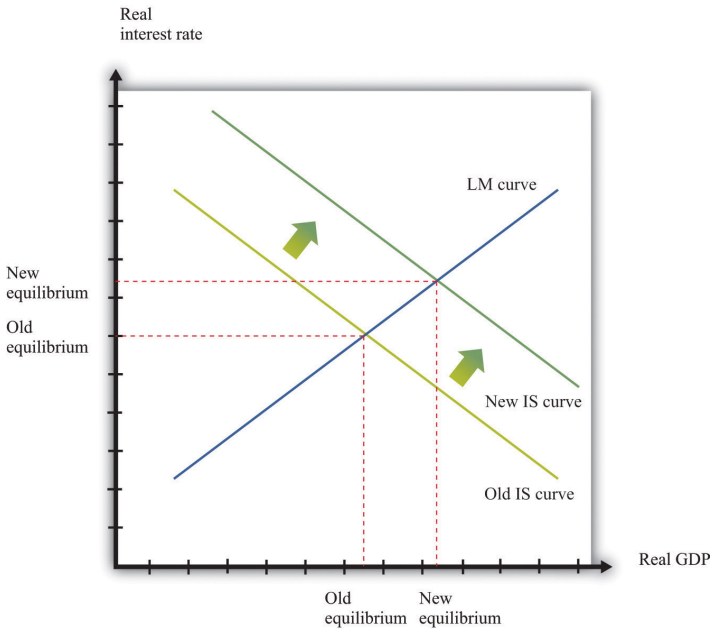


Figure 31.37.33: A Shift in the IS Curve

Variations in the real money supply shift the LM curve, as shown in Figure 31.37.34. If the money supply decreases, then the LM curve shifts in. This leads to a higher real interest rate and lower output as the LM curve shifts along the fixed IS curve.

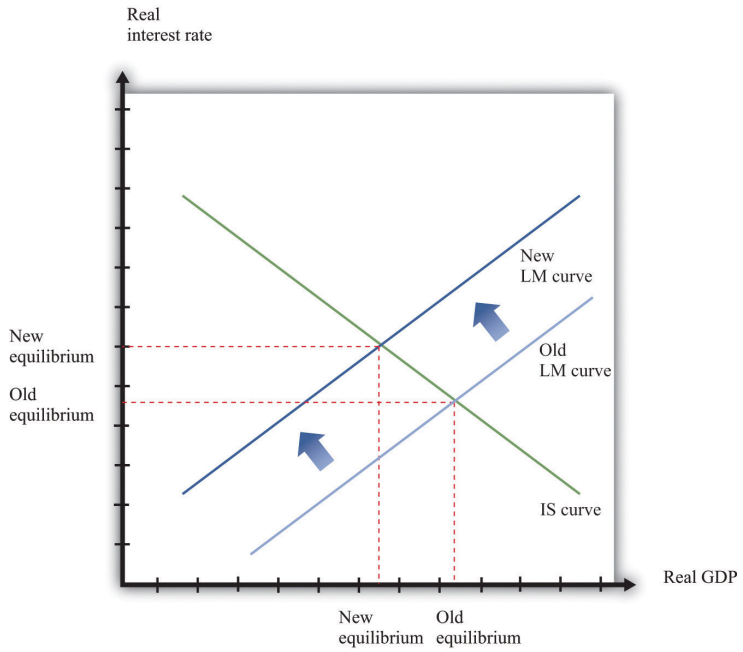


Figure 31.37.34: A Shift in the LM Curve

More Formally

We can represent the LM and IS curves algebraically.

LM Curve

Let L(*Y*,*r*) represent real money demand at a level of real GDP of *Y* and a real interest rate of *r*. (When we say "real" money demand, we mean that, as usual, we have deflated by the price level.) For simplicity, suppose that the inflation rate is zero, so the real interest rate is the opportunity cost of holding money.If we wanted to include inflation in our analysis, we could write the real demand for money as L(*Y*, *r* + π), where π is the inflation rate. Assume that real money demand takes a particular form:

\(L(Y,r) = L_{0} + L_{1}Y– L_{2}r. \nonumber \]

In this equation, L_0, L_1, and L_2 are all positive constants. Real money demand is increasing in income and decreasing in the interest rate. Letting *M/P* be the real stock of money in the economy, then money market equilibrium requires

$$M/P = L_0 + L_1Y–L_2r.$$

Given a level of real GDP and the real stock of money, this equation can be used to solve for the interest rate such that money supply and money demand are equal. This is given by

$$r = (1/L_2)[L_0 + L_1Y–M/P].$$

From this equation we learn that an increase in the real stock of money lowers the interest rate, given the level of real GDP. Further, an increase in the level of real GDP increases the interest rate, given the stock of money. This is another way of saying that the LM curve is upward sloping.

IS Curve

Recall the two equations from the aggregate expenditure model:

$$Y = E$$

and

$$E = E_0(r) + \betaY.$$

Here we have shown explicitly that the level of autonomous spending depends on the real interest rate *r*.

We can solve the two equations to find the values of *E* and *Y* that are consistent with both equations. We find

Given a level of the real interest rate, we solve for the level of autonomous spending (using the dependence of consumption and investment on the real interest rate) and then use this equation to find the level of output.

Here is an example. Suppose that

$C = 100 + 0.6Y$,

$I = 400 - 5r$,

$G = 300$,

and

$NX = 200 - 0.1Y$,

where C is consumption, I is investment, G is government purchases, and NX is net exports. First group the components of spending as follows:

$$C + I + G + NX = (100 + 400 - 5r + 300 + 200) + (0.6Y - 0.1Y)$$

Adding together the first group of terms, we find autonomous spending:

\[E_{0} = 100 + 400 + 300 + 200 − 5r = 1000 − 5r. \nonumber \]

Adding the coefficients on the income terms, we find the marginal propensity to spend:

$$\beta = 0.6 - 0.1 = 0.5.$$

Using β = 0.5, we calculate the multiplier:

We then calculate real GDP, given the real interest rate:

$$Y = 2 \times (1000 - 5r) = 2000 - 10r.$$

Equilibrium

Combining the discussion of the LM and the IS curves will generate equilibrium levels of interest rates and output. Note that both relationships are combinations of interest rates and output. Solving these two equations jointly determines the equilibrium.

Algebraically, we have an equation for the LM curve:

$$r = (1/L_2)\left[L_0 + L_1 Y - M/P\right]$$

And we have an equation for the IS curve:

$$Y = mE_0(r),$$

where we let $m = (1/(1-\times))$ denote the multiplier. If we assume that the dependence of spending in the interest rate is linear, so that \[E_{0}(r)=e_{0}-e_{1} r \nonumber \], then the equation for the IS curve is

$$Y = m\left(e_0 - e_1 r\right),$$

To solve the IS and LM curves simultaneously, we substitute Y from the IS curve into the LM curve to get

$$r = (1/L_2)\left[L_0 + L_1 m\left(e_0 - e_1 r\right) - M/P\right]$$

Solving this for r we get

$$r = A_r - B_r M/P$$

where both A_r and B_r are constants, with $A_r = \left(L_0 + L_1 m e_0\right)/\left(L_1 m e_1 + L_2\right)$ and $B_r = 1/\left(L_1 m e_1 + L_2\right)$. This equation gives us the equilibrium level of the real interest rate given the level of autonomous spending, summarized by e_0, and the real stock of money, summarized by *M/P*.

To find the equilibrium level of output, we substitute this equation for *r* back into the equation for the IS curve. This gives us

$$Y = A_y + B_y(M/P)$$

where both A_y and B_y are constants, with $A_y = m\left(e_0 - e_1 A_r\right)$ and $B_y = m e_1 B_1$. This equation gives us the equilibrium level of output given the level of autonomous spending, summarized by e_0, and the real stock of money, summarized by *M/P*.

Algebraically, we can use the equations to determine the magnitude of the responses of interest rates and output to exogenous changes. An increase in the autonomous spending, e_0, will increase both A_r and A_y, implying that both the interest rate and output increase.To see that A_y increases with e_0 requires a bit more algebra. An increase in the real money stock will reduce interest rates by B_r and increase output by B_y. A key part of monetary policy is the sensitivity of spending to the interest rate, given by e_1. The more sensitive is spending to the interest rate, the larger is e_1 and therefore the larger is B_y.

The Main Uses of This Tool

We do not explicitly use this tool in our chapter presentations. However, the tool can be used to support the discussions in the following chapters.

- Chapter 24
- Chapter 25
- Chapter 26
- Chapter 29

This page titled 31.37: The IS-LM Model is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous via source content that was edited to the style and standards of the LibreTexts platform.

Index IndexEdit Update titlecancel

A

absolute advantage
- 23.3: Job and Worker Flows

C

comparative advantage
- 23.3: Job and Worker Flows

comparative statics
- 31.17: Comparative Statics

credit market
- 8.5: Three Important Markets
- 29.5: The Costs of Deficits
- 31.7: The Credit Market

crowding out
- 29.5: The Costs of Deficits

D

disposable income
- 4.2: Individual Decision Making- How You Spend Your Income

L

loan market
- 8.5: Three Important Markets

N

national savings
- 29.5: The Costs of Deficits

R

real interest rate
- 29.5: The Costs of Deficits

S

Search theory
- 23.3: Job and Worker Flows

U

ultimatum game
- 13.4: Distributive Justice

V

Vickrey Auction
- 6.3: eBay

Index IndexEdit Update titlecancel

A

absolute advantage
- 23.3: Job and Worker Flows

C

comparative advantage
- 23.3: Job and Worker Flows

comparative statics
- 31.17: Comparative Statics

credit market
- 8.5: Three Important Markets
- 29.5: The Costs of Deficits
- 31.7: The Credit Market

crowding out
- 29.5: The Costs of Deficits

D

disposable income
- 4.2: Individual Decision Making- How You Spend Your Income

L

loan market
- 8.5: Three Important Markets

N

national savings
- 29.5: The Costs of Deficits

R

real interest rate
- 29.5: The Costs of Deficits

S

Search theory
- 23.3: Job and Worker Flows

U

ultimatum game
- 13.4: Distributive Justice

V

Vickrey Auction
- 6.3: eBay

Glossary GlossaryEdit Update titlecancel

Sample Word 1 | Sample Definition 1

Detailed Licensing Detailed LicensingEdit Update titlecancel

Loading...

https://socialsci.libretexts.org/@go/page/178615