openstax™

Principles of

Macro-economics

2e.

Principles of Macroeconomics 2e

SENIOR CONTRIBUTING AUTHORS

STEVEN A. GREENLAW, UNIVERSITY OF MARY WASHINGTON

DAVID SHAPIRO, PENNSYLVANIA STATE UNIVERSITY

Based on the 2nd edition of *Principles of Economics, Economics and the Economy, 2e* by Timothy Taylor, published in 2011.

openstax™

OpenStax
Rice University
6100 Main Street MS-375
Houston, Texas 77005

To learn more about OpenStax, visit https://openstax.org.
Individual print copies and bulk orders can be purchased through our website.

Portions of the Demand and Supply chapter were derived from "Why It Matters: Government in Action by Steve Greenlaw and Lumen Learning." Located at: https://courses.lumenlearning.com/waymakermacroxmasterfall2016/chapter/why-it-matters-government-action/. License: CC BY 4.0.

PRINT BOOK ISBN-10	**1-947172-38-7**
PRINT BOOK ISBN-13	**978-1-947172-38-8**
PDF VERSION ISBN-10	**1-947172-39-5**
PDF VERSION ISBN-13	**978-1-947172-39-5**
ENHANCED TEXTBOOK ISBN-10	**1-947172-50-6**
ENHANCED TEXTBOOK ISBN-13	**978-1-947172-50-0**
Revision Number	**MA2-2017-001-(03/18)-LC**
Original Publication Year	**2017**

OpenStax

OpenStax provides free, peer-reviewed, openly licensed textbooks for introductory college and Advanced Placement® courses and low-cost, personalized courseware that helps students learn. A nonprofit ed tech initiative based at Rice University, we're committed to helping students access the tools they need to complete their courses and meet their educational goals.

Rice University

OpenStax, OpenStax CNX, and OpenStax Tutor are initiatives of Rice University. As a leading research university with a distinctive commitment to undergraduate education, Rice University aspires to path-breaking research, unsurpassed teaching, and contributions to the betterment of our world. It seeks to fulfill this mission by cultivating a diverse community of learning and discovery that produces leaders across the spectrum of human endeavor.

Foundation Support

OpenStax is grateful for the tremendous support of our sponsors. Without their strong engagement, the goal of free access to high-quality textbooks would remain just a dream.

Laura and John Arnold Foundation (LJAF) actively seeks opportunities to invest in organizations and thought leaders that have a sincere interest in implementing fundamental changes that not only yield immediate gains, but also repair broken systems for future generations. LJAF currently focuses its strategic investments on education, criminal justice, research integrity, and public accountability.

The William and Flora Hewlett Foundation has been making grants since 1967 to help solve social and environmental problems at home and around the world. The Foundation concentrates its resources on activities in education, the environment, global development and population, performing arts, and philanthropy, and makes grants to support disadvantaged communities in the San Francisco Bay Area.

Calvin K. Kazanjian was the founder and president of Peter Paul (Almond Joy), Inc. He firmly believed that the more people understood about basic economics the happier and more prosperous they would be. Accordingly, he established the Calvin K. Kazanjian Economics Foundation Inc, in 1949 as a philanthropic, nonpolitical educational organization to support efforts that enhanced economic understanding.

Guided by the belief that every life has equal value, the Bill & Melinda Gates Foundation works to help all people lead healthy, productive lives. In developing countries, it focuses on improving people's health with vaccines and other life-saving tools and giving them the chance to lift themselves out of hunger and extreme poverty. In the United States, it seeks to significantly improve education so that all young people have the opportunity to reach their full potential. Based in Seattle, Washington, the foundation is led by CEO Jeff Raikes and Co-chair William H. Gates Sr., under the direction of Bill and Melinda Gates and Warren Buffett.

The Maxfield Foundation supports projects with potential for high impact in science, education, sustainability, and other areas of social importance.

Our mission at The Michelson 20MM Foundation is to grow access and success by eliminating unnecessary hurdles to affordability. We support the creation, sharing, and proliferation of more effective, more affordable educational content by leveraging disruptive technologies, open educational resources, and new models for collaboration between for-profit, nonprofit, and public entities.

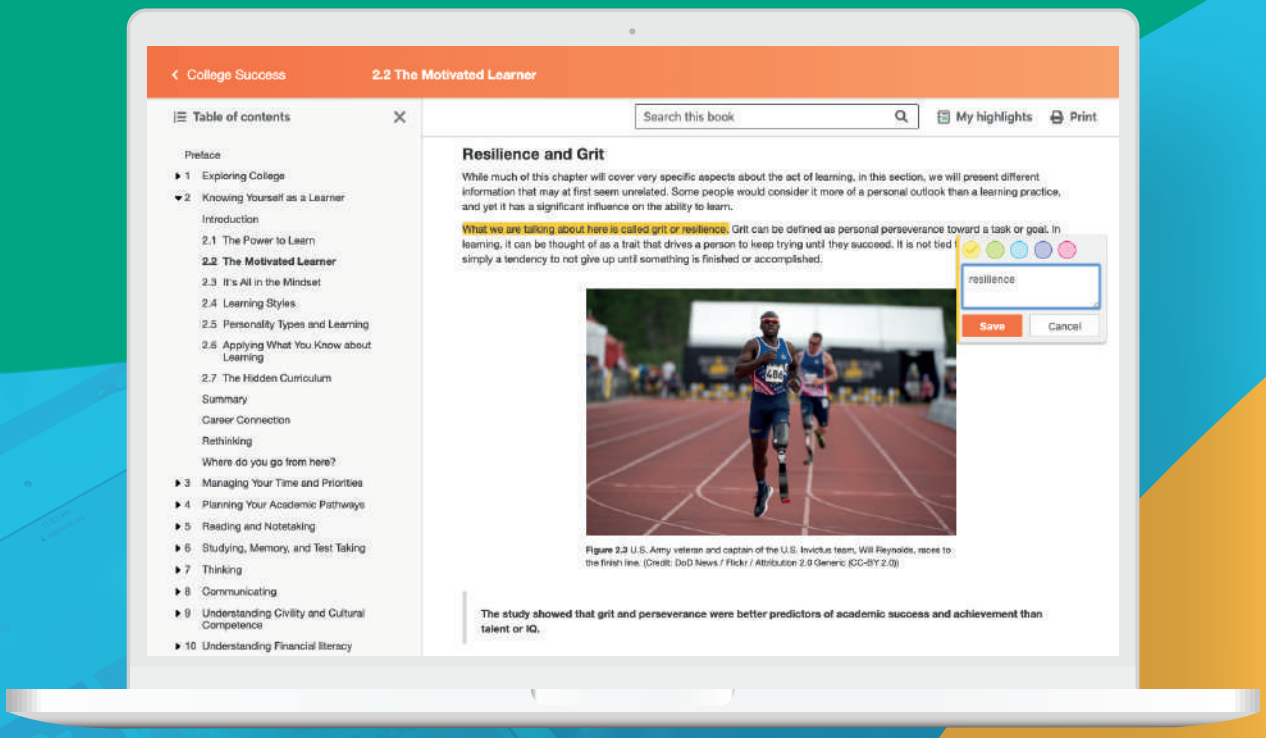
The Bill and Stephanie Sick Fund supports innovative projects in the areas of Education, Art, Science and Engineering.

openstax™

Study where you want, what you want, when you want.

When you access College Success in our web view, you can use our new online highlighting and note-taking features to create your own study guides.

Our books are free and flexible, forever.
Get started at openstax.org/details/books/principles-macroeconomics-2e



Access. The future of education.
openstax.org

RICE

Table of Contents

PREFACE

Welcome to *Principles of Macroeconomics 2e* (2nd Edition), an OpenStax resource. This textbook was written to increase student access to high-quality learning materials, maintaining highest standards of academic rigor at little to no cost.

About OpenStax

OpenStax is a nonprofit based at Rice University, and it's our mission to improve student access to education. Our first openly licensed college textbook was published in 2012, and our library has since scaled to over 25 books for college and AP® courses used by hundreds of thousands of students. OpenStax Tutor, our low-cost personalized learning tool, is being used in college courses throughout the country. Through our partnerships with philanthropic foundations and our alliance with other educational resource organizations, OpenStax is breaking down the most common barriers to learning and empowering students and instructors to succeed.

About OpenStax resources

Customization

Principles of Macroeconomics 2e is licensed under a Creative Commons Attribution 4.0 International (CC BY) license, which means that you can distribute, remix, and build upon the content, as long as you provide attribution to OpenStax and its content contributors.

Because our books are openly licensed, you are free to use the entire book or pick and choose the sections that are most relevant to the needs of your course. Feel free to remix the content by assigning your students certain chapters and sections in your syllabus, in the order that you prefer. You can even provide a direct link in your syllabus to the sections in the web view of your book.

Instructors also have the option of creating a customized version of their OpenStax book. The custom version can be made available to students in low-cost print or digital form through their campus bookstore. Visit the Instructor Resources section of your book page on OpenStax.org for more information.

Errata

All OpenStax textbooks undergo a rigorous review process. However, like any professional-grade textbook, errors sometimes occur. Since our books are web based, we can make updates periodically when deemed pedagogically necessary. If you have a correction to suggest, submit it through the link on your book page on OpenStax.org. Subject matter experts review all errata suggestions. OpenStax is committed to remaining transparent about all updates, so you will also find a list of past errata changes on your book page on OpenStax.org.

Format

You can access this textbook for free in web view or PDF through OpenStax.org, and for a low cost in print.

About *Principles of Macroeconomics 2e*

Principles of Macroeconomics 2e (2nd edition) covers the scope and sequence requirements of most introductory macroeconomics courses. The text includes many current examples, which are handled in a politically equitable way. The outcome is a balanced approach to both Keynesian and classical views, and to the theory and application of economics concepts. The second edition has been thoroughly revised to increase clarity, update data and current event impacts, and incorporate the feedback from many reviewers and adopters.

Coverage and scope

To develop the first edition of *Principles of Macroeconomics*, we acquired the rights to Timothy Taylor's *Principles of Economics* and solicited ideas from economics instructors at all levels of higher education, from community colleges to PhD-granting universities. For the second edition, we received even more expansive and actionable feedback from hundreds of adopters who had used the book for several academic terms. These knowledgeable instructors informed the pedagogical courses, learning objective development and fulfillment, and the chapter arrangements. Faculty who taught from the material provided critical and detailed commentary.

The result is a book that covers the breadth of economics topics and also provides the necessary depth to ensure the course is manageable for instructors and students alike. We strove to balance theory and application, as well as the amount of calculation and mathematical examples.

The book is organized into seven main parts:

What is Economics? The first two chapters introduce students to the study of economics with a focus on making choices in a world of scarce resources.

Supply and Demand, Chapters 3 and 4, introduces and explains the first analytical model in economics: supply, demand, and equilibrium, before showing applications in the markets for labor and finance.

Elasticity and Price, Chapter 5, introduces and explains elasticity and price, two key concepts in economics.

The Macroeconomic Perspective and Goals, Chapters 6 through 10, introduces a number of key concepts in macro: economic growth, unemployment and inflation, and international trade and capital flows.

A Framework for Macroeconomic Analysis, Chapters 11 through 13, introduces the principal analytic model in macro, namely the aggregate demand/aggregate supply Model. The model is then applied to the Keynesian and Neoclassical perspectives. The expenditure-output model is fully explained in a stand-alone appendix.

Monetary and Fiscal Policy, Chapters 14 through 18, explains the role of money and the banking system, as well as monetary policy and financial regulation. Then the discussion switches to government deficits and fiscal policy.

International Economics, Chapters 19 through 21, the final part of the text, introduces the international dimensions of economics, including international trade and protectionism.

Alternate Sequencing

Principles of Macroeconomics 2e was conceived and written to fit a particular topical sequence, but it can be used flexibly to accommodate other course structures. One such potential structure, which will fit reasonably well with the textbook content, is provided below. Please consider, however, that the chapters were not written to be completely independent, and that the proposed alternate sequence should be carefully considered for student preparation and textual consistency.

Chapter 1 Welcome to Economics!
Chapter 2 Choice in a World of Scarcity
Chapter 3 Demand and Supply
Chapter 4 Labor and Financial Markets
Chapter 5 Elasticity
Chapter 20 International Trade
Chapter 6 The Macroeconomic Perspective
Chapter 7 Economic Growth
Chapter 8 Unemployment
Chapter 9 Inflation
Chapter 10 The International Trade and Capital Flows
Chapter 12 The Keynesian Perspective
Chapter 13 The Neoclassical Perspective
Chapter 14 Money and Banking
Chapter 15 Monetary Policy and Bank Regulation
Chapter 16 Exchange Rates and International Capital Flows
Chapter 17 Government Budgets and Fiscal Policy
Chapter 11 The Aggregate Demand/Aggregate Supply Model
Chapter 18 The Impacts of Government Borrowing
Chapter 19 Macroeconomic Policy Around the World
Chapter 21 Globalization and Protectionism

Appendix A The Use of Mathematics in Principles of Economics
Appendix B The Expenditure-Output Model

Changes to the second edition

OpenStax only undertakes revisions when significant modifications to a text are necessary. In the case of *Principles of Macroeconomics*, we received a wealth of constructive feedback. Many of the book's users felt that consequential movement in economic data, coupled with the impacts of national and global events, warranted a full revision. We also took advantage of the opportunity to improve the writing and sequencing of the text, as well as many of the calculation examples. The major changes are summarized below.

Augmented explanations in chapters one through four provide a more comprehensive and informative foundation for the book.

A clearer explanation, using a numerical example, has been given for finding the utility maximizing combination of goods and services a consumer should choose.

The Theory of Production has been added to the chapter on costs & industry structure.

A more complete treatment has been given to labor markets, including the theories of competitive and monopsonistic labor markets, and bilateral monopoly; and the labor markets chapter and the poverty and economic inequality chapter have been resequenced.

Substantial revisions to the AD/AS model in chapters 11-13 present the core concepts of macroeconomics in a clearer, more dynamic manner.

Case studies and examples have been revised and, in some cases, replaced to provide more relevant and useful information for students.

Economic data, tables, and graphs, as well as discussion and analysis around that data, have been thoroughly updated.

Wherever possible, data from the Federal Reserve Economic Database (FRED) was included and referenced. In most of these uses, links to the direct source of the FRED data are provided, and students are encouraged to explore the information and the overall FRED resources more thoroughly.

Additional updates and revisions appear throughout the book. They reflect changes to economic realities and policies regarding international trade, taxation, insurance, and other topics. For issues that may change in the months or years following the textbook's publication, the authors often provided a more open-ended explanation, but we will update the text annually to address further changes.

The revision of *Principles of Macroeconomics* was undertaken by Steven Greenlaw (University of Mary Washington) and David Shapiro (Pennsylvania State University), with significant input by lead reviewer Daniel MacDonald (California State University, San Bernardino).

Pedagogical foundation

Throughout *Principles of Macroeconomics 2e*, you will find features that engage the students in economic inquiry and support their learning. Our features include:

Bring It Home: This added feature is a brief case study, specific to each chapter, which connects the chapter's main topic to the real word. It is broken up into two parts: the first at the beginning of the chapter (in the intro module) and the second at chapter's end, when students have learned what's necessary to understand the case and "bring home" the chapter's core concepts.

Work It Out: This added feature asks students to work through a generally analytical or computational problem, and guides them step by step to find out how its solution is derived.

Clear It Up: This boxed feature, which includes pre-existing features from Taylor's text, addresses common student misconceptions about the content. Clear It Ups are usually deeper explanations of something in the main body of the text. Each CIU starts with a question. The rest of the feature explains the answer.

Link It Up: This added feature is a very brief introduction to a website that is pertinent to students' understanding and enjoyment of the topic at hand.

Questions for each level of learning

Principles of Macroeconomics 2e offers four types of end-of-module questions for students:

Self-Checks are analytical self-assessment questions that appear at the end of each module. They "click to

reveal" an answer in the web view so students can check their understanding before moving on to the next module. Self-Check questions are not simple look-up questions. They push the student to think beyond what is said in the text. Self-Check questions are designed for formative (rather than summative) assessment. The questions and answers are explained so that students feel like they are being walked through the problem.

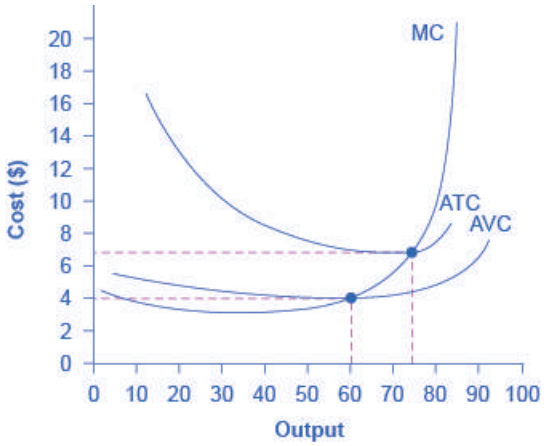
Review Questions have been retained from Taylor's version, and are simple recall questions from the chapter in open-response format (not multiple choice or true/false). The answers can be looked up in the text.

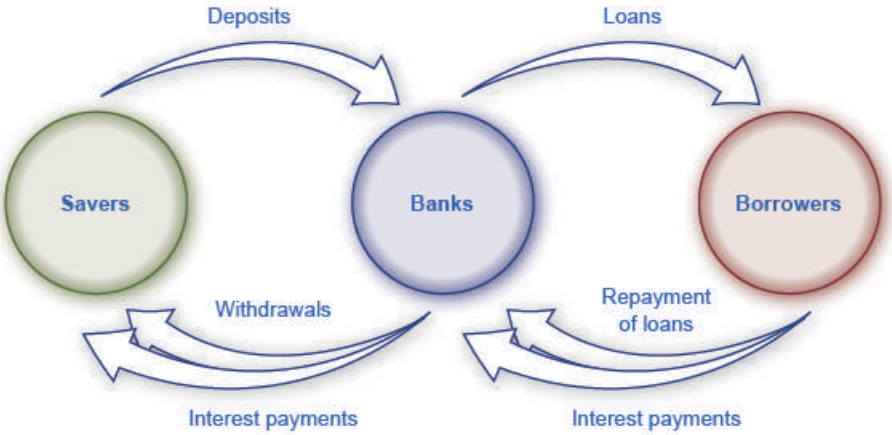
Critical Thinking Questions are new higher-level, conceptual questions that ask students to *demonstrate their understanding by applying* what they have learned in different contexts. They ask for outside-the-box thinking, for *reasoning* about the concepts. They push the student to places they wouldn't have thought of going themselves.

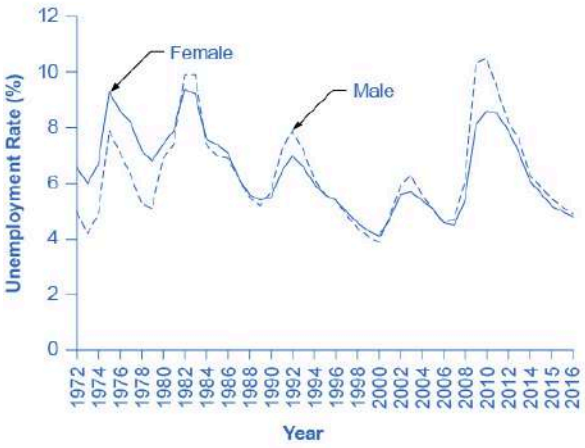
Problems are exercises that give students additional practice working with the analytic and computational concepts in the module.

Updated art

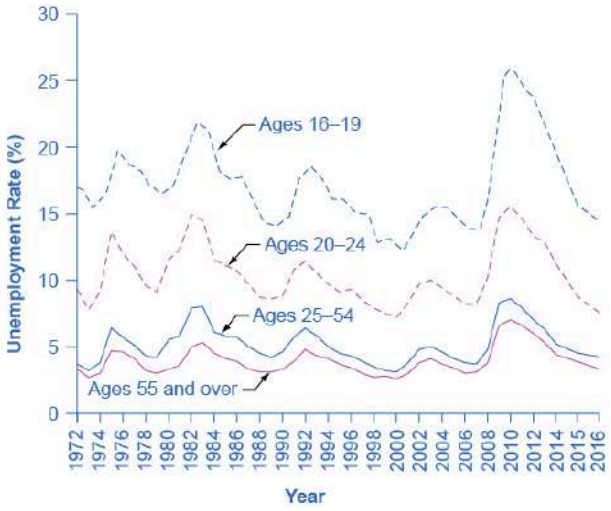
Principles of Macroeconomics 2e includes an updated art program to better inform today's student, providing the latest data on covered topics.



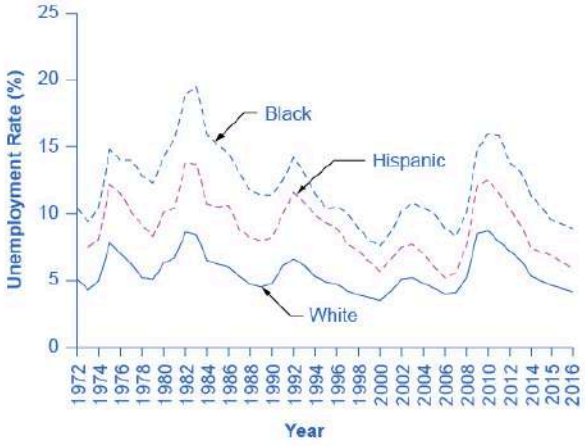




(a) Unemployment rates by gender



(b) Unemployment rates for women, by age



(c) Unemployment rates by race and ethnicity

Additional resources

Student and instructor resources

We've compiled additional resources for both students and instructors, including Getting Started Guides, an instructor solution manual, test bank, and PowerPoint slides. Instructor resources require a verified instructor account, which you can apply for when you log in or create your account on OpenStax.org. Take advantage of these resources to supplement your OpenStax book.

Community Hubs

OpenStax partners with the Institute for the Study of Knowledge Management in Education (ISKME) to offer Community Hubs on OER Commons – a platform for instructors to share community-created resources that support OpenStax books, free of charge. Through our Community Hubs, instructors can upload their own materials or download resources to use in their own courses, including additional ancillaries, teaching material, multimedia, and relevant course content. We encourage instructors to join the hubs for the subjects most relevant to your teaching and

research as an opportunity both to enrich your courses and to engage with other faculty.

To reach the Community Hubs, visit www.oercommons.org/hubs/OpenStax.

Technology partners

As allies in making high-quality learning materials accessible, our technology partners offer optional low-cost tools that are integrated with OpenStax books. To access the technology options for your text, visit your book page on OpenStax.org.

About the authors

Senior contributing authors

Steven A. Greenlaw, University of Mary Washington
Steven Greenlaw has been teaching principles of economics for more than 30 years. In 1999, he received the Grellet C. Simpson Award for Excellence in Undergraduate Teaching at the University of Mary Washington. He is the author of *Doing Economics: A Guide to Doing and Understanding Economic Research*, as well as a variety of articles on economics pedagogy and instructional technology, published in the *Journal of Economic Education*, the *International Review of Economic Education*, and other outlets. He wrote the module on Quantitative Writing for *Starting Point: Teaching and Learning Economics*, the web portal on best practices in teaching economics. Steven Greenlaw lives in Alexandria, Virginia with his wife Kathy and their three children.

David Shapiro, Pennsylvania State University
David Shapiro is Professor Emeritus of Economics, Demography, and Women's, Gender, and Sexuality Studies at the Pennsylvania State University. He received a BA in economics and political science from the University of Michigan, and an MA as well as a PhD in economics from Princeton University. He began his academic career at Ohio State University in 1971, and moved to Penn State in 1980. His early research focused on women and youth in the United States labor market. Following a 1978-79 stint as a Fulbright professor at the University of Kinshasa in the Democratic Republic of the Congo, his research shifted focus to fertility in Kinshasa and more broadly, in sub-Saharan Africa. He has also received the top prize for teaching at both Ohio State and Penn State.

Special thanks to Christian Potter from University of Mary Washington, who thoroughly researched and applied many of the data updates and provided the foundation for many new and revised illustrations.

Development editor

Thomas Sigel

Contributing authors

Eric Dodge, Hanover College
Cynthia Gamez, University of Texas at El Paso
Andres Jauregui, Columbus State University
Diane Keenan, Cerritos College
Dan MacDonald, California State University, San Bernardino
Amyaz Moledina, The College of Wooster
Craig Richardson, Winston-Salem State University
David Shapiro, Pennsylvania State University
Ralph Sonenshine, American University

Reviewers

Bryan Aguiar, Northwest Arkansas Community College
Basil Al Hashimi, Mesa Community College
Emil Berendt, Mount St. Mary's University
Zena Buser, Adams State University
Douglas Campbell, The University of Memphis
Sanjukta Chaudhuri, University of Wisconsin-Eau Claire
Xueyu Cheng, Alabama State University

Robert Cunningham, Alma College
Rosa Lea Danielson, College of DuPage
Steven Deloach, Elon University
Michael Enz, Framingham State University
Debbie Evercloud, University of Colorado Denver
Reza Ghorashi, Richard Stockton College of New Jersey
Robert Gillette, University of Kentucky
Shaomin Huang, Lewis-Clark State College
George Jones, University of Wisconsin-Rock County
Charles Kroncke, College of Mount St. Joseph
Teresa Laughlin, Palomar Community College
Carlos Liard-Muriente, Central Connecticut State University
Heather Luea, Kansas State University
Steven Lugauer, University of Notre Dame
William Mosher, Nashua Community College
Michael Netta, Hudson County Community College
Nick Noble, Miami University
Joe Nowakowski, Muskingum University
Shawn Osell, University of Wisconsin-Superior
Mark Owens, Middle Tennessee State University
Sonia Pereira, Barnard College
Jennifer Platania, Elon University
Robert Rycroft, University of Mary Washington
Adrienne Sachse, Florida State College at Jacksonville
Hans Schumann, Texas A&M University
Gina Shamshak, Goucher College
Chris Warburton, John Jay College of Criminal Justice, CUNY
Mark Witte, Northwestern University

1 | Welcome to Economics!

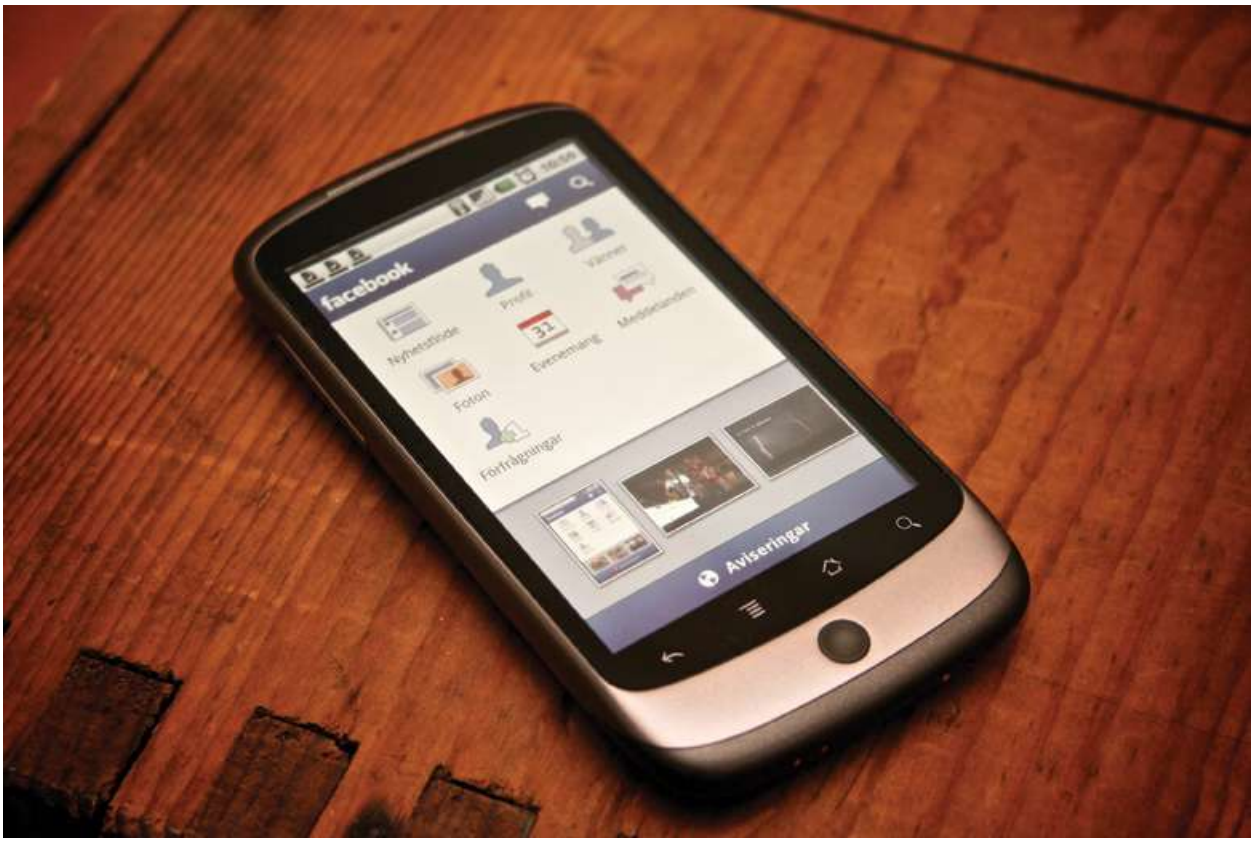


Figure 1.1 Do You Use Facebook? Economics is greatly impacted by how well information travels through society. Today, social media giants Twitter, Facebook, and Instagram are major forces on the information super highway. (Credit: Johan Larsson/Flickr)

Bring it Home

Decisions ... Decisions in the Social Media Age

To post or not to post? Every day we are faced with a myriad of decisions, from what to have for breakfast, to which route to take to class, to the more complex—"Should I double major and add possibly another semester of study to my education?" Our response to these choices depends on the information we have available at any given moment. Economists call this "imperfect" because we rarely have all the data we need to make perfect decisions. Despite the lack of perfect information, we still make hundreds of decisions a day.

Now we have another avenue in which to gather information—social media. Outlets like Facebook and Twitter are altering the process by which we make choices, how we spend our time, which movies we see, which products we buy, and more. How many of you chose a university without checking out its Facebook page or Twitter stream first for information and feedback?

As you will see in this course, what happens in economics is affected by how well and how fast information disseminates through a society, such as how quickly information travels through Facebook. "Economists love nothing better than when deep and liquid markets operate under conditions of perfect information," says Jessica Irvine, National Economics Editor for News Corp Australia.

This leads us to the topic of this chapter, an introduction to the world of making decisions, processing

information, and understanding behavior in markets —the world of economics. Each chapter in this book will start with a discussion about current (or sometimes past) events and revisit it at chapter's end—to "bring home" the concepts in play.

Introduction

In this chapter, you will learn about:

- What Is Economics, and Why Is It Important?
- Microeconomics and Macroeconomics
- How Economists Use Theories and Models to Understand Economic Issues
- How Economies Can Be Organized: An Overview of Economic Systems

What is economics and why should you spend your time learning it? After all, there are other disciplines you could be studying, and other ways you could be spending your time. As the Bring it Home feature just mentioned, making choices is at the heart of what economists study, and your decision to take this course is as much as economic decision as anything else.

Economics is probably not what you think. It is not primarily about money or finance. It is not primarily about business. It is not mathematics. What is it then? It is both a subject area and a way of viewing the world.

1.1 | What Is Economics, and Why Is It Important?

By the end of this section, you will be able to:

- Discuss the importance of studying economics
- Explain the relationship between production and division of labor
- Evaluate the significance of scarcity

Economics is the study of how humans make decisions in the face of scarcity. These can be individual decisions, family decisions, business decisions or societal decisions. If you look around carefully, you will see that scarcity is a fact of life. **Scarcity** means that human wants for goods, services and resources exceed what is available. Resources, such as labor, tools, land, and raw materials are necessary to produce the goods and services we want but they exist in limited supply. Of course, the ultimate scarce resource is time- everyone, rich or poor, has just 24 expendable hours in the day to earn income to acquire goods and services, for leisure time, or for sleep. At any point in time, there is only a finite amount of resources available.

Think about it this way: In 2015 the labor force in the United States contained over 158 million workers, according to the U.S. Bureau of Labor Statistics. The total land area was 3,794,101 square miles. While these are certainly large numbers, they are not infinite. Because these resources are limited, so are the numbers of goods and services we produce with them. Combine this with the fact that human wants seem to be virtually infinite, and you can see why scarcity is a problem.

Introduction to FRED

Data is very important in economics because it describes and measures the issues and problems that economics seek to understand. A variety of government agencies publish economic and social data. For this course, we will generally use data from the St. Louis Federal Reserve Bank's FRED database. FRED is very user friendly. It allows you to display data in tables or charges, and you can easily download it into spreadsheet form if you want to use the data for other purposes. The **FRED website (https://openstax.org/l/FRED/)** includes data on nearly 400,000 domestic and international variables over time, in the following broad categories:

- Money, Banking & Finance
- Population, Employment, & Labor Markets (including Income Distribution)

- National Accounts (Gross Domestic Product & its components), Flow of Funds, and International Accounts
- Production & Business Activity (including Business Cycles)
- Prices & Inflation (including the Consumer Price Index, the Producer Price Index, and the Employment Cost Index)
- International Data from other nations
- U.S. Regional Data
- Academic Data (including Penn World Tables & NBER Macrohistory database)

For more information about how to use FRED, see the variety of **videos (https://openstax.org/l/FRED_intro)** on YouTube starting with this introduction.

Figure 1.2 Scarcity of Resources Homeless people are a stark reminder that scarcity of resources is real. (Credit: "daveynin"/Flickr Creative Commons)

If you still do not believe that scarcity is a problem, consider the following: Does everyone require food to eat? Does everyone need a decent place to live? Does everyone have access to healthcare? In every country in the world, there are people who are hungry, homeless (for example, those who call park benches their beds, as **Figure 1.2** shows), and in need of healthcare, just to focus on a few critical goods and services. Why is this the case? It is because of scarcity. Let's delve into the concept of scarcity a little deeper, because it is crucial to understanding economics.

The Problem of Scarcity

Think about all the things you consume: food, shelter, clothing, transportation, healthcare, and entertainment. How do you acquire those items? You do not produce them yourself. You buy them. How do you afford the things you buy? You work for pay. If you do not, someone else does on your behalf. Yet most of us never have enough income to buy all the things we want. This is because of scarcity. So how do we solve it?

Link It Up

Visit this website (http://openstax.org/l/drought) to read about how the United States is dealing with scarcity in resources.

Every society, at every level, must make choices about how to use its resources. Families must decide whether to spend their money on a new car or a fancy vacation. Towns must choose whether to put more of the budget into police and fire protection or into the school system. Nations must decide whether to devote more funds to national defense or to protecting the environment. In most cases, there just isn't enough money in the budget to do everything. How do we use our limited resources the best way possible, that is, to obtain the most goods and services we can? There are a couple of options. First, we could each produce everything we each consume. Alternatively, we could each produce some of what we want to consume, and "trade" for the rest of what we want. Let's explore these options. Why do we not each just produce all of the things we consume? Think back to pioneer days, when individuals knew how to do so much more than we do today, from building their homes, to growing their crops, to hunting for food, to repairing their equipment. Most of us do not know how to do all—or any—of those things, but it is not because we could not learn. Rather, we do not have to. The reason why is something called *the division and specialization of labor*, a production innovation first put forth by Adam Smith (**Figure 1.3**) in his book, *The Wealth of Nations*.

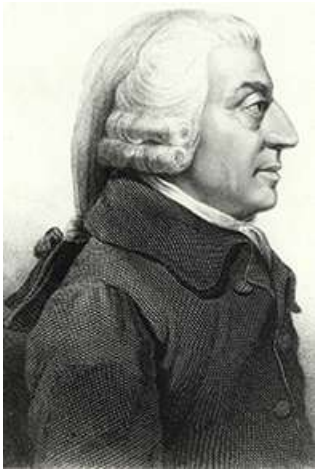


Figure 1.3 Adam Smith Adam Smith introduced the idea of dividing labor into discrete tasks. (Credit: Wikimedia Commons)

The Division of and Specialization of Labor

The formal study of economics began when Adam Smith (1723–1790) published his famous book *The Wealth of Nations* in 1776. Many authors had written on economics in the centuries before Smith, but he was the first to address the subject in a comprehensive way. In the first chapter, Smith introduces the concept of **division of labor**, which means that the way one produces a good or service is divided into a number of tasks that different workers perform, instead of all the tasks being done by the same person.

To illustrate division of labor, Smith counted how many tasks went into making a pin: drawing out a piece of wire, cutting it to the right length, straightening it, putting a head on one end and a point on the other, and packaging pins for sale, to name just a few. Smith counted 18 distinct tasks that different people performed—all for a pin, believe it or not!

Modern businesses divide tasks as well. Even a relatively simple business like a restaurant divides the task of serving meals into a range of jobs like top chef, sous chefs, less-skilled kitchen help, servers to wait on the tables, a greeter at the door, janitors to clean up, and a business manager to handle paychecks and bills—not to mention the economic

connections a restaurant has with suppliers of food, furniture, kitchen equipment, and the building where it is located. A complex business like a large manufacturing factory, such as the shoe factory (**Figure 1.4**), or a hospital can have hundreds of job classifications.

Figure 1.4 Division of Labor Workers on an assembly line are an example of the divisions of labor. (Credit: Nina Hale/Flickr Creative Commons)

Why the Division of Labor Increases Production

When we divide and subdivide the tasks involved with producing a good or service, workers and businesses can produce a greater quantity of output. In his observations of pin factories, Smith noticed that one worker alone might make 20 pins in a day, but that a small business of 10 workers (some of whom would need to complete two or three of the 18 tasks involved with pin-making), could make 48,000 pins in a day. How can a group of workers, each specializing in certain tasks, produce so much more than the same number of workers who try to produce the entire good or service by themselves? Smith offered three reasons.

First, **specialization** in a particular small job allows workers to focus on the parts of the production process where they have an advantage. (In later chapters, we will develop this idea by discussing comparative advantage.) People have different skills, talents, and interests, so they will be better at some jobs than at others. The particular advantages may be based on educational choices, which are in turn shaped by interests and talents. Only those with medical degrees qualify to become doctors, for instance. For some goods, geography affects specialization. For example, it is easier to be a wheat farmer in North Dakota than in Florida, but easier to run a tourist hotel in Florida than in North Dakota. If you live in or near a big city, it is easier to attract enough customers to operate a successful dry cleaning business or movie theater than if you live in a sparsely populated rural area. Whatever the reason, if people specialize in the production of what they do best, they will be more effective than if they produce a combination of things, some of which they are good at and some of which they are not.

Second, workers who specialize in certain tasks often learn to produce more quickly and with higher quality. This pattern holds true for many workers, including assembly line laborers who build cars, stylists who cut hair, and doctors who perform heart surgery. In fact, specialized workers often know their jobs well enough to suggest innovative ways to do their work faster and better.

A similar pattern often operates within businesses. In many cases, a business that focuses on one or a few products (sometimes called its “core competency”) is more successful than firms that try to make a wide range of products.

Third, specialization allows businesses to take advantage of **economies of scale**, which means that for many goods, as the level of production increases, the average cost of producing each individual unit declines. For example, if a factory produces only 100 cars per year, each car will be quite expensive to make on average. However, if a factory produces 50,000 cars each year, then it can set up an assembly line with huge machines and workers performing specialized tasks, and the average cost of production per car will be lower. The ultimate result of workers who can focus on their preferences and talents, learn to do their specialized jobs better, and work in larger organizations is that society as a whole can produce and consume far more than if each person tried to produce all of his or her own goods and services. The division and specialization of labor has been a force against the problem of scarcity.

Trade and Markets

Specialization only makes sense, though, if workers can use the pay they receive for doing their jobs to purchase the other goods and services that they need. In short, specialization requires trade.

You do not have to know anything about electronics or sound systems to play music—you just buy an iPod or MP3 player, download the music, and listen. You do not have to know anything about artificial fibers or the construction of sewing machines if you need a jacket—you just buy the jacket and wear it. You do not need to know anything about internal combustion engines to operate a car—you just get in and drive. Instead of trying to acquire all the knowledge and skills involved in producing all of the goods and services that you wish to consume, the market allows you to learn a specialized set of skills and then use the pay you receive to buy the goods and services you need or want. This is how our modern society has evolved into a strong economy.

Why Study Economics?

Now that you have an overview on what economics studies, let's quickly discuss why you are right to study it. Economics is not primarily a collection of facts to memorize, although there are plenty of important concepts to learn. Instead, think of economics as a collection of questions to answer or puzzles to work. Most importantly, economics provides the tools to solve those puzzles. If the economics "bug" has not bitten you yet, there are other reasons why you should study economics.

- Virtually every major problem facing the world today, from global warming, to world poverty, to the conflicts in Syria, Afghanistan, and Somalia, has an economic dimension. If you are going to be part of solving those problems, you need to be able to understand them. Economics is crucial.
- It is hard to overstate the importance of economics to good citizenship. You need to be able to vote intelligently on budgets, regulations, and laws in general. When the U.S. government came close to a standstill at the end of 2012 due to the "fiscal cliff," what were the issues? Did you know?
- A basic understanding of economics makes you a well-rounded thinker. When you read articles about economic issues, you will understand and be able to evaluate the writer's argument. When you hear classmates, co-workers, or political candidates talking about economics, you will be able to distinguish between common sense and nonsense. You will find new ways of thinking about current events and about personal and business decisions, as well as current events and politics.

The study of economics does not dictate the answers, but it can illuminate the different choices.

1.2 | Microeconomics and Macroeconomics

By the end of this section, you will be able to:

- Describe microeconomics
- Describe macroeconomics
- Contrast monetary policy and fiscal policy

Economics is concerned with the well-being of *all* people, including those with jobs and those without jobs, as well as those with high incomes and those with low incomes. Economics acknowledges that production of useful goods and services can create problems of environmental pollution. It explores the question of how investing in education helps to develop workers' skills. It probes questions like how to tell when big businesses or big labor unions are operating in a way that benefits society as a whole and when they are operating in a way that benefits their owners or members at the expense of others. It looks at how government spending, taxes, and regulations affect decisions about production and consumption.

It should be clear by now that economics covers considerable ground. We can divide that ground into two parts: **Microeconomics** focuses on the actions of individual agents within the economy, like households, workers, and businesses. **Macroeconomics** looks at the economy as a whole. It focuses on broad issues such as growth of production, the number of unemployed people, the inflationary increase in prices, government deficits, and levels of exports and imports. Microeconomics and macroeconomics are not separate subjects, but rather complementary perspectives on the overall subject of the economy.

To understand why both microeconomic and macroeconomic perspectives are useful, consider the problem of studying a biological ecosystem like a lake. One person who sets out to study the lake might focus on specific topics: certain kinds of algae or plant life; the characteristics of particular fish or snails; or the trees surrounding the lake. Another person might take an overall view and instead consider the lake's ecosystem from top to bottom; what eats what, how the system stays in a rough balance, and what environmental stresses affect this balance. Both approaches are useful, and both examine the same lake, but the viewpoints are different. In a similar way, both microeconomics and macroeconomics study the same economy, but each has a different viewpoint.

Whether you are scrutinizing lakes or economics, the micro and the macro insights should blend with each other. In studying a lake, the micro insights about particular plants and animals help to understand the overall food chain, while the macro insights about the overall food chain help to explain the environment in which individual plants and animals live.

In economics, the micro decisions of individual businesses are influenced by whether the macroeconomy is healthy. For example, firms will be more likely to hire workers if the overall economy is growing. In turn, macroeconomy's performance ultimately depends on the microeconomic decisions that individual households and businesses make.

Microeconomics

What determines how households and individuals spend their budgets? What combination of goods and services will best fit their needs and wants, given the budget they have to spend? How do people decide whether to work, and if so, whether to work full time or part time? How do people decide how much to save for the future, or whether they should borrow to spend beyond their current means?

What determines the products, and how many of each, a firm will produce and sell? What determines the prices a firm will charge? What determines how a firm will produce its products? What determines how many workers it will hire? How will a firm finance its business? When will a firm decide to expand, downsize, or even close? In the microeconomics part of this book, we will learn about the theory of consumer behavior, the theory of the firm, how markets for labor and other resources work, and how markets sometimes fail to work properly.

Macroeconomics

What determines the level of economic activity in a society? In other words, what determines how many goods and services a nation actually produces? What determines how many jobs are available in an economy? What determines a nation’s standard of living? What causes the economy to speed up or slow down? What causes firms to hire more workers or to lay them off? Finally, what causes the economy to grow over the long term?

We can determine an economy's macroeconomic health by examining a number of goals: growth in the standard of living, low unemployment, and low inflation, to name the most important. How can we use government macroeconomic policy to pursue these goals? A nation's central bank conducts **monetary policy**, which involves policies that affect bank lending, interest rates, and financial capital markets. For the United States, this is the Federal Reserve. A nation's legislative body determines **fiscal policy**, which involves government spending and taxes. For the United States, this is the Congress and the executive branch, which originates the federal budget. These are the government's main tools. Americans tend to expect that government can fix whatever economic problems we encounter, but to what extent is that expectation realistic? These are just some of the issues that we will explore in the macroeconomic chapters of this book.

1.3 | How Economists Use Theories and Models to Understand Economic Issues

By the end of this section, you will be able to:

- Interpret a circular flow diagram
- Explain the importance of economic theories and models
- Describe goods and services markets and labor markets

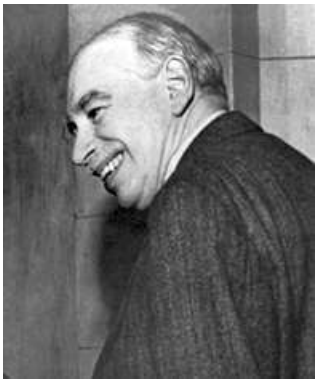


Figure 1.5 John Maynard Keynes One of the most influential economists in modern times was John Maynard Keynes. (Credit: Wikimedia Commons)

John Maynard Keynes (1883–1946), one of the greatest economists of the twentieth century, pointed out that economics is not just a subject area but also a way of thinking. Keynes (**Figure 1.5**) famously wrote in the introduction to a fellow economist's book: "[Economics] is a method rather than a doctrine, an apparatus of the mind, a technique of thinking, which helps its possessor to draw correct conclusions." In other words, economics teaches you how to think, not what to think.

Link It Up

Watch this video (http://openstax.org/l/Keynes) about John Maynard Keynes and his influence on economics.

Economists see the world through a different lens than anthropologists, biologists, classicists, or practitioners of any other discipline. They analyze issues and problems using economic theories that are based on particular assumptions about human behavior. These assumptions tend to be different than the assumptions an anthropologist or psychologist might use. A **theory** is a simplified representation of how two or more variables interact with each other. The purpose of a theory is to take a complex, real-world issue and simplify it down to its essentials. If done well, this enables the analyst to understand the issue and any problems around it. A good theory is simple enough to understand, while complex enough to capture the key features of the object or situation you are studying.

Sometimes economists use the term **model** instead of theory. Strictly speaking, a theory is a more abstract representation, while a model is a more applied or empirical representation. We use models to test theories, but for this course we will use the terms interchangeably.

For example, an architect who is planning a major office building will often build a physical model that sits on a tabletop to show how the entire city block will look after the new building is constructed. Companies often build models of their new products, which are more rough and unfinished than the final product, but can still demonstrate how the new product will work.

A good model to start with in economics is the **circular flow diagram** (**Figure 1.6**). It pictures the economy as consisting of two groups—households and firms—that interact in two markets: the **goods and services market** in which firms sell and households buy and the **labor market** in which households sell labor to business firms or other employees.

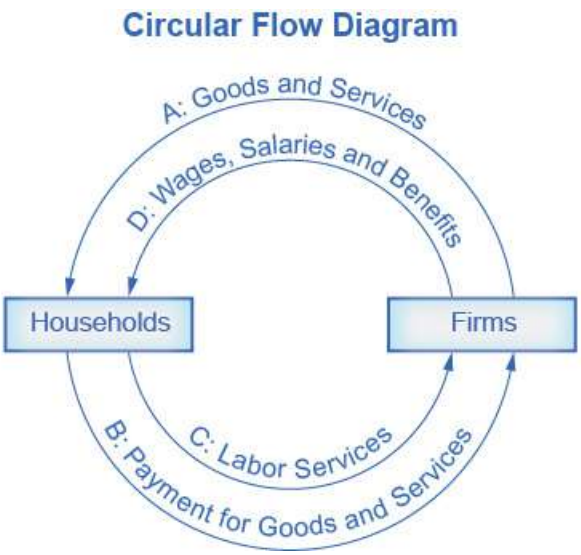


Figure 1.6 The Circular Flow Diagram The circular flow diagram shows how households and firms interact in the goods and services market, and in the labor market. The direction of the arrows shows that in the goods and services market, households receive goods and services and pay firms for them. In the labor market, households provide labor and receive payment from firms through wages, salaries, and benefits.

Firms produce and sell goods and services to households in the market for goods and services (or product market). Arrow "A" indicates this. Households pay for goods and services, which becomes the revenues to firms. Arrow "B" indicates this. Arrows A and B represent the two sides of the product market. Where do households obtain the income to buy goods and services? They provide the labor and other resources (e.g. land, capital, raw materials) firms need to produce goods and services in the market for inputs (or factors of production). Arrow "C" indicates this. In return, firms pay for the inputs (or resources) they use in the form of wages and other factor payments. Arrow "D" indicates this. Arrows "C" and "D" represent the two sides of the factor market.

Of course, in the real world, there are many different markets for goods and services and markets for many different types of labor. The circular flow diagram simplifies this to make the picture easier to grasp. In the diagram, firms produce goods and services, which they sell to households in return for revenues. The outer circle shows this, and represents the two sides of the product market (for example, the market for goods and services) in which households demand and firms supply. Households sell their labor as workers to firms in return for wages, salaries, and benefits. The inner circle shows this and represents the two sides of the labor market in which households supply and firms demand.

This version of the circular flow model is stripped down to the essentials, but it has enough features to explain how the product and labor markets work in the economy. We could easily add details to this basic model if we wanted to introduce more real-world elements, like financial markets, governments, and interactions with the rest of the globe (imports and exports).

Economists carry a set of theories in their heads like a carpenter carries around a toolkit. When they see an economic issue or problem, they go through the theories they know to see if they can find one that fits. Then they use the theory to derive insights about the issue or problem. Economists express theories as diagrams, graphs, or even as mathematical equations. (Do not worry. In this course, we will mostly use graphs.) Economists do not figure out the answer to the problem first and then draw the graph to illustrate. Rather, they use the graph of the theory to help them figure out the answer. Although at the introductory level, you can sometimes figure out the right answer without applying a model, if you keep studying economics, before too long you will run into issues and problems that you will need to graph to solve. We explain both micro and macroeconomics in terms of theories and models. The most well-known theories are probably those of supply and demand, but you will learn a number of others.

1.4 | How To Organize Economies: An Overview of Economic Systems

By the end of this section, you will be able to:

- Contrast traditional economies, command economies, and market economies
- Explain gross domestic product (GDP)
- Assess the importance and effects of globalization

Think about what a complex system a modern economy is. It includes all production of goods and services, all buying and selling, all employment. The economic life of every individual is interrelated, at least to a small extent, with the economic lives of thousands or even millions of other individuals. Who organizes and coordinates this system? Who insures that, for example, the number of televisions a society provides is the same as the amount it needs and wants? Who insures that the right number of employees work in the electronics industry? Who insures that televisions are produced in the best way possible? How does it all get done?

There are at least three ways that societies organize an economy. The first is the **traditional economy**, which is the oldest economic system and is used in parts of Asia, Africa, and South America. Traditional economies organize their economic affairs the way they have always done (i.e., tradition). Occupations stay in the family. Most families are farmers who grow the crops using traditional methods. What you produce is what you consume. Because tradition drives the way of life, there is little economic progress or development.

Figure 1.7 A Command Economy Ancient Egypt was an example of a command economy. (Credit: Jay Bergesen/ Flickr Creative Commons)

Command economies are very different. In a **command economy**, economic effort is devoted to goals passed down from a ruler or ruling class. Ancient Egypt was a good example: a large part of economic life was devoted to building pyramids, like those in **Figure 1.7**, for the pharaohs. Medieval manor life is another example: the lord provided the land for growing crops and protection in the event of war. In return, vassals provided labor and soldiers to do the lord's bidding. In the last century, communism emphasized command economies.

In a command economy, the government decides what goods and services will be produced and what prices it will charge for them. The government decides what methods of production to use and sets wages for workers. The government provides many necessities like healthcare and education for free. Currently, Cuba and North Korea have command economies.

Figure 1.8 A Market Economy Nothing says “market” more than The New York Stock Exchange. (Credit: Erik Drost/ Flickr Creative Commons)

Although command economies have a very centralized structure for economic decisions, market economies have a very decentralized structure. A **market** is an institution that brings together buyers and sellers of goods or services, who may be either individuals or businesses. The New York Stock Exchange (**Figure 1.8**) is a prime example of a market which brings buyers and sellers together. In a **market economy**, decision-making is decentralized. Market economies are based on **private enterprise**: the private individuals or groups of private individuals own and operate the means of production (resources and businesses). Businesses supply goods and services based on demand. (In a command economy, by contrast, the government owns resources and businesses.) Supply of goods and services depends on what the demands. A person’s income is based on his or her ability to convert resources (especially labor) into something that society values. The more society values the person’s output, the higher the income (think Lady Gaga or LeBron James). In this scenario, market forces, not governments, determine economic decisions.

Most economies in the real world are mixed. They combine elements of command and market (and even traditional) systems. The U.S. economy is positioned toward the market-oriented end of the spectrum. Many countries in Europe and Latin America, while primarily market-oriented, have a greater degree of government involvement in economic decisions than the U.S. economy. China and Russia, while over the past several decades have moved more in the direction of having a market-oriented system, remain closer to the command economy end of the spectrum. The Heritage Foundation provides information about how free and thus market-oriented different countries' are, as the following Clear It Up feature discusses. For a similar ranking, but one that defines freedom more broadly, see the Cato Foundation's Human Freedom **Index (https://openstax.org/l/cato)** .

Clear It Up

What countries are considered economically free?

Who is in control of economic decisions? Are people free to do what they want and to work where they want? Are businesses free to produce when they want and what they choose, and to hire and fire as they wish? Are banks free to choose who will receive loans, or does the government control these kinds of choices? Each year, researchers at the Heritage Foundation and the *Wall Street Journal* look at 50 different categories of economic freedom for countries around the world. They give each nation a score based on the extent of economic freedom in each category.

The 2016 Heritage Foundation’s Index of Economic Freedom report ranked 178 countries around the world: **Table 1.1** lists some examples of the most free and the least free countries. Several additional countries were not ranked because of extreme instability that made judgments about economic freedom impossible. These countries include Afghanistan, Iraq, Libya, Syria, Somalia, and Yemen.

The assigned rankings are inevitably based on estimates, yet even these rough measures can be useful for discerning trends. In 2015, 101 of the 178 included countries shifted toward greater economic freedom, although 77 of the countries shifted toward less economic freedom. In recent decades, the overall trend has been a *higher level of economic freedom around the world*.

Most Economic Freedom	Least Economic Freedom
1. Hong Kong	167. Timor-Leste
2. Singapore	168. Democratic Republic of Congo
3. New Zealand	169. Argentina
4. Switzerland	170. Equatorial Guinea
5. Australia	171. Iran
6. Canada	172. Republic of Congo
7. Chile	173. Eritrea
8. Ireland	174. Turkmenistan
9. Estonia	175. Zimbabwe
10. United Kingdom	176. Venezuela
11. United States	177. Cuba
12. Denmark	178. North Korea

Table 1.1 Economic Freedoms, 2016 (Source: The Heritage Foundation, 2016 Index of Economic Freedom, Country Rankings, http://www.heritage.org/index/ranking)

Regulations: The Rules of the Game

Markets and government regulations are always entangled. There is no such thing as an absolutely free market. Regulations always define the "rules of the game" in the economy. Economies that are primarily market-oriented have fewer regulations—ideally just enough to maintain an even playing field for participants. At a minimum, these laws govern matters like safeguarding private property against theft, protecting people from violence, enforcing legal contracts, preventing fraud, and collecting taxes. Conversely, even the most command-oriented economies operate using markets. How else would buying and selling occur? The government heavily regulates decisions of what to produce and prices to charge. Heavily regulated economies often have **underground economies** (or black markets), which are markets where the buyers and sellers make transactions without the government's approval.

The question of how to organize economic institutions is typically not a black-or-white choice between all market or all government, but instead involves a balancing act over the appropriate combination of market freedom and government rules.

Figure 1.9 Globalization Cargo ships are one mode of transportation for shipping goods in the global economy. (Credit: Raul Valdez/Flickr Creative Commons)

The Rise of Globalization

Recent decades have seen a trend toward **globalization**, which is the expanding cultural, political, and economic connections between people around the world. One measure of this is the increased buying and selling of goods, services, and assets across national borders—in other words, international trade and financial capital flows.

Globalization has occurred for a number of reasons. Improvements in shipping, as illustrated by the container ship in **Figure 1.9**, and air cargo have driven down transportation costs. Innovations in computing and telecommunications have made it easier and cheaper to manage long-distance economic connections of production and sales. Many valuable products and services in the modern economy can take the form of information—for example: computer software; financial advice; travel planning; music, books and movies; and blueprints for designing a building. These products and many others can be transported over telephones and computer networks at ever-lower costs. Finally, international agreements and treaties between countries have encouraged greater trade.

Table 1.2 presents one measure of globalization. It shows the percentage of domestic economic production that was exported for a selection of countries from 2010 to 2015, according to an entity known as The World Bank. **Exports** are the goods and services that one produces domestically and sells abroad. **Imports** are the goods and services that one produces abroad and then sells domestically. **Gross domestic product (GDP)** measures the size of total production in an economy. Thus, the ratio of exports divided by GDP measures what share of a country's total economic production is sold in other countries.

Country	2010	2011	2012	2013	2014	2015
Higher Income Countries						
United States	12.4	13.6	13.6	13.5	13.5	12.6
Belgium	76.2	81.4	82.2	82.8	84.0	84.4
Canada	29.1	30.7	30.0	30.1	31.7	31.5
France	26.0	27.8	28.1	28.3	29.0	30.0
Middle Income Countries						
Brazil	10.9	11.9	12.6	12.6	11.2	13.0
Mexico	29.9	31.2	32.6	31.7	32.3	35.3
South Korea	49.4	55.7	56.3	53.9	50.3	45.9

Table 1.2 The Extent of Globalization (exports/GDP) (Source: http://databank.worldbank.org/data/)

Country	2010	2011	2012	2013	2014	2015
Lower Income Countries						
Chad	36.8	38.9	36.9	32.2	34.2	29.8
China	29.4	28.5	27.3	26.4	23.9	22.4
India	22.0	23.9	24.0	24.8	22.9	-
Nigeria	25.3	31.3	31.4	18.0	18.4	-

Table 1.2 The Extent of Globalization (exports/GDP) (Source: http://databank.worldbank.org/data/)

In recent decades, the export/GDP ratio has generally risen, both worldwide and for the U.S. economy. Interestingly, the share of U.S. exports in proportion to the U.S. economy is well below the global average, in part because large economies like the United States can contain more of the division of labor inside their national borders. However, smaller economies like Belgium, Korea, and Canada need to trade across their borders with other countries to take full advantage of division of labor, specialization, and economies of scale. In this sense, the enormous U.S. economy is less affected by globalization than most other countries.

Table 1.2 indicates that many medium and low income countries around the world, like Mexico and China, have also experienced a surge of globalization in recent decades. If an astronaut in orbit could put on special glasses that make all economic transactions visible as brightly colored lines and look down at Earth, the astronaut would see the planet covered with connections.

Despite the rise in globalization over the last few decades, in recent years we've seen significant pushback against globalization from people across the world concerned about loss of jobs, loss of political sovereignty, and increased economic inequality. Prominent examples of this pushback include the 2016 vote in Great Britain to exit the European Union (i.e. Brexit), and the election of Donald J. Trump for President of the United States.

Hopefully, you now have an idea about economics. Before you move to any other chapter of study, be sure to read the very important appendix to this chapter called **The Use of Mathematics in Principles of Economics**. It is essential that you learn more about how to read and use models in economics.

Bring it Home

Decisions ... Decisions in the Social Media Age

The world we live in today provides nearly instant access to a wealth of information. Consider that as recently as the late 1970s, the *Farmer's Almanac*, along with the Weather Bureau of the U.S. Department of Agriculture, were the primary sources American farmers used to determine when to plant and harvest their crops. Today, farmers are more likely to access, online, weather forecasts from the National Oceanic and Atmospheric Administration or watch the Weather Channel. After all, knowing the upcoming forecast could drive when to harvest crops. Consequently, knowing the upcoming weather could change the amount of crop harvested.

Some relatively new information forums, such as Facebook, are rapidly changing how information is distributed; hence, influencing decision making. In 2014, the Pew Research Center reported that 71% of online adults use Facebook. This social media forum posts topics ranging from the National Basketball Association, to celebrity singers and performers, to farmers.

Information helps us make decisions as simple as what to wear today to how many reporters the media should send to cover a crash. Each of these decisions is an economic decision. After all, resources are scarce. If the media send ten reporters to cover an accident, they are not available to cover other stories or complete other tasks. Information provides the necessary knowledge to make the best possible decisions on how to utilize scarce resources. Welcome to the world of economics!

KEY TERMS

circular flow diagram a diagram that views the economy as consisting of households and firms interacting in a goods and services market and a labor market

command economy an economy where economic decisions are passed down from government authority and where the government owns the resources

division of labor the way in which different workers divide required tasks to produce a good or service

economics the study of how humans make choices under conditions of scarcity

economies of scale when the average cost of producing each individual unit declines as total output increases

exports products (goods and services) made domestically and sold abroad

fiscal policy economic policies that involve government spending and taxes

globalization the trend in which buying and selling in markets have increasingly crossed national borders

goods and services market a market in which firms are sellers of what they produce and households are buyers

gross domestic product (GDP) measure of the size of total production in an economy

imports products (goods and services) made abroad and then sold domestically

labor market the market in which households sell their labor as workers to business firms or other employers

macroeconomics the branch of economics that focuses on broad issues such as growth, unemployment, inflation, and trade balance

market interaction between potential buyers and sellers; a combination of demand and supply

market economy an economy where economic decisions are decentralized, private individuals own resources, and businesses supply goods and services based on demand

microeconomics the branch of economics that focuses on actions of particular agents within the economy, like households, workers, and business firms

model see theory

monetary policy policy that involves altering the level of interest rates, the availability of credit in the economy, and the extent of borrowing

private enterprise system where private individuals or groups of private individuals own and operate the means of production (resources and businesses)

scarcity when human wants for goods and services exceed the available supply

specialization when workers or firms focus on particular tasks for which they are well-suited within the overall production process

theory a representation of an object or situation that is simplified while including enough of the key features to help us understand the object or situation

traditional economy typically an agricultural economy where things are done the same as they have always been done

underground economy a market where the buyers and sellers make transactions in violation of one or more

government regulations

KEY CONCEPTS AND SUMMARY

1.1 What Is Economics, and Why Is It Important?

Economics seeks to solve the problem of scarcity, which is when human wants for goods and services exceed the available supply. A modern economy displays a division of labor, in which people earn income by specializing in what they produce and then use that income to purchase the products they need or want. The division of labor allows individuals and firms to specialize and to produce more for several reasons: a) It allows the agents to focus on areas of advantage due to natural factors and skill levels; b) It encourages the agents to learn and invent; c) It allows agents to take advantage of economies of scale. Division and specialization of labor only work when individuals can purchase what they do not produce in markets. Learning about economics helps you understand the major problems facing the world today, prepares you to be a good citizen, and helps you become a well-rounded thinker.

1.2 Microeconomics and Macroeconomics

Microeconomics and macroeconomics are two different perspectives on the economy. The microeconomic perspective focuses on parts of the economy: individuals, firms, and industries. The macroeconomic perspective looks at the economy as a whole, focusing on goals like growth in the standard of living, unemployment, and inflation. Macroeconomics has two types of policies for pursuing these goals: monetary policy and fiscal policy.

1.3 How Economists Use Theories and Models to Understand Economic Issues

Economists analyze problems differently than do other disciplinary experts. The main tools economists use are economic theories or models. A theory is not an illustration of the answer to a problem. Rather, a theory is a tool for determining the answer.

1.4 How To Organize Economies: An Overview of Economic Systems

We can organize societies as traditional, command, or market-oriented economies. Most societies are a mix. The last few decades have seen globalization evolve as a result of growth in commercial and financial networks that cross national borders, making businesses and workers from different economies increasingly interdependent.

SELF-CHECK QUESTIONS

1. What is scarcity? Can you think of two causes of scarcity?

2. Residents of the town of Smithfield like to consume hams, but each ham requires 10 people to produce it and takes a month. If the town has a total of 100 people, what is the maximum amount of ham the residents can consume in a month?

3. A consultant works for $200 per hour. She likes to eat vegetables, but is not very good at growing them. Why does it make more economic sense for her to spend her time at the consulting job and shop for her vegetables?

4. A computer systems engineer could paint his house, but it makes more sense for him to hire a painter to do it. Explain why.

5. What would be another example of a “system” in the real world that could serve as a metaphor for micro and macroeconomics?

6. Suppose we extend the circular flow model to add imports and exports. Copy the circular flow diagram onto a sheet of paper and then add a foreign country as a third agent. Draw a rough sketch of the flows of imports, exports, and the payments for each on your diagram.

7. What is an example of a problem in the world today, not mentioned in the chapter, that has an economic dimension?

8. The chapter defines *private enterprise* as a characteristic of market-oriented economies. What would *public enterprise* be? *Hint*: It is a characteristic of command economies.

9. Why might Belgium, France, Italy, and Sweden have a higher export to GDP ratio than the United States?

REVIEW QUESTIONS

10. Give the three reasons that explain why the division of labor increases an economy's level of production.

11. What are three reasons to study economics?

12. What is the difference between microeconomics and macroeconomics?

13. What are examples of individual economic agents?

14. What are the three main goals of macroeconomics?

15. How did John Maynard Keynes define economics?

16. Are households primarily buyers or sellers in the goods and services market? In the labor market?

17. Are firms primarily buyers or sellers in the goods and services market? In the labor market?

18. What are the three ways that societies can organize themselves economically?

19. What is globalization? How do you think it might have affected the economy over the past decade?

CRITICAL THINKING QUESTIONS

20. Suppose you have a team of two workers: one is a baker and one is a chef. Explain why the kitchen can produce more meals in a given period of time if each worker specializes in what they do best than if each worker tries to do everything from appetizer to dessert.

21. Why would division of labor without trade not work?

22. Can you think of any examples of *free* goods, that is, goods or services that are not scarce?

23. A balanced federal budget and a balance of trade are secondary goals of macroeconomics, while growth in the standard of living (for example) is a primary goal. Why do you think that is so?

24. Macroeconomics is an aggregate of what happens at the microeconomic level. Would it be possible for what happens at the macro level to differ from how economic agents would react to some stimulus at the micro level? *Hint:* Think about the behavior of crowds.

25. Why is it unfair or meaningless to criticize a theory as "unrealistic?"

26. Suppose, as an economist, you are asked to analyze an issue unlike anything you have ever done before. Also, suppose you do not have a specific model for analyzing that issue. What should you do? *Hint:* What would a carpenter do in a similar situation?

27. Why do you think that most modern countries' economies are a mix of command and market types?

28. Can you think of ways that globalization has helped you economically? Can you think of ways that it has not?

2 | Choice in a World of Scarcity

Figure 2.1 Choices and Tradeoffs In general, the higher the degree, the higher the salary, so why aren't more people pursuing higher degrees? The short answer: choices and tradeoffs. (Credit: modification of work by "Jim, the Photographer"/Flickr Creative Commons)

Bring it Home

Choices ... To What Degree?

In 2015, the median income for workers who hold master's degrees varies from males to females. The average of the two is $2,951 weekly. Multiply this average by 52 weeks, and you get an average salary of $153,452. Compare that to the median weekly earnings for a full-time worker over 25 with no higher than a bachelor's degree: $1,224 weekly and $63,648 a year. What about those with no higher than a high school diploma in 2015? They earn just $664 weekly and $34,528 over 12 months. In other words, says the Bureau of Labor Statistics (BLS), earning a bachelor's degree boosted salaries 54% over what you would have earned if you had stopped your education after high school. A master's degree yields a salary almost double that of a high school diploma.

Given these statistics, we might expect many people to choose to go to college and at least earn a bachelor's degree. Assuming that people want to improve their material well-being, it seems like they would make those choices that provide them with the greatest opportunity to consume goods and services. As it turns out, the analysis is not nearly as simple as this. In fact, in 2014, the BLS reported that while almost 88% of the population in the United States had a high school diploma, only 33.6% of 25–65 year olds had bachelor's degrees, and only 7.4% of 25–65 year olds in 2014 had earned a master's.

This brings us to the subject of this chapter: why people make the choices they make and how economists explain those choices.

Introduction to Choice in a World of Scarcity

In this chapter, you will learn about:

- How Individuals Make Choices Based on Their Budget Constraint
- The Production Possibilities Frontier and Social Choices
- Confronting Objections to the Economic Approach

You will learn quickly when you examine the relationship between economics and scarcity that choices involve tradeoffs. Every choice has a cost.

In 1968, the Rolling Stones recorded "You Can't Always Get What You Want." Economists chuckled, because they had been singing a similar tune for decades. English economist Lionel Robbins (1898–1984), in his *Essay on the Nature and Significance of Economic Science* in 1932, described not always getting what you want in this way:

> The time at our disposal is limited. There are only twenty-four hours in the day. We have to choose between the different uses to which they may be put. ... Everywhere we turn, if we choose one thing we must relinquish others which, in different circumstances, we would wish not to have relinquished. Scarcity of means to satisfy given ends is an almost ubiquitous condition of human nature.

Because people live in a world of scarcity, they cannot have all the time, money, possessions, and experiences they wish. Neither can society.

This chapter will continue our discussion of scarcity and the economic way of thinking by first introducing three critical concepts: opportunity cost, marginal decision making, and diminishing returns. Later, it will consider whether the economic way of thinking accurately describes either how we *make* choices and how we *should* make them.

2.1 | How Individuals Make Choices Based on Their Budget Constraint

By the end of this section, you will be able to:

- Calculate and graph budget constraints
- Explain opportunity sets and opportunity costs
- Evaluate the law of diminishing marginal utility
- Explain how marginal analysis and utility influence choices

Consider the typical consumer's budget problem. Consumers have a limited amount of income to spend on the things they need and want. Suppose Alphonso has $10 in spending money each week that he can allocate between bus tickets for getting to work and the burgers that he eats for lunch. Burgers cost $2 each, and bus tickets are 50 cents each. We can see Alphonso's budget problem in **Figure 2.2**.

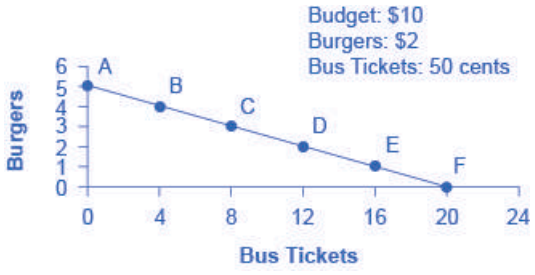


Figure 2.2 The Budget Constraint: Alphonso's Consumption Choice Opportunity Frontier Each point on the budget constraint represents a combination of burgers and bus tickets whose total cost adds up to Alphonso's budget of $10. The relative price of burgers and bus tickets determines the slope of the budget constraint. All along the budget set, giving up one burger means gaining four bus tickets.

The vertical axis in the figure shows burger purchases and the horizontal axis shows bus ticket purchases. If Alphonso spends all his money on burgers, he can afford five per week. ($10 per week/$2 per burger = 5 burgers per week.) However, if he does this, he will not be able to afford any bus tickets. Point A in the figure shows the choice (zero bus tickets and five burgers). Alternatively, if Alphonso spends all his money on bus tickets, he can afford 20 per week. ($10 per week/$0.50 per bus ticket = 20 bus tickets per week.) Then, however, he will not be able to afford any burgers. Point F shows this alternative choice (20 bus tickets and zero burgers).

If we connect all the points between A and F, we get Alphonso's **budget constraint**. This indicates all the combination of burgers and bus tickets Alphonso can afford, given the price of the two goods and his budget amount.

If Alphonso is like most people, he will choose some combination that includes both bus tickets and burgers. That is, he will choose some combination on the budget constraint that is between points A and F. Every point on (or inside) the constraint shows a combination of burgers and bus tickets that Alphonso can afford. Any point outside the constraint is not affordable, because it would cost more money than Alphonso has in his budget.

The budget constraint clearly shows the tradeoff Alphonso faces in choosing between burgers and bus tickets. Suppose he is currently at point D, where he can afford 12 bus tickets and two burgers. What would it cost Alphonso for one more burger? It would be natural to answer $2, but that's not the way economists think. Instead they ask, how many bus tickets would Alphonso have to give up to get one more burger, while staying within his budget? Since bus tickets cost 50 cents, Alphonso would have to give up four to afford one more burger. That is the true cost to Alphonso.

The Concept of Opportunity Cost

Economists use the term **opportunity cost** to indicate what one must give up to obtain what he or she desires. The idea behind opportunity cost is that the cost of one item is the lost opportunity to do or consume something else. In short, opportunity cost is the value of the next best alternative. For Alphonso, the opportunity cost of a burger is the four bus tickets he would have to give up. He would decide whether or not to choose the burger depending on whether the value of the burger exceeds the value of the forgone alternative—in this case, bus tickets. Since people must choose, they inevitably face tradeoffs in which they have to give up things they desire to obtain other things they desire more.

Link It Up

View this website (http://openstaxcollege.org/l/linestanding) for an example of opportunity cost—paying someone else to wait in line for you.

A fundamental principle of economics is that every choice has an opportunity cost. If you sleep through your economics class, the opportunity cost is the learning you miss from not attending class. If you spend your income on video games, you cannot spend it on movies. If you choose to marry one person, you give up the opportunity to marry anyone else. In short, opportunity cost is all around us and part of human existence.

The following Work It Out feature shows a step-by-step analysis of a budget constraint calculation. Read through it to understand another important concept—slope—that we further explain in the appendix **The Use of Mathematics in Principles of Economics**.

Work It Out

Understanding Budget Constraints

Budget constraints are easy to understand if you apply a little math. The appendix **The Use of Mathematics in Principles of Economics** explains all the math you are likely to need in this book. Therefore, if math is not your strength, you might want to take a look at the appendix.

Step 1: The equation for any budget constraint is:

$$\text{Budget} = P_1 \times Q_1 + P_2 \times Q_2$$

where P and Q are the price and quantity of items purchased (which we assume here to be two items) and Budget is the amount of income one has to spend.

Step 2. Apply the budget constraint equation to the scenario. In Alphonso's case, this works out to be:

$$\begin{aligned} \text{Budget} &= P_1 \times Q_1 + P_2 \times Q_2 \\ \$10 \text{ budget} &= \$2 \text{ per burger} \times \text{quantity of burgers} + \$0.50 \text{ per bus ticket} \times \text{quantity of bus tickets} \\ \$10 &= \$2 \times Q_{\text{burgers}} + \$0.50 \times Q_{\text{bus tickets}} \end{aligned}$$

Step 3. Using a little algebra, we can turn this into the familiar equation of a line:

$$y = b + mx$$

For Alphonso, this is:

$$\$10 = \$2 \times Q_{\text{burgers}} + \$0.50 \times Q_{\text{bus tickets}}$$

Step 4. Simplify the equation. Begin by multiplying both sides of the equation by 2:

$$\begin{aligned} 2 \times 10 &= 2 \times 2 \times Q_{\text{burgers}} + 2 \times 0.5 \times Q_{\text{bus tickets}} \\ 20 &= 4 \times Q_{\text{burgers}} + 1 \times Q_{\text{bus tickets}} \end{aligned}$$

Step 5. Subtract one bus ticket from both sides:

$$20 - Q_{\text{bus tickets}} = 4 \times Q_{\text{burgers}}$$

Divide each side by 4 to yield the answer:

$$\begin{aligned} 5 - 0.25 \times Q_{\text{bus tickets}} &= Q_{\text{burgers}} \\ \text{or} \\ Q_{\text{burgers}} &= 5 - 0.25 \times Q_{\text{bus tickets}} \end{aligned}$$

Step 6. Notice that this equation fits the budget constraint in **Figure 2.2**. The vertical intercept is 5 and the slope is –0.25, just as the equation says. If you plug 20 bus tickets into the equation, you get 0 burgers. If you plug other numbers of bus tickets into the equation, you get the results (see **Table 2.1**), which are the points on Alphonso's budget constraint.

Point	Quantity of Burgers (at $2)	Quantity of Bus Tickets (at 50 cents)
A	5	0
B	4	4
C	3	8
D	2	12

Table 2.1

Point	Quantity of Burgers (at $2)	Quantity of Bus Tickets (at 50 cents)
E	1	16
F	0	20

Table 2.1

Step 7. Notice that the slope of a budget constraint always shows the opportunity cost of the good which is on the horizontal axis. For Alphonso, the slope is −0.25, indicating that for every bus ticket he buys, he must give up 1/4 burger. To phrase it differently, for every four tickets he buys, Alphonso must give up 1 burger.

There are two important observations here. First, the algebraic sign of the slope is negative, which means that the only way to get more of one good is to give up some of the other. Second, we define the slope as the price of bus tickets (whatever is on the horizontal axis in the graph) divided by the price of burgers (whatever is on the vertical axis), in this case $0.50/$2 = 0.25. If you want to determine the opportunity cost quickly, just divide the two prices.

Identifying Opportunity Cost

In many cases, it is reasonable to refer to the opportunity cost as the price. If your cousin buys a new bicycle for $300, then $300 measures the amount of "other consumption" that he has forsaken. For practical purposes, there may be no special need to identify the specific alternative product or products that he could have bought with that $300, but sometimes the price as measured in dollars may not accurately capture the true opportunity cost. This problem can loom especially large when costs of time are involved.

For example, consider a boss who decides that all employees will attend a two-day retreat to "build team spirit." The out-of-pocket monetary cost of the event may involve hiring an outside consulting firm to run the retreat, as well as room and board for all participants. However, an opportunity cost exists as well: during the two days of the retreat, none of the employees are doing any other work.

Attending college is another case where the opportunity cost exceeds the monetary cost. The out-of-pocket costs of attending college include tuition, books, room and board, and other expenses. However, in addition, during the hours that you are attending class and studying, it is impossible to work at a paying job. Thus, college imposes both an out-of-pocket cost and an opportunity cost of lost earnings.

Clear It Up

What is the opportunity cost associated with increased airport security measures?

After the terrorist plane hijackings on September 11, 2001, many steps were proposed to improve air travel safety. For example, the federal government could provide armed "sky marshals" who would travel inconspicuously with the rest of the passengers. The cost of having a sky marshal on every flight would be roughly $3 billion per year. Retrofitting all U.S. planes with reinforced cockpit doors to make it harder for terrorists to take over the plane would have a price tag of $450 million. Buying more sophisticated security equipment for airports, like three-dimensional baggage scanners and cameras linked to face recognition software, could cost another $2 billion.

However, the single biggest cost of greater airline security does not involve spending money. It is the opportunity cost of additional waiting time at the airport. According to the United States Department of Transportation (DOT), there were 895.5 million systemwide (domestic and international) scheduled service passengers in 2015. Since the 9/11 hijackings, security screening has become more intensive, and consequently, the procedure takes longer than in the past. Say that, on average, each air passenger spends

an extra 30 minutes in the airport per trip. Economists commonly place a value on time to convert an opportunity cost in time into a monetary figure. Because many air travelers are relatively high-paid business people, conservative estimates set the average price of time for air travelers at $20 per hour. By these back-of-the-envelope calculations, the opportunity cost of delays in airports could be as much as 800 million × 0.5 hours × $20/hour, or $8 billion per year. Clearly, the opportunity costs of waiting time can be just as important as costs that involve direct spending.

In some cases, realizing the opportunity cost can alter behavior. Imagine, for example, that you spend $8 on lunch every day at work. You may know perfectly well that bringing a lunch from home would cost only $3 a day, so the opportunity cost of buying lunch at the restaurant is $5 each day (that is, the $8 buying lunch costs minus the $3 your lunch from home would cost). Five dollars each day does not seem to be that much. However, if you project what that adds up to in a year—250 days a year × $5 per day equals $1,250, the cost, perhaps, of a decent vacation. If you describe the opportunity cost as "a nice vacation" instead of "$5 a day," you might make different choices.

Marginal Decision-Making and Diminishing Marginal Utility

The budget constraint framework helps to emphasize that most choices in the real world are not about getting all of one thing or all of another; that is, they are not about choosing either the point at one end of the budget constraint or else the point all the way at the other end. Instead, most choices involve **marginal analysis**, which means examining the benefits and costs of choosing a little more or a little less of a good. People naturally compare costs and benefits, but often we look at total costs and total benefits, when the optimal choice necessitates comparing how costs and benefits change from one option to another. You might think of marginal analysis as "change analysis." Marginal analysis is used throughout economics.

We now turn to the notion of **utility**. People desire goods and services for the satisfaction or utility those goods and services provide. Utility, as we will see in the chapter on **Consumer Choices (http://cnx.org/content/m63895/latest/)** , is subjective but that does not make it less real. Economists typically assume that the more of some good one consumes (for example, slices of pizza), the more utility one obtains. At the same time, the utility a person receives from consuming the first unit of a good is typically more than the utility received from consuming the fifth or the tenth unit of that same good. When Alphonso chooses between burgers and bus tickets, for example, the first few bus rides that he chooses might provide him with a great deal of utility—perhaps they help him get to a job interview or a doctor's appointment. However, later bus rides might provide much less utility—they may only serve to kill time on a rainy day. Similarly, the first burger that Alphonso chooses to buy may be on a day when he missed breakfast and is ravenously hungry. However, if Alphonso has a burger every single day, the last few burgers may taste pretty boring. The general pattern that consumption of the first few units of any good tends to bring a higher level of utility to a person than consumption of later units is a common pattern. Economists refer to this pattern as the **law of diminishing marginal utility**, which means that as a person receives more of a good, the additional (or marginal) utility from each additional unit of the good declines. In other words, the first slice of pizza brings more satisfaction than the sixth.

The law of diminishing marginal utility explains why people and societies rarely make all-or-nothing choices. You would not say, "My favorite food is ice cream, so I will eat nothing but ice cream from now on." Instead, even if you get a very high level of utility from your favorite food, if you ate it exclusively, the additional or marginal utility from those last few servings would not be very high. Similarly, most workers do not say: "I enjoy leisure, so I'll never work." Instead, workers recognize that even though some leisure is very nice, a combination of all leisure and no income is not so attractive. The budget constraint framework suggests that when people make choices in a world of scarcity, they will use marginal analysis and think about whether they would prefer a little more or a little less.

A rational consumer would only purchase additional units of some product as long as the marginal utility exceeds the opportunity cost. Suppose Alphonso moves down his budget constraint from Point A to Point B to Point C and further. As he consumes more bus tickets, the marginal utility of bus tickets will diminish, while the opportunity cost, that is, the marginal utility of foregone burgers, will increase. Eventually, the opportunity cost will exceed the marginal utility of an additional bus ticket. If Alphonso is rational, he won't purchase more bus tickets once the marginal utility just equals the opportunity cost. While we can't (yet) say exactly how many bus tickets Alphonso will buy, that number is unlikely to be the most he can afford, 20.

Sunk Costs

In the budget constraint framework, all decisions involve what will happen next: that is, what quantities of goods will

you consume, how many hours will you work, or how much will you save. These decisions do not look back to past choices. Thus, the budget constraint framework assumes that **sunk costs**, which are costs that were incurred in the past and cannot be recovered, should not affect the current decision.

Consider the case of Selena, who pays $8 to see a movie, but after watching the film for 30 minutes, she knows that it is truly terrible. Should she stay and watch the rest of the movie because she paid for the ticket, or should she leave? The money she spent is a sunk cost, and unless the theater manager is sympathetic, Selena will not get a refund. However, staying in the movie still means paying an opportunity cost in time. Her choice is whether to spend the next 90 minutes suffering through a cinematic disaster or to do something—anything—else. The lesson of sunk costs is to forget about the money and time that is irretrievably gone and instead to focus on the marginal costs and benefits of current and future options.

For people and firms alike, dealing with sunk costs can be frustrating. It often means admitting an earlier error in judgment. Many firms, for example, find it hard to give up on a new product that is doing poorly because they spent so much money in creating and launching the product. However, the lesson of sunk costs is to ignore them and make decisions based on what will happen in the future.

From a Model with Two Goods to One of Many Goods

The budget constraint diagram containing just two goods, like most models used in this book, is not realistic. After all, in a modern economy people choose from thousands of goods. However, thinking about a model with many goods is a straightforward extension of what we discussed here. Instead of drawing just one budget constraint, showing the tradeoff between two goods, you can draw multiple budget constraints, showing the possible tradeoffs between many different pairs of goods. In more advanced classes in economics, you would use mathematical equations that include many possible goods and services that can be purchased, together with their quantities and prices, and show how the total spending on all goods and services is limited to the overall budget available. The graph with two goods that we presented here clearly illustrates that every choice has an opportunity cost, which is the point that does carry over to the real world.

2.2 | The Production Possibilities Frontier and Social Choices

By the end of this section, you will be able to:

- Interpret production possibilities frontier graphs
- Contrast a budget constraint and a production possibilities frontier
- Explain the relationship between a production possibilities frontier and the law of diminishing returns
- Contrast productive efficiency and allocative efficiency
- Define comparative advantage

Just as individuals cannot have everything they want and must instead make choices, society as a whole cannot have everything it might want, either. This section of the chapter will explain the constraints society faces, using a model called the **production possibilities frontier (PPF)**. There are more similarities than differences between individual choice and social choice. As you read this section, focus on the similarities.

Because society has limited resources (e.g., labor, land, capital, raw materials) at any point in time, there is a limit to the quantities of goods and services it can produce. Suppose a society desires two products, healthcare and education. The production possibilities frontier in **Figure 2.3** illustrates this situation.

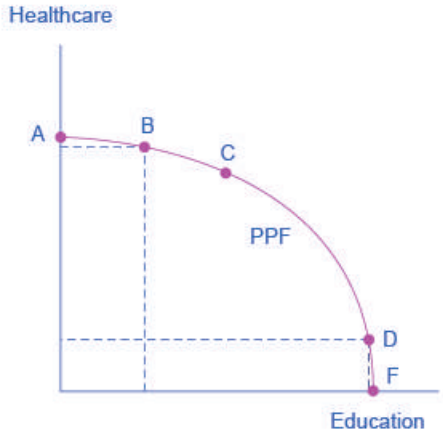


Figure 2.3 A Healthcare vs. Education Production Possibilities Frontier This production possibilities frontier shows a tradeoff between devoting social resources to healthcare and devoting them to education. At A all resources go to healthcare and at B, most go to healthcare. At D most resources go to education, and at F, all go to education.

Figure 2.3 shows healthcare on the vertical axis and education on the horizontal axis. If the society were to allocate all of its resources to healthcare, it could produce at point A. However, it would not have any resources to produce education. If it were to allocate all of its resources to education, it could produce at point F. Alternatively, the society could choose to produce any combination of healthcare and education on the production possibilities frontier. In effect, the production possibilities frontier plays the same role for society as the budget constraint plays for Alphonso. Society can choose any combination of the two goods on or inside the PPF. However, it does not have enough resources to produce outside the PPF.

Most importantly, the production possibilities frontier clearly shows the tradeoff between healthcare and education. Suppose society has chosen to operate at point B, and it is considering producing more education. Because the PPF is downward sloping from left to right, the only way society can obtain more education is by giving up some healthcare. That is the tradeoff society faces. Suppose it considers moving from point B to point C. What would the opportunity cost be for the additional education? The opportunity cost would be the healthcare society has to forgo. Just as with Alphonso's budget constraint, the **slope** of the production possibilities frontier shows the opportunity cost. By now you might be saying, "Hey, this PPF is sounding like the budget constraint." If so, read the following Clear It Up feature.

Clear It Up

What's the difference between a budget constraint and a PPF?

There are two major differences between a budget constraint and a production possibilities frontier. The first is the fact that the budget constraint is a straight line. This is because its slope is given by the relative prices of the two goods, which from the point of view of an individual consumer, are fixed, so the slope doesn't change. In contrast, the PPF has a curved shape because of the law of the diminishing returns. Thus, the slope is different at various points on the PPF. The second major difference is the absence of specific numbers on the axes of the PPF. There are no specific numbers because we do not know the exact amount of resources this imaginary economy has, nor do we know how many resources it takes to produce healthcare and how many resources it takes to produce education. If this were a real world example, that data would be available.

Whether or not we have specific numbers, conceptually we can measure the opportunity cost of additional education as society moves from point B to point C on the PPF. We measure the additional education by the horizontal distance between B and C. The foregone healthcare is given by the vertical distance between B and C. The slope of the PPF between B and C is (approximately) the vertical distance (the "rise") over the horizontal distance (the "run"). This is the opportunity cost of the additional education.

The Shape of the PPF and the Law of Diminishing Returns

The budget constraints that we presented earlier in this chapter, showing individual choices about what quantities of goods to consume, were all straight lines. The reason for these straight lines was that the relative prices of the two goods in the **consumption budget constraint** determined the slope of the budget constraint. However, we drew the production possibilities frontier for healthcare and education as a curved line. Why does the PPF have a different shape?

To understand why the PPF is curved, start by considering point A at the top left-hand side of the PPF. At point A, all available resources are devoted to healthcare and none are left for education. This situation would be extreme and even ridiculous. For example, children are seeing a doctor every day, whether they are sick or not, but not attending school. People are having cosmetic surgery on every part of their bodies, but no high school or college education exists. Now imagine that some of these resources are diverted from healthcare to education, so that the economy is at point B instead of point A. Diverting some resources away from A to B causes relatively little reduction in health because the last few marginal dollars going into healthcare services are not producing much additional gain in health. However, putting those marginal dollars into education, which is completely without resources at point A, can produce relatively large gains. For this reason, the shape of the PPF from A to B is relatively flat, representing a relatively small drop-off in health and a relatively large gain in education.

Now consider the other end, at the lower right, of the production possibilities frontier. Imagine that society starts at choice D, which is devoting nearly all resources to education and very few to healthcare, and moves to point F, which is devoting *all* spending to education and none to healthcare. For the sake of concreteness, you can imagine that in the movement from D to F, the last few doctors must become high school science teachers, the last few nurses must become school librarians rather than dispensers of vaccinations, and the last few emergency rooms are turned into kindergartens. The gains to education from adding these last few resources to education are very small. However, the opportunity cost lost to health will be fairly large, and thus the slope of the PPF between D and F is steep, showing a large drop in health for only a small gain in education.

The lesson is not that society is likely to make an extreme choice like devoting no resources to education at point A or no resources to health at point F. Instead, the lesson is that the gains from committing additional marginal resources to education depend on how much is already being spent. If on the one hand, very few resources are currently committed to education, then an increase in resources used can bring relatively large gains. On the other hand, if a large number of resources are already committed to education, then committing additional resources will bring relatively smaller gains.

This pattern is common enough that economists have given it a name: the **law of diminishing returns**, which holds that as additional increments of resources are added to a certain purpose, the marginal benefit from those additional increments will decline. (The law of diminishing marginal utility that we introduced in the last section is a more specific case of the law of diminishing returns.) When government spends a certain amount more on reducing crime, for example, the original gains in reducing crime could be relatively large. However, additional increases typically cause relatively smaller reductions in crime, and paying for enough police and security to reduce crime to nothing at all would be tremendously expensive.

The curvature of the production possibilities frontier shows that as we add more resources to education, moving from left to right along the horizontal axis, the original gains are fairly large, but gradually diminish. Thus, the slope of the PPF is relatively flat. By contrast, as we add more resources to healthcare, moving from bottom to top on the vertical axis, the original gains are fairly large, but again gradually diminish. Thus, the slope of the PPF is relatively steep. In this way, the law of diminishing returns produces the outward-bending shape of the production possibilities frontier.

Productive Efficiency and Allocative Efficiency

The study of economics does not presume to tell a society what choice it should make along its production possibilities frontier. In a market-oriented economy with a democratic government, the choice will involve a mixture of decisions by individuals, firms, and government. However, economics can point out that some choices are unambiguously better than others. This observation is based on the concept of efficiency. In everyday usage, efficiency refers to lack of waste. An inefficient machine operates at high cost, while an efficient machine operates at lower cost, because it is not wasting energy or materials. An inefficient organization operates with long delays and high costs, while an efficient organization meets schedules, is focused, and performs within budget.

The production possibilities frontier can illustrate two kinds of efficiency: productive efficiency and allocative efficiency. **Figure 2.4** illustrates these ideas using a production possibilities frontier between healthcare and

education.

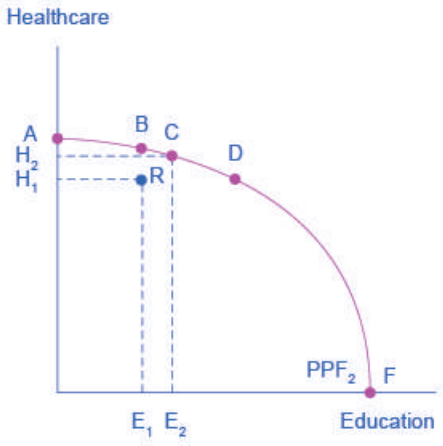


Figure 2.4 Productive and Allocative Efficiency Productive efficiency means it is impossible to produce more of one good without decreasing the quantity that is produced of another good. Thus, all choices along a given PPF like B, C, and D display productive efficiency, but R does not. Allocative efficiency means that the particular mix of goods being produced—that is, the specific choice along the production possibilities frontier—represents the allocation that society most desires.

Productive efficiency means that, given the available inputs and technology, it is impossible to produce more of one good without decreasing the quantity that is produced of another good. All choices on the PPF in **Figure 2.4**, including A, B, C, D, and F, display productive efficiency. As a firm moves from any one of these choices to any other, either healthcare increases and education decreases or vice versa. However, any choice inside the production possibilities frontier is productively inefficient and wasteful because it is possible to produce more of one good, the other good, or some combination of both goods.

For example, point R is productively inefficient because it is possible at choice C to have more of both goods: education on the horizontal axis is higher at point C than point R (E_2 is greater than E_1), and healthcare on the vertical axis is also higher at point C than point R (H_2 is great than H_1).

We can show the particular mix of goods and services produced—that is, the specific combination of selected healthcare and education along the production possibilities frontier—as a ray (line) from the origin to a specific point on the PPF. Output mixes that had more healthcare (and less education) would have a steeper ray, while those with more education (and less healthcare) would have a flatter ray.

Allocative efficiency means that the particular combination of goods and services on the production possibility curve that a society produces represents the combination that society most desires. How to determine what a society desires can be a controversial question, and is usually a discussion in political science, sociology, and philosophy classes as well as in economics. At its most basic, allocative efficiency means producers supply the quantity of each product that consumers demand. Only one of the productively efficient choices will be the allocatively efficient choice for society as a whole.

Why Society Must Choose

In **Welcome to Economics!** we learned that every society faces the problem of scarcity, where limited resources conflict with unlimited needs and wants. The production possibilities curve illustrates the choices involved in this dilemma.

Every economy faces two situations in which it may be able to expand consumption of all goods. In the first case, a society may discover that it has been using its resources inefficiently, in which case by improving efficiency and producing on the production possibilities frontier, it can have more of all goods (or at least more of some and less of none). In the second case, as resources grow over a period of years (e.g., more labor and more capital), the economy grows. As it does, the production possibilities frontier for a society will tend to shift outward and society will be able to afford more of all goods.

However, improvements in productive efficiency take time to discover and implement, and economic growth happens only gradually. Thus, a society must choose between tradeoffs in the present. For government, this process often involves trying to identify where additional spending could do the most good and where reductions in spending would do the least harm. At the individual and firm level, the market economy coordinates a process in which firms seek to produce goods and services in the quantity, quality, and price that people want. However, for both the government and the market economy in the short term, increases in production of one good typically mean offsetting decreases somewhere else in the economy.

The PPF and Comparative Advantage

While every society must choose how much of each good or service it should produce, it does not need to produce every single good it consumes. Often how much of a good a country decides to produce depends on how expensive it is to produce it versus buying it from a different country. As we saw earlier, the curvature of a country's PPF gives us information about the tradeoff between devoting resources to producing one good versus another. In particular, its slope gives the opportunity cost of producing one more unit of the good in the x-axis in terms of the other good (in the y-axis). Countries tend to have different opportunity costs of producing a specific good, either because of different climates, geography, technology, or skills.

Suppose two countries, the US and Brazil, need to decide how much they will produce of two crops: sugar cane and wheat. Due to its climatic conditions, Brazil can produce quite a bit of sugar cane per acre but not much wheat. Conversely, the U.S. can produce large amounts of wheat per acre, but not much sugar cane. Clearly, Brazil has a lower opportunity cost of producing sugar cane (in terms of wheat) than the U.S. The reverse is also true: the U.S. has a lower opportunity cost of producing wheat than Brazil. We illustrate this by the PPFs of the two countries in **Figure 2.5**.

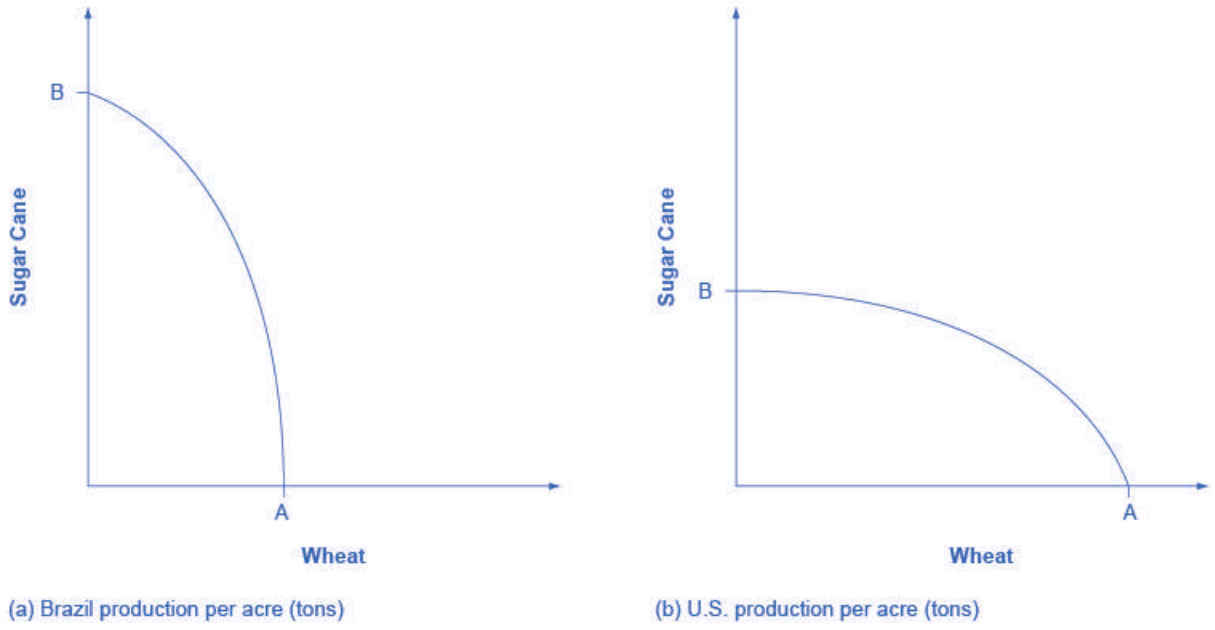


Figure 2.5 Production Possibility Frontier for the U.S. and Brazil The U.S. PPF is flatter than the Brazil PPF implying that the opportunity cost of wheat in terms of sugar cane is lower in the U.S. than in Brazil. Conversely, the opportunity cost of sugar cane is lower in Brazil. The U.S. has comparative advantage in wheat and Brazil has comparative advantage in sugar cane.

When a country can produce a good at a lower opportunity cost than another country, we say that this country has a **comparative advantage** in that good. Comparative advantage is not the same as absolute advantage, which is when a country can produce more of a good. Comparative advantage is not the same as absolute advantage, which is when a country can produce more of a good. In our example, Brazil has an absolute advantage in sugar cane and the U.S. has an absolute advantage in wheat. One can easily see this with a simple observation of the extreme production points in the PPFs of the two countries. If Brazil devoted all of its resources to producing wheat, it would be producing at point A. If however it had devoted all of its resources to producing sugar cane instead, it would be producing a much larger

amount than the U.S., at point B.

The slope of the PPF gives the opportunity cost of producing an additional unit of wheat. While the slope is not constant throughout the PPFs, it is quite apparent that the PPF in Brazil is much steeper than in the U.S., and therefore the opportunity cost of wheat generally higher in Brazil. In the chapter on **International Trade** you will learn that countries' differences in comparative advantage determine which goods they will choose to produce and trade. When countries engage in trade, they specialize in the production of the goods in which they have comparative advantage, and trade part of that production for goods in which they do not have comparative advantage. With trade, manufacturers produce goods where the opportunity cost is lowest, so total production increases, benefiting both trading parties.

2.3 | Confronting Objections to the Economic Approach

By the end of this section, you will be able to:

- Analyze arguments against economic approaches to decision-making
- Interpret a tradeoff diagram
- Contrast normative statements and positive statements

It is one thing to understand the economic approach to decision-making and another thing to feel comfortable applying it. The sources of discomfort typically fall into two categories: that people do not act in the way that fits the economic way of thinking, and that even if people did act that way, they should try not to. Let's consider these arguments in turn.

First Objection: People, Firms, and Society Do Not Act Like This

The economic approach to decision-making seems to require more information than most individuals possess and more careful decision-making than most individuals actually display. After all, do you or any of your friends draw a budget constraint and mutter to yourself about maximizing utility before you head to the shopping mall? Do members of the U.S. Congress contemplate production possibilities frontiers before they vote on the annual budget? The messy ways in which people and societies operate somehow doesn't look much like neat budget constraints or smoothly curving production possibilities frontiers.

However, the economics approach can be a useful way to analyze and understand the tradeoffs of economic decisions. To appreciate this point, imagine for a moment that you are playing basketball, dribbling to the right, and throwing a bounce-pass to the left to a teammate who is running toward the basket. A physicist or engineer could work out the correct speed and trajectory for the pass, given the different movements involved and the weight and bounciness of the ball. However, when you are playing basketball, you do not perform any of these calculations. You just pass the ball, and if you are a good player, you will do so with high accuracy.

Someone might argue: "The scientist's formula of the bounce-pass requires a far greater knowledge of physics and far more specific information about speeds of movement and weights than the basketball player actually has, so it must be an unrealistic description of how basketball passes actually occur." This reaction would be wrongheaded. The fact that a good player can throw the ball accurately because of practice and skill, without making a physics calculation, does not mean that the physics calculation is wrong.

Similarly, from an economic point of view, someone who shops for groceries every week has a great deal of practice with how to purchase the combination of goods that will provide that person with utility, even if the shopper does not phrase decisions in terms of a budget constraint. Government institutions may work imperfectly and slowly, but in general, a democratic form of government feels pressure from voters and social institutions to make the choices that are most widely preferred by people in that society. Thus, when thinking about the economic actions of groups of people, firms, and society, it is reasonable, as a first approximation, to analyze them with the tools of economic analysis. For more on this, read about behavioral economics in the chapter on **Consumer Choices (http://cnx.org/content/m63895/latest/)** .

Second Objection: People, Firms, and Society Should Not Act This Way

The economics approach portrays people as self-interested. For some critics of this approach, even if self-interest is an accurate description of how people behave, these behaviors are not moral. Instead, the critics argue that people

should be taught to care more deeply about others. Economists offer several answers to these concerns.

First, economics is not a form of moral instruction. Rather, it seeks to describe economic behavior as it actually exists. Philosophers draw a distinction between **positive statements**, which describe the world as it is, and **normative statements**, which describe how the world should be. Positive statements are factual. They may be true or false, but we can test them, at least in principle. Normative statements are subjective questions of opinion. We cannot test them since we cannot prove opinions to be true or false. They just are opinions based on one's values. For example, an economist could analyze a proposed subway system in a certain city. If the expected benefits exceed the costs, he concludes that the project is worthy—an example of positive analysis. Another economist argues for extended unemployment compensation during the Great Depression because a rich country like the United States should take care of its less fortunate citizens—an example of normative analysis.

Even if the line between positive and normative statements is not always crystal clear, economic analysis does try to remain rooted in the study of the actual people who inhabit the actual economy. Fortunately however, the assumption that individuals are purely self-interested is a simplification about human nature. In fact, we need to look no further than to Adam Smith, the very father of modern economics to find evidence of this. The opening sentence of his book, *The Theory of Moral Sentiments*, puts it very clearly: "How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it except the pleasure of seeing it." Clearly, individuals are both self-interested and altruistic.

Second, we can label self-interested behavior and profit-seeking with other names, such as personal choice and freedom. The ability to make personal choices about buying, working, and saving is an important personal freedom. Some people may choose high-pressure, high-paying jobs so that they can earn and spend considerable amounts of money on themselves. Others may allocate large portions of their earnings to charity or spend it on their friends and family. Others may devote themselves to a career that can require much time, energy, and expertise but does not offer high financial rewards, like being an elementary school teacher or a social worker. Still others may choose a job that does consume much of their time or provide a high level of income, but still leaves time for family, friends, and contemplation. Some people may prefer to work for a large company; others might want to start their own business. People's freedom to make their own economic choices has a moral value worth respecting.

Clear It Up

Is a diagram by any other name the same?

When you study economics, you may feel buried under an avalanche of diagrams. Your goal should be to recognize the common underlying logic and pattern of the diagrams, not to memorize each one.

This chapter uses only one basic diagram, although we present it with different sets of labels. The consumption budget constraint and the production possibilities frontier for society, as a whole, are the same basic diagram. **Figure 2.6** shows an individual budget constraint and a production possibilities frontier for two goods, Good 1 and Good 2. The tradeoff diagram always illustrates three basic themes: scarcity, tradeoffs, and economic efficiency.

The first theme is scarcity. It is not feasible to have unlimited amounts of both goods. Even if the budget constraint or a PPF shifts, scarcity remains—just at a different level. The second theme is tradeoffs. As depicted in the budget constraint or the production possibilities frontier, it is necessary to forgo some of one good to gain more of the other good. The details of this tradeoff vary. In a budget constraint we determine, the tradeoff is determined by the relative prices of the goods: that is, the relative price of two goods in the consumption choice budget constraint. These tradeoffs appear as a straight line. However, a curved line represents the tradeoffs in many production possibilities frontiers because the law of diminishing returns holds that as we add resources to an area, the marginal gains tend to diminish. Regardless of the specific shape, tradeoffs remain.

The third theme is economic efficiency, or getting the most benefit from scarce resources. All choices on the production possibilities frontier show productive efficiency because in such cases, there is no way to increase the quantity of one good without decreasing the quantity of the other. Similarly, when an individual makes a

choice along a budget constraint, there is no way to increase the quantity of one good without decreasing the quantity of the other. The choice on a production possibilities set that is socially preferred, or the choice on an individual's budget constraint that is personally preferred, will display allocative efficiency.

The basic budget constraint/production possibilities frontier diagram will recur throughout this book. Some examples include using these tradeoff diagrams to analyze trade, environmental protection and economic output, equality of incomes and economic output, and the macroeconomic tradeoff between consumption and investment. Do not allow the different labels to confuse you. The budget constraint/production possibilities frontier diagram is always just a tool for thinking carefully about scarcity, tradeoffs, and efficiency in a particular situation.

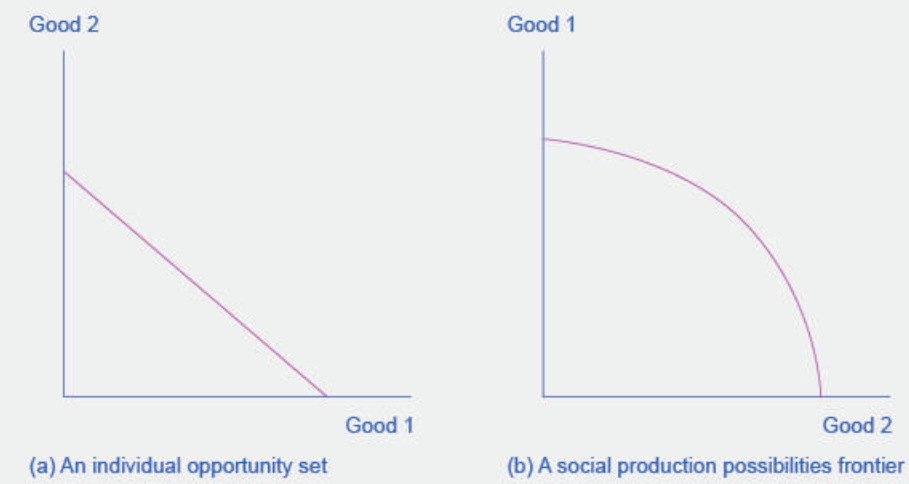


Figure 2.6 The Tradeoff Diagram Both the individual opportunity set (or budget constraint) and the social production possibilities frontier show the constraints under which individual consumers and society as a whole operate. Both diagrams show the tradeoff in choosing more of one good at the cost of less of the other.

Third, self-interested behavior can lead to positive social results. For example, when people work hard to make a living, they create economic output. Consumers who are looking for the best deals will encourage businesses to offer goods and services that meet their needs. Adam Smith, writing in *The Wealth of Nations*, named this property the **invisible hand**. In describing how consumers and producers interact in a market economy, Smith wrote:

> Every individual…generally, indeed, neither intends to promote the public interest, nor knows how much he is promoting it. By preferring the support of domestic to that of foreign industry, he intends only his own security; and by directing that industry in such a manner as its produce may be of the greatest value, he intends only his own gain. And he is in this, as in many other cases, led by an invisible hand to promote an end which was no part of his intention…By pursuing his own interest he frequently promotes that of the society more effectually than when he really intends to promote it.

The metaphor of the invisible hand suggests the remarkable possibility that broader social good can emerge from selfish individual actions.

Fourth, even people who focus on their own self-interest in the economic part of their life often set aside their own narrow self-interest in other parts of life. For example, you might focus on your own self-interest when asking your employer for a raise or negotiating to buy a car. Then you might turn around and focus on other people when you volunteer to read stories at the local library, help a friend move to a new apartment, or donate money to a charity. Self-interest is a reasonable starting point for analyzing many economic decisions, without needing to imply that people never do anything that is not in their own immediate self-interest.

Bring it Home

Choices ... To What Degree?

What have we learned? We know that scarcity impacts all the choices we make. An economist might argue that people do not obtain a bachelor's or master's degree because they do not have the resources to make those choices or because their incomes are too low and/or the price of these degrees is too high. A bachelor's or a master's degree may not be available in their opportunity set.

The price of these degrees may be too high not only because the actual price, college tuition (and perhaps room and board), is too high. An economist might also say that for many people, the full opportunity cost of a bachelor's or a master's degree is too high. For these people, they are unwilling or unable to make the tradeoff of forfeiting years of working, and earning an income, to earn a degree.

Finally, the statistics we introduced at the start of the chapter reveal information about intertemporal choices. An economist might say that people choose not to obtain a college degree because they may have to borrow money to attend college, and the interest they have to pay on that loan in the future will affect their decisions today. Also, it could be that some people have a preference for current consumption over future consumption, so they choose to work now at a lower salary and consume now, rather than postponing that consumption until after they graduate college.

KEY TERMS

allocative efficiency when the mix of goods produced represents the mix that society most desires

budget constraint all possible consumption combinations of goods that someone can afford, given the prices of goods, when all income is spent; the boundary of the opportunity set

comparative advantage when a country can produce a good at a lower cost in terms of other goods; or, when a country has a lower opportunity cost of production

invisible hand Adam Smith's concept that individuals' self-interested behavior can lead to positive social outcomes

law of diminishing marginal utility as we consume more of a good or service, the utility we get from additional units of the good or service tends to become smaller than what we received from earlier units

law of diminishing returns as we add additional increments of resources to producing a good or service, the marginal benefit from those additional increments will decline

marginal analysis examination of decisions on the margin, meaning a little more or a little less from the status quo

normative statement statement which describes how the world should be

opportunity cost measures cost by what we give up/forfeit in exchange; opportunity cost measures the value of the forgone alternative

opportunity set all possible combinations of consumption that someone can afford given the prices of goods and the individual’s income

positive statement statement which describes the world as it is

production possibilities frontier (PPF) a diagram that shows the productively efficient combinations of two products that an economy can produce given the resources it has available.

productive efficiency when it is impossible to produce more of one good (or service) without decreasing the quantity produced of another good (or service)

sunk costs costs that we make in the past that we cannot recover

utility satisfaction, usefulness, or value one obtains from consuming goods and services

KEY CONCEPTS AND SUMMARY

2.1 How Individuals Make Choices Based on Their Budget Constraint

Economists see the real world as one of scarcity: that is, a world in which people’s desires exceed what is possible. As a result, economic behavior involves tradeoffs in which individuals, firms, and society must forgo something that they desire to obtain things that they desire more. Individuals face the tradeoff of what quantities of goods and services to consume. The budget constraint, which is the frontier of the opportunity set, illustrates the range of available choices. The relative price of the choices determines the slope of the budget constraint. Choices beyond the budget constraint are not affordable.

Opportunity cost measures cost by what we forgo in exchange. Sometimes we can measure opportunity cost in money, but it is often useful to consider time as well, or to measure it in terms of the actual resources that we must forfeit.

Most economic decisions and tradeoffs are not all-or-nothing. Instead, they involve marginal analysis, which means they are about decisions on the margin, involving a little more or a little less. The law of diminishing marginal utility points out that as a person receives more of something—whether it is a specific good or another resource—the

additional marginal gains tend to become smaller. Because sunk costs occurred in the past and cannot be recovered, they should be disregarded in making current decisions.

2.2 The Production Possibilities Frontier and Social Choices

A production possibilities frontier defines the set of choices society faces for the combinations of goods and services it can produce given the resources available. The shape of the PPF is typically curved outward, rather than straight. Choices outside the PPF are unattainable and choices inside the PPF are wasteful. Over time, a growing economy will tend to shift the PPF outwards.

The law of diminishing returns holds that as increments of additional resources are devoted to producing something, the marginal increase in output will become increasingly smaller. All choices along a production possibilities frontier display productive efficiency; that is, it is impossible to use society's resources to produce more of one good without decreasing production of the other good. The specific choice along a production possibilities frontier that reflects the mix of goods society prefers is the choice with allocative efficiency. The curvature of the PPF is likely to differ by country, which results in different countries having comparative advantage in different goods. Total production can increase if countries specialize in the goods in which they have comparative advantage and trade some of their production for the remaining goods.

2.3 Confronting Objections to the Economic Approach

The economic way of thinking provides a useful approach to understanding human behavior. Economists make the careful distinction between positive statements, which describe the world as it is, and normative statements, which describe how the world should be. Even when economics analyzes the gains and losses from various events or policies, and thus draws normative conclusions about how the world should be, the analysis of economics is rooted in a positive analysis of how people, firms, and governments actually behave, not how they should behave.

SELF-CHECK QUESTIONS

1. Suppose Alphonso's town raised the price of bus tickets to $1 per trip (while the price of burgers stayed at $2 and his budget remained $10 per week.) Draw Alphonso's new budget constraint. What happens to the opportunity cost of bus tickets?

2. Return to the example in **Figure 2.4**. Suppose there is an improvement in medical technology that enables more healthcare with the same amount of resources. How would this affect the production possibilities curve and, in particular, how would it affect the opportunity cost of education?

3. Could a nation be producing in a way that is allocatively efficient, but productively inefficient?

4. What are the similarities between a consumer's budget constraint and society's production possibilities frontier, not just graphically but analytically?

5. Individuals may not act in the rational, calculating way described by the economic model of decision making, measuring utility and costs at the margin, but can you make a case that they behave approximately that way?

6. Would an op-ed piece in a newspaper urging the adoption of a particular economic policy be a positive or normative statement?

7. Would a research study on the effects of soft drink consumption on children's cognitive development be a positive or normative statement?

REVIEW QUESTIONS

8. Explain why scarcity leads to tradeoffs.

9. Explain why individuals make choices that are directly on the budget constraint, rather than inside the budget constraint or outside it.

10. What is comparative advantage?

11. What does a production possibilities frontier illustrate?

12. Why is a production possibilities frontier typically drawn as a curve, rather than a straight line?

13. Explain why societies cannot make a choice above their production possibilities frontier and should not make a choice below it.

14. What are diminishing marginal returns?

15. What is productive efficiency? Allocative efficiency?

16. What is the difference between a positive and a normative statement?

17. Is the economic model of decision-making intended as a literal description of how individuals, firms, and the governments actually make decisions?

18. What are four responses to the claim that people should not behave in the way described in this chapter?

CRITICAL THINKING QUESTIONS

19. Suppose Alphonso's town raises the price of bus tickets from $0.50 to $1 and the price of burgers rises from $2 to $4. Why is the opportunity cost of bus tickets unchanged? Suppose Alphonso's weekly spending money increases from $10 to $20. How is his budget constraint affected from all three changes? Explain.

20. During the Second World War, Germany's factories were decimated. It also suffered many human casualties, both soldiers and civilians. How did the war affect Germany's production possibilities curve?

21. It is clear that productive inefficiency is a waste since resources are used in a way that produces less goods and services than a nation is capable of. Why is allocative inefficiency also wasteful?

22. What assumptions about the economy must be true for the invisible hand to work? To what extent are those assumptions valid in the real world?

23. Do economists have any particular expertise at making normative arguments? In other words, they have expertise at making positive statements (i.e., what *will* happen) about some economic policy, for example, but do they have special expertise to judge whether or not the policy *should* be undertaken?

PROBLEMS

Use this information to answer the following 4 questions: Marie has a weekly budget of $24, which she likes to spend on magazines and pies.

24. If the price of a magazine is $4 each, what is the maximum number of magazines she could buy in a week?

25. If the price of a pie is $12, what is the maximum number of pies she could buy in a week?

26. Draw Marie's budget constraint with pies on the horizontal axis and magazines on the vertical axis. What is the slope of the budget constraint?

27. What is Marie's opportunity cost of purchasing a pie?

3 | Demand and Supply

Figure 3.1 Farmer's Market Organic vegetables and fruits that are grown and sold within a specific geographical region should, in theory, cost less than conventional produce because the transportation costs are less. That is not, however, usually the case. (Credit: Modification of work by Natalie Maynor/Flickr Creative Commons)

Bring it Home

Why Can We Not Get Enough of Organic?

Organic food is increasingly popular, not just in the United States, but worldwide. At one time, consumers had to go to specialty stores or farmers' markets to find organic produce. Now it is available in most grocery stores. In short, organic is part of the mainstream.

Ever wonder why organic food costs more than conventional food? Why, say, does an organic Fuji apple cost $1.99 a pound, while its conventional counterpart costs $1.49 a pound? The same price relationship is true for just about every organic product on the market. If many organic foods are locally grown, would they not take less time to get to market and therefore be cheaper? What are the forces that keep those prices from coming down? Turns out those forces have quite a bit to do with this chapter's topic: demand and supply.

Introduction to Demand and Supply

In this chapter, you will learn about:

- Demand, Supply, and Equilibrium in Markets for Goods and Services
- Shifts in Demand and Supply for Goods and Services
- Changes in Equilibrium Price and Quantity: The Four-Step Process

- Price Ceilings and Price Floors

An auction bidder pays thousands of dollars for a dress Whitney Houston wore. A collector spends a small fortune for a few drawings by John Lennon. People usually react to purchases like these in two ways: their jaw drops because they think these are high prices to pay for such goods or they think these are rare, desirable items and the amount paid seems right.

Link It Up

Visit this website (http://openstaxcollege.org/l/celebauction) to read a list of bizarre items that have been purchased for their ties to celebrities. These examples represent an interesting facet of demand and supply.

When economists talk about prices, they are less interested in making judgments than in gaining a practical understanding of what determines prices and why prices change. Consider a price most of us contend with weekly: that of a gallon of gas. Why was the average price of gasoline in the United States $3.71 per gallon in June 2014? Why did the price for gasoline fall sharply to $1.96 per gallon by January 2016? To explain these price movements, economists focus on the determinants of what gasoline buyers are willing to pay and what gasoline sellers are willing to accept.

As it turns out, the price of gasoline in June of any given year is nearly always higher than the price in January of that same year. Over recent decades, gasoline prices in midsummer have averaged about 10 cents per gallon more than their midwinter low. The likely reason is that people drive more in the summer, and are also willing to pay more for gas, but that does not explain how steeply gas prices fell. Other factors were at work during those 18 months, such as increases in supply and decreases in the demand for crude oil.

This chapter introduces the economic model of demand and supply—one of the most powerful models in all of economics. The discussion here begins by examining how demand and supply determine the price and the quantity sold in markets for goods and services, and how changes in demand and supply lead to changes in prices and quantities.

3.1 | Demand, Supply, and Equilibrium in Markets for Goods and Services

By the end of this section, you will be able to:

- Explain demand, quantity demanded, and the law of demand
- Identify a demand curve and a supply curve
- Explain supply, quantity supplied, and the law of supply
- Explain equilibrium, equilibrium price, and equilibrium quantity

First let’s first focus on what economists mean by demand, what they mean by supply, and then how demand and supply interact in a market.

Demand for Goods and Services

Economists use the term **demand** to refer to the amount of some good or service consumers are willing and able to purchase at each price. Demand is fundamentally based on needs and wants—if you have no need or want for something, you won't buy it. While a consumer may be able to differentiate between a need and a want, but from an economist's perspective they are the same thing. Demand is also based on ability to pay. If you cannot pay for it, you have no effective demand. By this definition, a homeless person probably has no effective demand for shelter.

What a buyer pays for a unit of the specific good or service is called **price**. The total number of units that consumers would purchase at that price is called the **quantity demanded**. A rise in price of a good or service almost always decreases the quantity demanded of that good or service. Conversely, a fall in price will increase the quantity demanded. When the price of a gallon of gasoline increases, for example, people look for ways to reduce their consumption by combining several errands, commuting by carpool or mass transit, or taking weekend or vacation trips closer to home. Economists call this inverse relationship between price and quantity demanded the **law of demand**. The law of demand assumes that all other variables that affect demand (which we explain in the next module) are held constant.

We can show an example from the market for gasoline in a table or a graph. Economist call a table that shows the quantity demanded at each price, such as **Table 3.1**, a **demand schedule**. In this case we measure price in dollars per gallon of gasoline. We measure the quantity demanded in millions of gallons over some time period (for example, per day or per year) and over some geographic area (like a state or a country). A **demand curve** shows the relationship between price and quantity demanded on a graph like **Figure 3.2**, with quantity on the horizontal axis and the price per gallon on the vertical axis. (Note that this is an exception to the normal rule in mathematics that the independent variable (x) goes on the horizontal axis and the dependent variable (y) goes on the vertical. Economics is not math.)

Table 3.1 shows the demand schedule and the graph in **Figure 3.2** shows the demand curve. These are two ways to describe the same relationship between price and quantity demanded.

Price (per gallon)	Quantity Demanded (millions of gallons)
$1.00	800
$1.20	700
$1.40	600
$1.60	550
$1.80	500
$2.00	460
$2.20	420

Table 3.1 Price and Quantity Demanded of Gasoline

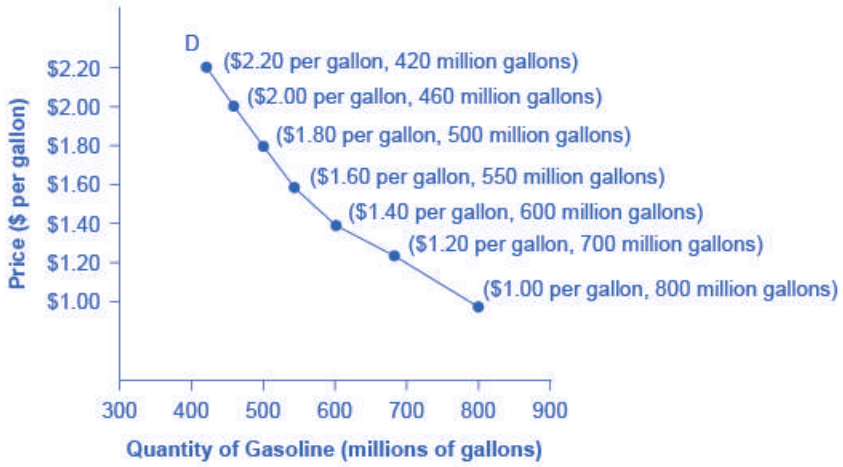


Figure 3.2 A Demand Curve for Gasoline The demand schedule shows that as price rises, quantity demanded decreases, and vice versa. We graph these points, and the line connecting them is the demand curve (D). The downward slope of the demand curve again illustrates the law of demand—the inverse relationship between prices and quantity demanded.

Demand curves will appear somewhat different for each product. They may appear relatively steep or flat, or they may be straight or curved. Nearly all demand curves share the fundamental similarity that they slope down from left to right. Demand curves embody the law of demand: As the price increases, the quantity demanded decreases, and conversely, as the price decreases, the quantity demanded increases.

Confused about these different types of demand? Read the next Clear It Up feature.

Clear It Up

Is demand the same as quantity demanded?

In economic terminology, demand is not the same as quantity demanded. When economists talk about demand, they mean the relationship between a range of prices and the quantities demanded at those prices, as illustrated by a demand curve or a demand schedule. When economists talk about quantity demanded, they mean only a certain point on the demand curve, or one quantity on the demand schedule. In short, demand refers to the curve and quantity demanded refers to the (specific) point on the curve.

Supply of Goods and Services

When economists talk about **supply**, they mean the amount of some good or service a producer is willing to supply at each price. Price is what the producer receives for selling one unit of a good or service. A rise in price almost always leads to an increase in the **quantity supplied** of that good or service, while a fall in price will decrease the quantity supplied. When the price of gasoline rises, for example, it encourages profit-seeking firms to take several actions: expand exploration for oil reserves; drill for more oil; invest in more pipelines and oil tankers to bring the oil to plants for refining into gasoline; build new oil refineries; purchase additional pipelines and trucks to ship the gasoline to gas stations; and open more gas stations or keep existing gas stations open longer hours. Economists call this positive relationship between price and quantity supplied—that a higher price leads to a higher quantity supplied and a lower price leads to a lower quantity supplied—the **law of supply**. The law of supply assumes that all other variables that affect supply (to be explained in the next module) are held constant.

Still unsure about the different types of supply? See the following Clear It Up feature.

Clear It Up

Is supply the same as quantity supplied?

In economic terminology, supply is not the same as quantity supplied. When economists refer to supply, they mean the relationship between a range of prices and the quantities supplied at those prices, a relationship that we can illustrate with a supply curve or a supply schedule. When economists refer to quantity supplied, they mean only a certain point on the supply curve, or one quantity on the supply schedule. In short, supply refers to the curve and quantity supplied refers to the (specific) point on the curve.

Figure 3.3 illustrates the law of supply, again using the market for gasoline as an example. Like demand, we can illustrate supply using a table or a graph. A **supply schedule** is a table, like **Table 3.2**, that shows the quantity supplied at a range of different prices. Again, we measure price in dollars per gallon of gasoline and we measure quantity supplied in millions of gallons. A **supply curve** is a graphic illustration of the relationship between price, shown on the vertical axis, and quantity, shown on the horizontal axis. The supply schedule and the supply curve are just two different ways of showing the same information. Notice that the horizontal and vertical axes on the graph for the supply curve are the same as for the demand curve.

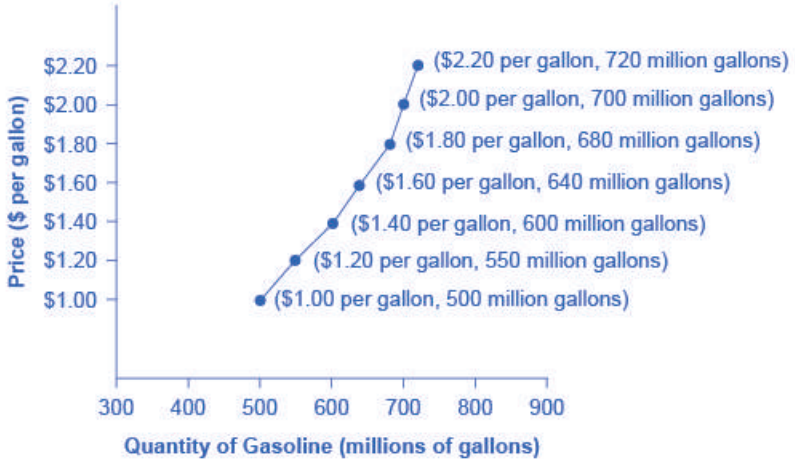


Figure 3.3 A Supply Curve for Gasoline The supply schedule is the table that shows quantity supplied of gasoline at each price. As price rises, quantity supplied also increases, and vice versa. The supply curve (S) is created by graphing the points from the supply schedule and then connecting them. The upward slope of the supply curve illustrates the law of supply—that a higher price leads to a higher quantity supplied, and vice versa.

Price (per gallon)	Quantity Supplied (millions of gallons)
$1.00	500
$1.20	550
$1.40	600
$1.60	640
$1.80	680
$2.00	700

Table 3.2 Price and Supply of Gasoline

Price (per gallon)	Quantity Supplied (millions of gallons)
$2.20	720

Table 3.2 Price and Supply of Gasoline

The shape of supply curves will vary somewhat according to the product: steeper, flatter, straighter, or curved. Nearly all supply curves, however, share a basic similarity: they slope up from left to right and illustrate the law of supply: as the price rises, say, from $1.00 per gallon to $2.20 per gallon, the quantity supplied increases from 500 gallons to 720 gallons. Conversely, as the price falls, the quantity supplied decreases.

Equilibrium—Where Demand and Supply Intersect

Because the graphs for demand and supply curves both have price on the vertical axis and quantity on the horizontal axis, the demand curve and supply curve for a particular good or service can appear on the same graph. Together, demand and supply determine the price and the quantity that will be bought and sold in a market.

Figure 3.4 illustrates the interaction of demand and supply in the market for gasoline. The demand curve (D) is identical to **Figure 3.2**. The supply curve (S) is identical to **Figure 3.3**. **Table 3.3** contains the same information in tabular form.

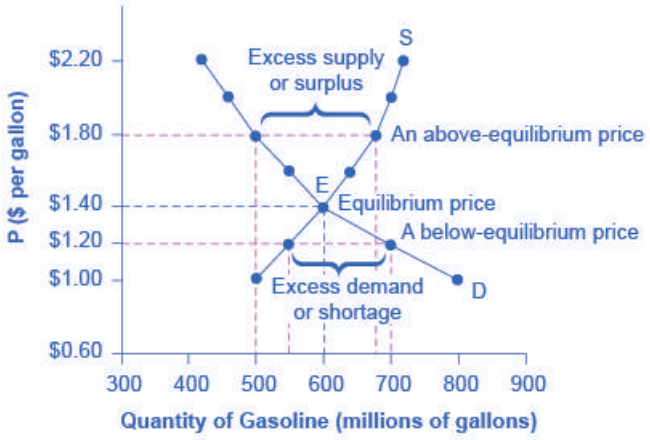


Figure 3.4 Demand and Supply for Gasoline The demand curve (D) and the supply curve (S) intersect at the equilibrium point E, with a price of $1.40 and a quantity of 600. The equilibrium is the only price where quantity demanded is equal to quantity supplied. At a price above equilibrium like $1.80, quantity supplied exceeds the quantity demanded, so there is excess supply. At a price below equilibrium such as $1.20, quantity demanded exceeds quantity supplied, so there is excess demand.

Price (per gallon)	Quantity demanded (millions of gallons)	Quantity supplied (millions of gallons)
$1.00	800	500
$1.20	700	550
$1.40	**600**	**600**
$1.60	550	640
$1.80	500	680

Table 3.3 Price, Quantity Demanded, and Quantity Supplied

Price (per gallon)	Quantity demanded (millions of gallons)	Quantity supplied (millions of gallons)
$2.00	460	700
$2.20	420	720

Table 3.3 Price, Quantity Demanded, and Quantity Supplied

Remember this: When two lines on a diagram cross, this intersection usually means something. The point where the supply curve (S) and the demand curve (D) cross, designated by point E in **Figure 3.4**, is called the **equilibrium**. The **equilibrium price** is the only price where the plans of consumers and the plans of producers agree—that is, where the amount of the product consumers want to buy (quantity demanded) is equal to the amount producers want to sell (quantity supplied). Economists call this common quantity the **equilibrium quantity**. At any other price, the quantity demanded does not equal the quantity supplied, so the market is not in equilibrium at that price.

In **Figure 3.4**, the equilibrium price is $1.40 per gallon of gasoline and the equilibrium quantity is 600 million gallons. If you had only the demand and supply schedules, and not the graph, you could find the equilibrium by looking for the price level on the tables where the quantity demanded and the quantity supplied are equal.

The word “equilibrium” means “balance.” If a market is at its equilibrium price and quantity, then it has no reason to move away from that point. However, if a market is not at equilibrium, then economic pressures arise to move the market toward the equilibrium price and the equilibrium quantity.

Imagine, for example, that the price of a gallon of gasoline was above the equilibrium price—that is, instead of $1.40 per gallon, the price is $1.80 per gallon. The dashed horizontal line at the price of $1.80 in **Figure 3.4** illustrates this above equilibrium price. At this higher price, the quantity demanded drops from 600 to 500. This decline in quantity reflects how consumers react to the higher price by finding ways to use less gasoline.

Moreover, at this higher price of $1.80, the quantity of gasoline supplied rises from the 600 to 680, as the higher price makes it more profitable for gasoline producers to expand their output. Now, consider how quantity demanded and quantity supplied are related at this above-equilibrium price. Quantity demanded has fallen to 500 gallons, while quantity supplied has risen to 680 gallons. In fact, at any above-equilibrium price, the quantity supplied exceeds the quantity demanded. We call this an **excess supply** or a **surplus**.

With a surplus, gasoline accumulates at gas stations, in tanker trucks, in pipelines, and at oil refineries. This accumulation puts pressure on gasoline sellers. If a surplus remains unsold, those firms involved in making and selling gasoline are not receiving enough cash to pay their workers and to cover their expenses. In this situation, some producers and sellers will want to cut prices, because it is better to sell at a lower price than not to sell at all. Once some sellers start cutting prices, others will follow to avoid losing sales. These price reductions in turn will stimulate a higher quantity demanded. Therefore, if the price is above the equilibrium level, incentives built into the structure of demand and supply will create pressures for the price to fall toward the equilibrium.

Now suppose that the price is below its equilibrium level at $1.20 per gallon, as the dashed horizontal line at this price in **Figure 3.4** shows. At this lower price, the quantity demanded increases from 600 to 700 as drivers take longer trips, spend more minutes warming up the car in the driveway in wintertime, stop sharing rides to work, and buy larger cars that get fewer miles to the gallon. However, the below-equilibrium price reduces gasoline producers’ incentives to produce and sell gasoline, and the quantity supplied falls from 600 to 550.

When the price is below equilibrium, there is **excess demand**, or a **shortage**—that is, at the given price the quantity demanded, which has been stimulated by the lower price, now exceeds the quantity supplied, which had been depressed by the lower price. In this situation, eager gasoline buyers mob the gas stations, only to find many stations running short of fuel. Oil companies and gas stations recognize that they have an opportunity to make higher profits by selling what gasoline they have at a higher price. As a result, the price rises toward the equilibrium level. Read **Demand, Supply, and Efficiency** for more discussion on the importance of the demand and supply model.

3.2 | Shifts in Demand and Supply for Goods and

Services

By the end of this section, you will be able to:

- Identify factors that affect demand
- Graph demand curves and demand shifts
- Identify factors that affect supply
- Graph supply curves and supply shifts

The previous module explored how price affects the quantity demanded and the quantity supplied. The result was the demand curve and the supply curve. Price, however, is not the only factor that influences demand, nor is it the only thing that influences supply. For example, how is demand for vegetarian food affected if, say, health concerns cause more consumers to avoid eating meat? How is the supply of diamonds affected if diamond producers discover several new diamond mines? What are the major factors, in addition to the price, that influence demand or supply?

Link It Up

Visit this website (http://openstaxcollege.org/l/toothfish) to read a brief note on how marketing strategies can influence supply and demand of products.

What Factors Affect Demand?

We defined demand as the amount of some product a consumer is willing and able to purchase at each price. That suggests at least two factors in addition to price that affect demand. Willingness to purchase suggests a desire, based on what economists call tastes and preferences. If you neither need nor want something, you will not buy it. Ability to purchase suggests that income is important. Professors are usually able to afford better housing and transportation than students, because they have more income. Prices of related goods can affect demand also. If you need a new car, the price of a Honda may affect your demand for a Ford. Finally, the size or composition of the population can affect demand. The more children a family has, the greater their demand for clothing. The more driving-age children a family has, the greater their demand for car insurance, and the less for diapers and baby formula.

These factors matter for both individual and market demand as a whole. Exactly how do these various factors affect demand, and how do we show the effects graphically? To answer those questions, we need the *ceteris paribus* assumption.

The *Ceteris Paribus* Assumption

A demand curve or a supply curve is a relationship between two, and only two, variables: quantity on the horizontal axis and price on the vertical axis. The assumption behind a demand curve or a supply curve is that no relevant economic factors, other than the product's price, are changing. Economists call this assumption **ceteris paribus**, a Latin phrase meaning "other things being equal." Any given demand or supply curve is based on the *ceteris paribus* assumption that all else is held equal. A demand curve or a supply curve is a relationship between two, and only two, variables when all other variables are kept constant. If all else is not held equal, then the laws of supply and demand will not necessarily hold, as the following Clear It Up feature shows.

Clear It Up

When does *ceteris paribus* apply?

We typically apply *ceteris paribus* when we observe how changes in price affect demand or supply, but we can apply *ceteris paribus* more generally. In the real world, demand and supply depend on more factors than just price. For example, a consumer's demand depends on income and a producer's supply depends on the cost of producing the product. How can we analyze the effect on demand or supply if multiple factors are changing at the same time—say price rises and income falls? The answer is that we examine the changes one at a time, assuming the other factors are held constant.

For example, we can say that an increase in the price reduces the amount consumers will buy (assuming income, and anything else that affects demand, is unchanged). Additionally, a decrease in income reduces the amount consumers can afford to buy (assuming price, and anything else that affects demand, is unchanged). This is what the *ceteris paribus* assumption really means. In this particular case, after we analyze each factor separately, we can combine the results. The amount consumers buy falls for two reasons: first because of the higher price and second because of the lower income.

How Does Income Affect Demand?

Let's use income as an example of how factors other than price affect demand. **Figure 3.5** shows the initial demand for automobiles as D_0. At point Q, for example, if the price is $20,000 per car, the quantity of cars demanded is 18 million. D_0 also shows how the quantity of cars demanded would change as a result of a higher or lower price. For example, if the price of a car rose to $22,000, the quantity demanded would decrease to 17 million, at point R.

The original demand curve D_0, like every demand curve, is based on the *ceteris paribus* assumption that no other economically relevant factors change. Now imagine that the economy expands in a way that raises the incomes of many people, making cars more affordable. How will this affect demand? How can we show this graphically?

Return to **Figure 3.5**. The price of cars is still $20,000, but with higher incomes, the quantity demanded has now increased to 20 million cars, shown at point S. As a result of the higher income levels, the demand curve shifts to the right to the new demand curve D_1, indicating an increase in demand. **Table 3.4** shows clearly that this increased demand would occur at every price, not just the original one.

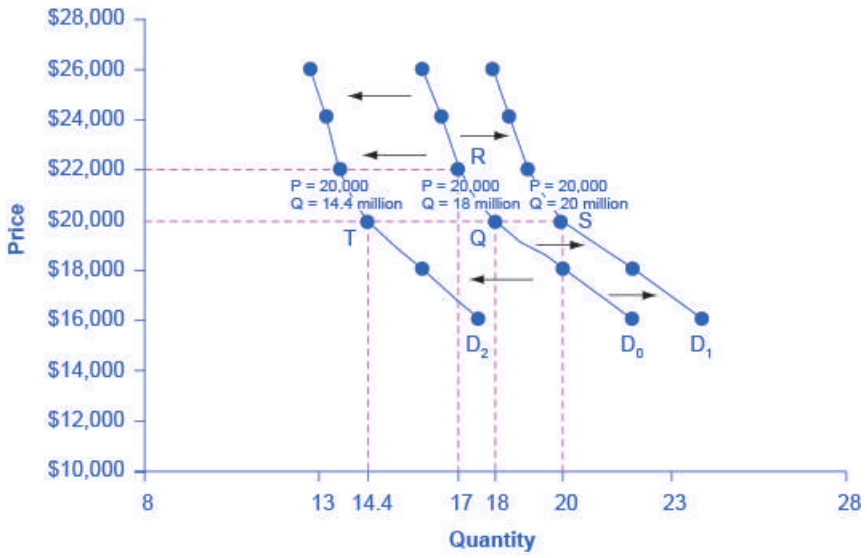


Figure 3.5 Shifts in Demand: A Car Example Increased demand means that at every given price, the quantity demanded is higher, so that the demand curve shifts to the right from D_0 to D_1. Decreased demand means that at every given price, the quantity demanded is lower, so that the demand curve shifts to the left from D_0 to D_2.

Price	Decrease to D_2	Original Quantity Demanded D_0	Increase to D_1
$16,000	17.6 million	22.0 million	24.0 million
$18,000	16.0 million	20.0 million	22.0 million
$20,000	14.4 million	18.0 million	20.0 million
$22,000	13.6 million	17.0 million	19.0 million
$24,000	13.2 million	16.5 million	18.5 million
$26,000	12.8 million	16.0 million	18.0 million

Table 3.4 Price and Demand Shifts: A Car Example

Now, imagine that the economy slows down so that many people lose their jobs or work fewer hours, reducing their incomes. In this case, the decrease in income would lead to a lower quantity of cars demanded at every given price, and the original demand curve D_0 would shift left to D_2. The shift from D_0 to D_2 represents such a decrease in demand: At any given price level, the quantity demanded is now lower. In this example, a price of $20,000 means 18 million cars sold along the original demand curve, but only 14.4 million sold after demand fell.

When a demand curve shifts, it does not mean that the quantity demanded by every individual buyer changes by the same amount. In this example, not everyone would have higher or lower income and not everyone would buy or not buy an additional car. Instead, a shift in a demand curve captures a pattern for the market as a whole.

In the previous section, we argued that higher income causes greater demand at every price. This is true for most goods and services. For some—luxury cars, vacations in Europe, and fine jewelry—the effect of a rise in income can be especially pronounced. A product whose demand rises when income rises, and vice versa, is called a **normal good**. A few exceptions to this pattern do exist. As incomes rise, many people will buy fewer generic brand groceries and more name brand groceries. They are less likely to buy used cars and more likely to buy new cars. They will be less likely to rent an apartment and more likely to own a home. A product whose demand falls when income rises, and vice versa, is called an **inferior good**. In other words, when income increases, the demand curve shifts to the left.

Other Factors That Shift Demand Curves

Income is not the only factor that causes a shift in demand. Other factors that change demand include tastes and preferences, the composition or size of the population, the prices of related goods, and even expectations. A change in any one of the underlying factors that determine what quantity people are willing to buy at a given price will cause a shift in demand. Graphically, the new demand curve lies either to the right (an increase) or to the left (a decrease) of the original demand curve. Let’s look at these factors.

Changing Tastes or Preferences

From 1980 to 2014, the per-person consumption of chicken by Americans rose from 48 pounds per year to 85 pounds per year, and consumption of beef fell from 77 pounds per year to 54 pounds per year, according to the U.S. Department of Agriculture (USDA). Changes like these are largely due to movements in taste, which change the quantity of a good demanded at every price: that is, they shift the demand curve for that good, rightward for chicken and leftward for beef.

Changes in the Composition of the Population

The proportion of elderly citizens in the United States population is rising. It rose from 9.8% in 1970 to 12.6% in 2000, and will be a projected (by the U.S. Census Bureau) 20% of the population by 2030. A society with relatively more children, like the United States in the 1960s, will have greater demand for goods and services like tricycles and day care facilities. A society with relatively more elderly persons, as the United States is projected to have by 2030, has a higher demand for nursing homes and hearing aids. Similarly, changes in the size of the population can affect the demand for housing and many other goods. Each of these changes in demand will be shown as a shift in the demand curve.

Changes in the prices of related goods such as substitutes or complements also can affect the demand for a product. A

substitute is a good or service that we can use in place of another good or service. As electronic books, like this one, become more available, you would expect to see a decrease in demand for traditional printed books. A lower price for a substitute decreases demand for the other product. For example, in recent years as the price of tablet computers has fallen, the quantity demanded has increased (because of the law of demand). Since people are purchasing tablets, there has been a decrease in demand for laptops, which we can show graphically as a leftward shift in the demand curve for laptops. A higher price for a substitute good has the reverse effect.

Other goods are **complements** for each other, meaning we often use the goods together, because consumption of one good tends to enhance consumption of the other. Examples include breakfast cereal and milk; notebooks and pens or pencils, golf balls and golf clubs; gasoline and sport utility vehicles; and the five-way combination of bacon, lettuce, tomato, mayonnaise, and bread. If the price of golf clubs rises, since the quantity demanded of golf clubs falls (because of the law of demand), demand for a complement good like golf balls decreases, too. Similarly, a higher price for skis would shift the demand curve for a complement good like ski resort trips to the left, while a lower price for a complement has the reverse effect.

Changes in Expectations about Future Prices or Other Factors that Affect Demand

While it is clear that the price of a good affects the quantity demanded, it is also true that expectations about the future price (or expectations about tastes and preferences, income, and so on) can affect demand. For example, if people hear that a hurricane is coming, they may rush to the store to buy flashlight batteries and bottled water. If people learn that the price of a good like coffee is likely to rise in the future, they may head for the store to stock up on coffee now. We show these changes in demand as shifts in the curve. Therefore, a **shift in demand** happens when a change in some economic factor (other than price) causes a different quantity to be demanded at every price. The following Work It Out feature shows how this happens.

Work It Out

Shift in Demand

A shift in demand means that at any price (and at every price), the quantity demanded will be different than it was before. Following is an example of a shift in demand due to an income increase.

Step 1. Draw the graph of a demand curve for a normal good like pizza. Pick a price (like P_0). Identify the corresponding Q_0. See an example in **Figure 3.6**.

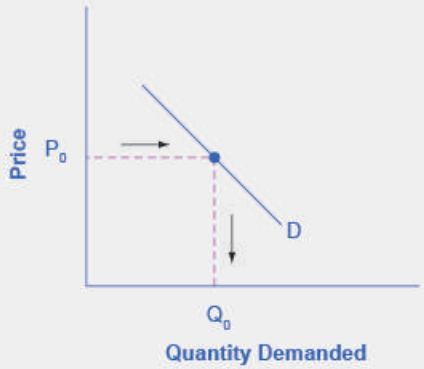


Figure 3.6 Demand Curve We can use the demand curve to identify how much consumers would buy at any given price.

Step 2. Suppose income increases. As a result of the change, are consumers going to buy more or less pizza? The answer is more. Draw a dotted horizontal line from the chosen price, through the original quantity demanded, to the new point with the new Q_1. Draw a dotted vertical line down to the horizontal axis and label the new Q_1. **Figure 3.7** provides an example.

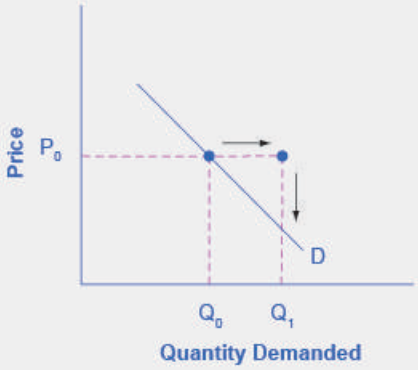


Figure 3.7 Demand Curve with Income Increase With an increase in income, consumers will purchase larger quantities, pushing demand to the right.

Step 3. Now, shift the curve through the new point. You will see that an increase in income causes an upward (or rightward) shift in the demand curve, so that at any price the quantities demanded will be higher, as **Figure 3.8** illustrates.

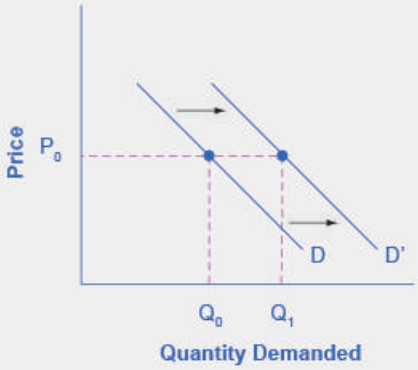


Figure 3.8 Demand Curve Shifted Right With an increase in income, consumers will purchase larger quantities, pushing demand to the right, and causing the demand curve to shift right.

Summing Up Factors That Change Demand

Figure 3.9 summarizes six factors that can shift demand curves. The direction of the arrows indicates whether the demand curve shifts represent an increase in demand or a decrease in demand. Notice that a change in the price of the good or service itself is not listed among the factors that can shift a demand curve. A change in the price of a good or service causes a movement along a specific demand curve, and it typically leads to some change in the quantity demanded, but it does not shift the demand curve.

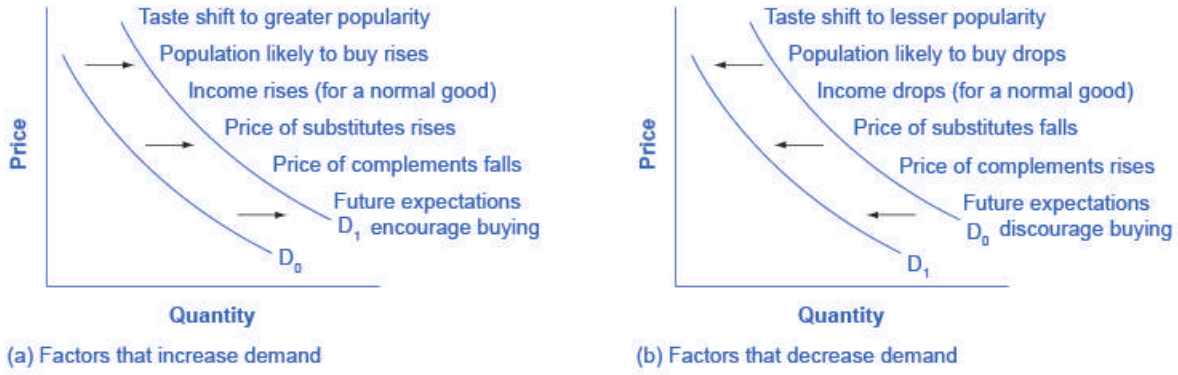


Figure 3.9 Factors That Shift Demand Curves (a) A list of factors that can cause an increase in demand from D_0 to D_1. (b) The same factors, if their direction is reversed, can cause a decrease in demand from D_0 to D_1.

When a demand curve shifts, it will then intersect with a given supply curve at a different equilibrium price and quantity. We are, however, getting ahead of our story. Before discussing how changes in demand can affect equilibrium price and quantity, we first need to discuss shifts in supply curves.

How Production Costs Affect Supply

A supply curve shows how quantity supplied will change as the price rises and falls, assuming *ceteris paribus* so that no other economically relevant factors are changing. If other factors relevant to supply do change, then the entire supply curve will shift. Just as we described a shift in demand as a change in the quantity demanded at every price, a **shift in supply** means a change in the quantity supplied at every price.

In thinking about the factors that affect supply, remember what motivates firms: profits, which are the difference between revenues and costs. A firm produces goods and services using combinations of labor, materials, and machinery, or what we call **inputs** or **factors of production**. If a firm faces lower costs of production, while the prices for the good or service the firm produces remain unchanged, a firm's profits go up. When a firm's profits increase, it is more motivated to produce output, since the more it produces the more profit it will earn. When costs of production fall, a firm will tend to supply a larger quantity at any given price for its output. We can show this by the supply curve shifting to the right.

Take, for example, a messenger company that delivers packages around a city. The company may find that buying gasoline is one of its main costs. If the price of gasoline falls, then the company will find it can deliver messages more cheaply than before. Since lower costs correspond to higher profits, the messenger company may now supply more of its services at any given price. For example, given the lower gasoline prices, the company can now serve a greater area, and increase its supply.

Conversely, if a firm faces higher costs of production, then it will earn lower profits at any given selling price for its products. As a result, a higher cost of production typically causes a firm to supply a smaller quantity at any given price. In this case, the supply curve shifts to the left.

Consider the supply for cars, shown by curve S_0 in **Figure 3.10**. Point J indicates that if the price is $20,000, the quantity supplied will be 18 million cars. If the price rises to $22,000 per car, *ceteris paribus*, the quantity supplied will rise to 20 million cars, as point K on the S_0 curve shows. We can show the same information in table form, as in **Table 3.5**.

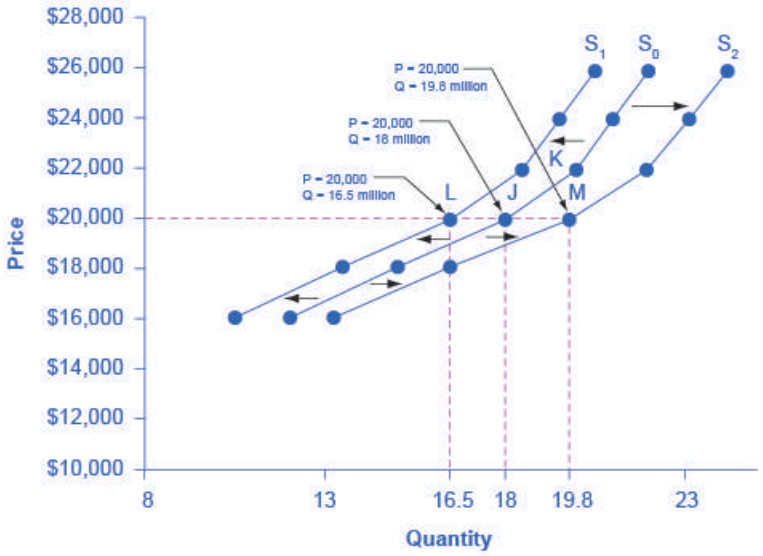


Figure 3.10 Shifts in Supply: A Car Example Decreased supply means that at every given price, the quantity supplied is lower, so that the supply curve shifts to the left, from S_0 to S_1. Increased supply means that at every given price, the quantity supplied is higher, so that the supply curve shifts to the right, from S_0 to S_2.

Price	Decrease to S_1	Original Quantity Supplied S_0	Increase to S_2
$16,000	10.5 million	12.0 million	13.2 million
$18,000	13.5 million	15.0 million	16.5 million
$20,000	16.5 million	18.0 million	19.8 million
$22,000	18.5 million	20.0 million	22.0 million
$24,000	19.5 million	21.0 million	23.1 million
$26,000	20.5 million	22.0 million	24.2 million

Table 3.5 Price and Shifts in Supply: A Car Example

Now, imagine that the price of steel, an important ingredient in manufacturing cars, rises, so that producing a car has become more expensive. At any given price for selling cars, car manufacturers will react by supplying a lower quantity. We can show this graphically as a leftward shift of supply, from S_0 to S_1, which indicates that at any given price, the quantity supplied decreases. In this example, at a price of $20,000, the quantity supplied decreases from 18 million on the original supply curve (S_0) to 16.5 million on the supply curve S_1, which is labeled as point L.

Conversely, if the price of steel decreases, producing a car becomes less expensive. At any given price for selling cars, car manufacturers can now expect to earn higher profits, so they will supply a higher quantity. The shift of supply to the right, from S_0 to S_2, means that at all prices, the quantity supplied has increased. In this example, at a price of $20,000, the quantity supplied increases from 18 million on the original supply curve (S_0) to 19.8 million on the supply curve S_2, which is labeled M.

Other Factors That Affect Supply

In the example above, we saw that changes in the prices of inputs in the production process will affect the cost of production and thus the supply. Several other things affect the cost of production, too, such as changes in weather or other natural conditions, new technologies for production, and some government policies.

Changes in weather and climate will affect the cost of production for many agricultural products. For example, in

2014 the Manchurian Plain in Northeastern China, which produces most of the country's wheat, corn, and soybeans, experienced its most severe drought in 50 years. A drought decreases the supply of agricultural products, which means that at any given price, a lower quantity will be supplied. Conversely, especially good weather would shift the supply curve to the right.

When a firm discovers a new technology that allows the firm to produce at a lower cost, the supply curve will shift to the right, as well. For instance, in the 1960s a major scientific effort nicknamed the Green Revolution focused on breeding improved seeds for basic crops like wheat and rice. By the early 1990s, more than two-thirds of the wheat and rice in low-income countries around the world used these Green Revolution seeds—and the harvest was twice as high per acre. A technological improvement that reduces costs of production will shift supply to the right, so that a greater quantity will be produced at any given price.

Government policies can affect the cost of production and the supply curve through taxes, regulations, and subsidies. For example, the U.S. government imposes a tax on alcoholic beverages that collects about $8 billion per year from producers. Businesses treat taxes as costs. Higher costs decrease supply for the reasons we discussed above. Other examples of policy that can affect cost are the wide array of government regulations that require firms to spend money to provide a cleaner environment or a safer workplace. Complying with regulations increases costs.

A government subsidy, on the other hand, is the opposite of a tax. A subsidy occurs when the government pays a firm directly or reduces the firm's taxes if the firm carries out certain actions. From the firm's perspective, taxes or regulations are an additional cost of production that shifts supply to the left, leading the firm to produce a lower quantity at every given price. Government subsidies reduce the cost of production and increase supply at every given price, shifting supply to the right. The following Work It Out feature shows how this shift happens.

Work It Out

Shift in Supply

We know that a supply curve shows the minimum price a firm will accept to produce a given quantity of output. What happens to the supply curve when the cost of production goes up? Following is an example of a shift in supply due to a production cost increase.

Step 1. Draw a graph of a supply curve for pizza. Pick a quantity (like Q_0). If you draw a vertical line up from Q_0 to the supply curve, you will see the price the firm chooses. **Figure 3.11** provides an example.

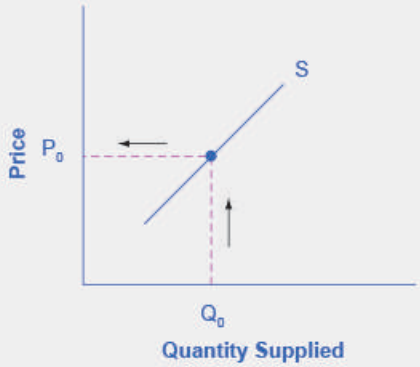


Figure 3.11 Supply Curve You can use a supply curve to show the minimum price a firm will accept to produce a given quantity of output.

Step 2. Why did the firm choose that price and not some other? One way to think about this is that the price is composed of two parts. The first part is the cost of producing pizzas at the margin; in this case, the cost of producing the pizza, including cost of ingredients (e.g., dough, sauce, cheese, and pepperoni), the cost of the pizza oven, the shop rent, and the workers' wages. The second part is the firm's desired profit, which is determined, among other factors, by the profit margins in that particular business. If you add these two parts together, you get the price the firm wishes to charge. The quantity Q0 and associated price P0 give you one point on the firm's supply curve, as **Figure 3.12** illustrates.

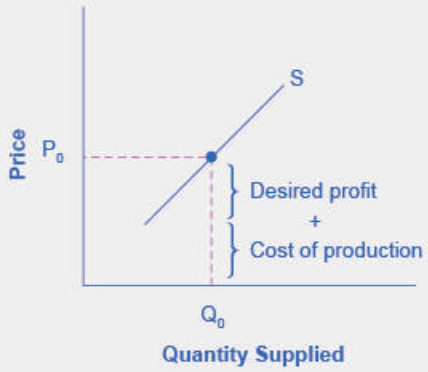


Figure 3.12 Setting Prices The cost of production and the desired profit equal the price a firm will set for a product.

Step 3. Now, suppose that the cost of production increases. Perhaps cheese has become more expensive by $0.75 per pizza. If that is true, the firm will want to raise its price by the amount of the increase in cost ($0.75). Draw this point on the supply curve directly above the initial point on the curve, but $0.75 higher, as **Figure 3.13** shows.

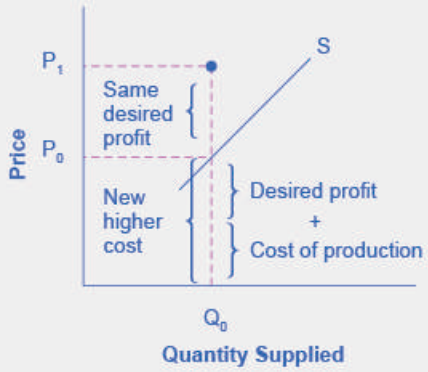


Figure 3.13 Increasing Costs Leads to Increasing Price Because the cost of production and the desired profit equal the price a firm will set for a product, if the cost of production increases, the price for the product will also need to increase.

Step 4. Shift the supply curve through this point. You will see that an increase in cost causes an upward (or a leftward) shift of the supply curve so that at any price, the quantities supplied will be smaller, as **Figure 3.14** illustrates.

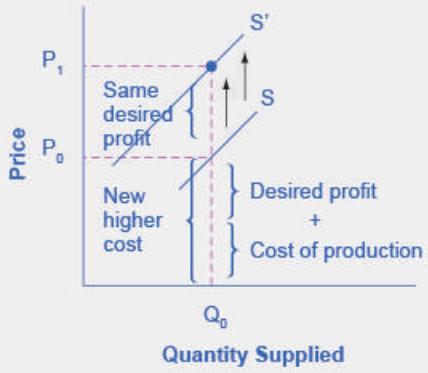


Figure 3.14 Supply Curve Shifts When the cost of production increases, the supply curve shifts upwardly to a new price level.

Summing Up Factors That Change Supply

Changes in the cost of inputs, natural disasters, new technologies, and the impact of government decisions all affect the cost of production. In turn, these factors affect how much firms are willing to supply at any given price.

Figure 3.15 summarizes factors that change the supply of goods and services. Notice that a change in the price of the product itself is not among the factors that shift the supply curve. Although a change in price of a good or service typically causes a change in quantity supplied or a movement along the supply curve for that specific good or service, it does not cause the supply curve itself to shift.

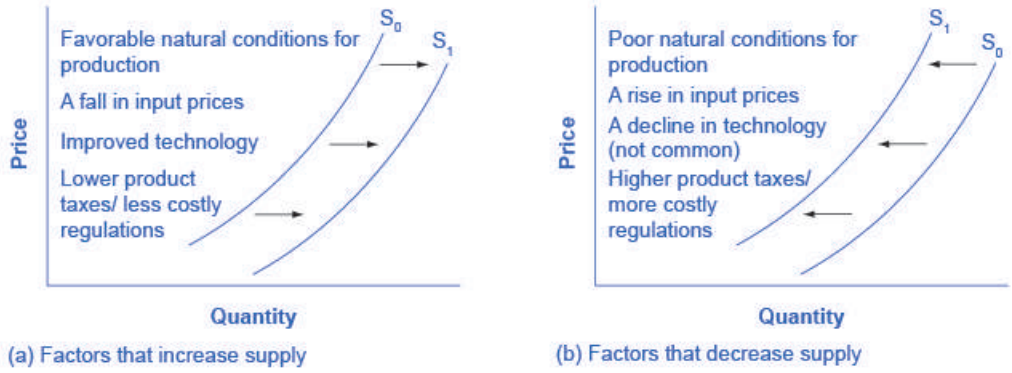


Figure 3.15 Factors That Shift Supply Curves (a) A list of factors that can cause an increase in supply from S_0 to S_1. (b) The same factors, if their direction is reversed, can cause a decrease in supply from S_0 to S_1.

Because demand and supply curves appear on a two-dimensional diagram with only price and quantity on the axes, an unwary visitor to the land of economics might be fooled into believing that economics is about only four topics: demand, supply, price, and quantity. However, demand and supply are really "umbrella" concepts: demand covers all the factors that affect demand, and supply covers all the factors that affect supply. We include factors other than price that affect demand and supply are included by using shifts in the demand or the supply curve. In this way, the two-dimensional demand and supply model becomes a powerful tool for analyzing a wide range of economic circumstances.

3.3 | Changes in Equilibrium Price and Quantity: The Four-Step Process

By the end of this section, you will be able to:

- Identify equilibrium price and quantity through the four-step process
- Graph equilibrium price and quantity
- Contrast shifts of demand or supply and movements along a demand or supply curve
- Graph demand and supply curves, including equilibrium price and quantity, based on real-world examples

Let's begin this discussion with a single economic event. It might be an event that affects demand, like a change in income, population, tastes, prices of substitutes or complements, or expectations about future prices. It might be an event that affects supply, like a change in natural conditions, input prices, or technology, or government policies that affect production. How does this economic event affect equilibrium price and quantity? We will analyze this question using a four-step process.

Step 1. Draw a demand and supply model before the economic change took place. To establish the model requires four standard pieces of information: The law of demand, which tells us the slope of the demand curve; the law of supply, which gives us the slope of the supply curve; the shift variables for demand; and the shift variables for supply. From this model, find the initial equilibrium values for price and quantity.

Step 2. Decide whether the economic change you are analyzing affects demand or supply. In other words, does the

event refer to something in the list of demand factors or supply factors?

Step 3. Decide whether the effect on demand or supply causes the curve to shift to the right or to the left, and sketch the new demand or supply curve on the diagram. In other words, does the event increase or decrease the amount consumers want to buy or producers want to sell?

Step 4. Identify the new equilibrium and then compare the original equilibrium price and quantity to the new equilibrium price and quantity.

Let's consider one example that involves a shift in supply and one that involves a shift in demand. Then we will consider an example where both supply and demand shift.

Good Weather for Salmon Fishing

Supposed that during the summer of 2015, weather conditions were excellent for commercial salmon fishing off the California coast. Heavy rains meant higher than normal levels of water in the rivers, which helps the salmon to breed. Slightly cooler ocean temperatures stimulated the growth of plankton, the microscopic organisms at the bottom of the ocean food chain, providing everything in the ocean with a hearty food supply. The ocean stayed calm during fishing season, so commercial fishing operations did not lose many days to bad weather. How did these climate conditions affect the quantity and price of salmon? **Figure 3.16** illustrates the four-step approach, which we explain below, to work through this problem. **Table 3.6** also provides the information to work the problem.

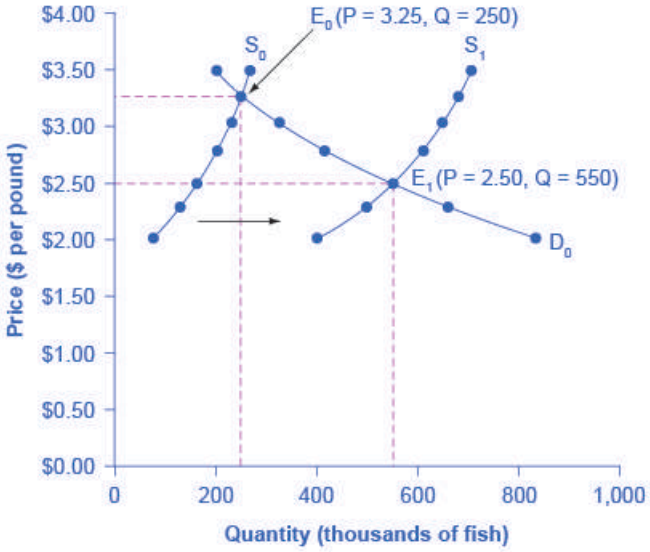


Figure 3.16 Good Weather for Salmon Fishing: The Four-Step Process Unusually good weather leads to changes in the price and quantity of salmon.

Price per Pound	Quantity Supplied in 2014	Quantity Supplied in 2015	Quantity Demanded
$2.00	80	400	840
$2.25	120	480	680
$2.50	160	550	550
$2.75	200	600	450
$3.00	230	640	350
$3.25	250	670	250

Table 3.6 Salmon Fishing

Price per Pound	Quantity Supplied in 2014	Quantity Supplied in 2015	Quantity Demanded
$3.50	270	700	200

Table 3.6 Salmon Fishing

Step 1. Draw a demand and supply model to illustrate the market for salmon in the year before the good weather conditions began. The demand curve D_0 and the supply curve S_0 show that the original equilibrium price is $3.25 per pound and the original equilibrium quantity is 250,000 fish. (This price per pound is what commercial buyers pay at the fishing docks. What consumers pay at the grocery is higher.)

Step 2. Did the economic event affect supply or demand? Good weather is an example of a natural condition that affects supply.

Step 3. Was the effect on supply an increase or a decrease? Good weather is a change in natural conditions that increases the quantity supplied at any given price. The supply curve shifts to the right, moving from the original supply curve S_0 to the new supply curve S_1, which **Figure 3.16** and **Table 3.6** show.

Step 4. Compare the new equilibrium price and quantity to the original equilibrium. At the new equilibrium E_1, the equilibrium price falls from $3.25 to $2.50, but the equilibrium quantity increases from 250,000 to 550,000 salmon. Notice that the equilibrium quantity demanded increased, even though the demand curve did not move.

In short, good weather conditions increased supply of the California commercial salmon. The result was a higher equilibrium quantity of salmon bought and sold in the market at a lower price.

Newspapers and the Internet

According to the Pew Research Center for People and the Press, increasingly more people, especially younger people, are obtaining their news from online and digital sources. The majority of U.S. adults now own smartphones or tablets, and most of those Americans say they use them in part to access the news. From 2004 to 2012, the share of Americans who reported obtaining their news from digital sources increased from 24% to 39%. How has this affected consumption of print news media, and radio and television news? **Figure 3.17** and the text below illustrates using the four-step analysis to answer this question.

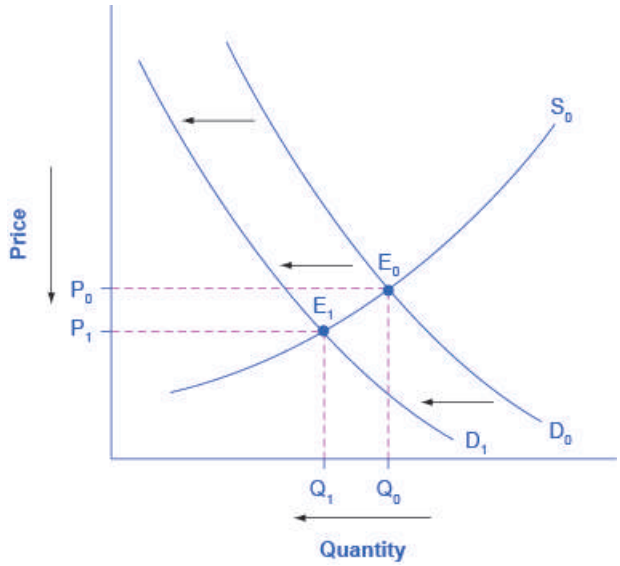


Figure 3.17 The Print News Market: A Four-Step Analysis A change in tastes from print news sources to digital sources results in a leftward shift in demand for the former. The result is a decrease in both equilibrium price and quantity.

Step 1. Develop a demand and supply model to think about what the market looked like before the event. The demand curve D_0 and the supply curve S_0 show the original relationships. In this case, we perform the analysis without

specific numbers on the price and quantity axis.

Step 2. Did the described change affect supply or demand? A change in tastes, from traditional news sources (print, radio, and television) to digital sources, caused a change in demand for the former.

Step 3. Was the effect on demand positive or negative? A shift to digital news sources will tend to mean a lower quantity demanded of traditional news sources at every given price, causing the demand curve for print and other traditional news sources to shift to the left, from D_0 to D_1.

Step 4. Compare the new equilibrium price and quantity to the original equilibrium price. The new equilibrium (E_1) occurs at a lower quantity and a lower price than the original equilibrium (E_0).

The decline in print news reading predates 2004. Print newspaper circulation peaked in 1973 and has declined since then due to competition from television and radio news. In 1991, 55% of Americans indicated they received their news from print sources, while only 29% did so in 2012. Radio news has followed a similar path in recent decades, with the share of Americans obtaining their news from radio declining from 54% in 1991 to 33% in 2012. Television news has held its own over the last 15 years, with a market share staying in the mid to upper fifties. What does this suggest for the future, given that two-thirds of Americans under 30 years old say they do not obtain their news from television at all?

The Interconnections and Speed of Adjustment in Real Markets

In the real world, many factors that affect demand and supply can change all at once. For example, the demand for cars might increase because of rising incomes and population, and it might decrease because of rising gasoline prices (a complementary good). Likewise, the supply of cars might increase because of innovative new technologies that reduce the cost of car production, and it might decrease as a result of new government regulations requiring the installation of costly pollution-control technology.

Moreover, rising incomes and population or changes in gasoline prices will affect many markets, not just cars. How can an economist sort out all these interconnected events? The answer lies in the *ceteris paribus* assumption. Look at how each economic event affects each market, one event at a time, holding all else constant. Then combine the analyses to see the net effect.

A Combined Example

The U.S. Postal Service is facing difficult challenges. Compensation for postal workers tends to increase most years due to cost-of-living increases. At the same time, increasingly more people are using email, text, and other digital message forms such as Facebook and Twitter to communicate with friends and others. What does this suggest about the continued viability of the Postal Service? **Figure 3.18** and the text below illustrate this using the four-step analysis to answer this question.

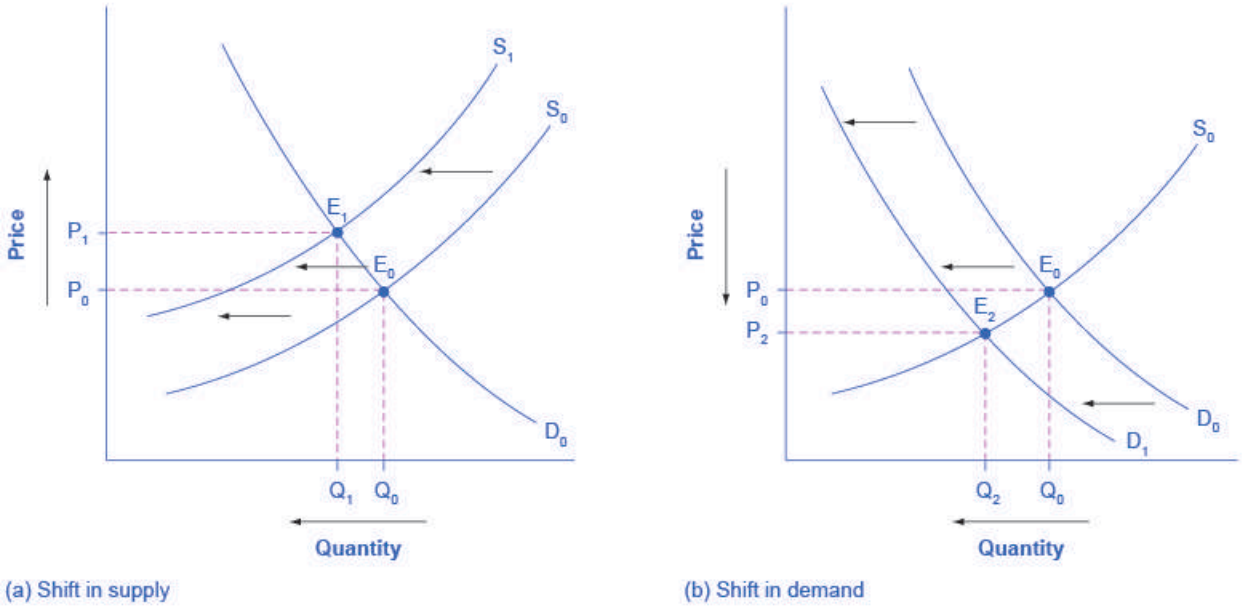


Figure 3.18 Higher Compensation for Postal Workers: A Four-Step Analysis (a) Higher labor compensation causes a leftward shift in the supply curve, a decrease in the equilibrium quantity, and an increase in the equilibrium price. (b) A change in tastes away from Postal Services causes a leftward shift in the demand curve, a decrease in the equilibrium quantity, and a decrease in the equilibrium price.

Since this problem involves two disturbances, we need two four-step analyses, the first to analyze the effects of higher compensation for postal workers, the second to analyze the effects of many people switching from "snail mail" to email and other digital messages.

Figure 3.18 (a) shows the shift in supply discussed in the following steps.

Step 1. Draw a demand and supply model to illustrate what the market for the U.S. Postal Service looked like before this scenario starts. The demand curve D_0 and the supply curve S_0 show the original relationships.

Step 2. Did the described change affect supply or demand? Labor compensation is a cost of production. A change in production costs caused a change in supply for the Postal Service.

Step 3. Was the effect on supply positive or negative? Higher labor compensation leads to a lower quantity supplied of postal services at every given price, causing the supply curve for postal services to shift to the left, from S_0 to S_1.

Step 4. Compare the new equilibrium price and quantity to the original equilibrium price. The new equilibrium (E_1) occurs at a lower quantity and a higher price than the original equilibrium (E_0).

Figure 3.18 (b) shows the shift in demand in the following steps.

Step 1. Draw a demand and supply model to illustrate what the market for U.S. Postal Services looked like before this scenario starts. The demand curve D_0 and the supply curve S_0 show the original relationships. Note that this diagram is independent from the diagram in panel (a).

Step 2. Did the change described affect supply or demand? A change in tastes away from snail mail toward digital messages will cause a change in demand for the Postal Service.

Step 3. Was the effect on demand positive or negative? A change in tastes away from snailmail toward digital messages causes lower quantity demanded of postal services at every given price, causing the demand curve for postal services to shift to the left, from D_0 to D_1.

Step 4. Compare the new equilibrium price and quantity to the original equilibrium price. The new equilibrium (E_2) occurs at a lower quantity and a lower price than the original equilibrium (E_0).

The final step in a scenario where both supply and demand shift is to combine the two individual analyses to determine what happens to the equilibrium quantity and price. Graphically, we superimpose the previous two diagrams one on top of the other, as in **Figure 3.19**.

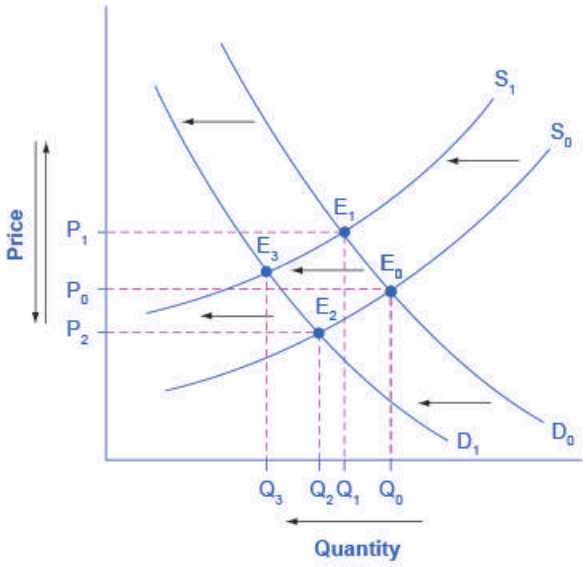


Figure 3.19 Combined Effect of Decreased Demand and Decreased Supply Supply and demand shifts cause changes in equilibrium price and quantity.

Following are the results:

Effect on Quantity: The effect of higher labor compensation on Postal Services because it raises the cost of production is to decrease the equilibrium quantity. The effect of a change in tastes away from snail mail is to decrease the equilibrium quantity. Since both shifts are to the left, the overall impact is a decrease in the equilibrium quantity of Postal Services (Q_3). This is easy to see graphically, since Q_3 is to the left of Q_0.

Effect on Price: The overall effect on price is more complicated. The effect of higher labor compensation on Postal Services, because it raises the cost of production, is to increase the equilibrium price. The effect of a change in tastes away from snail mail is to decrease the equilibrium price. Since the two effects are in opposite directions, unless we know the magnitudes of the two effects, the overall effect is unclear. This is not unusual. When both curves shift, typically we can determine the overall effect on price or on quantity, but not on both. In this case, we determined the overall effect on the equilibrium quantity, but not on the equilibrium price. In other cases, it might be the opposite.

The next Clear It Up feature focuses on the difference between shifts of supply or demand and movements along a curve.

Clear It Up

What is the difference between shifts of demand or supply versus movements along a demand or supply curve?

One common mistake in applying the demand and supply framework is to confuse the shift of a demand or a supply curve with movement along a demand or supply curve. As an example, consider a problem that asks whether a drought will increase or decrease the equilibrium quantity and equilibrium price of wheat. Lee, a student in an introductory economics class, might reason:

"Well, it is clear that a drought reduces supply, so I will shift back the supply curve, as in the shift from the original supply curve S_0 to S_1 on the diagram (Shift 1). The equilibrium moves from E_0 to E_1, the equilibrium quantity is lower and the equilibrium price is higher. Then, a higher price makes farmers more likely to supply the good, so the supply curve shifts right, as shows the shift from S_1 to S_2, shows on the diagram (Shift 2), so that the equilibrium now moves from E_1 to E_2. The higher price, however, also reduces demand and so causes demand to shift back, like the shift from the original demand curve, D_0 to D_1 on the diagram (labeled

Shift 3), and the equilibrium moves from E_2 to E_3."

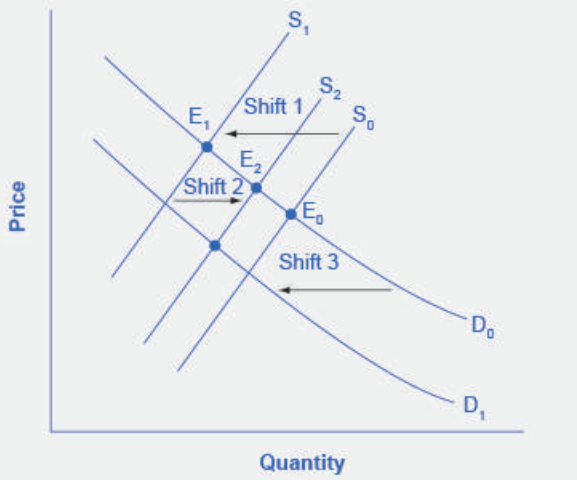


Figure 3.20 Shifts of Demand or Supply versus Movements along a Demand or Supply Curve A shift in one curve never causes a shift in the other curve. Rather, a shift in one curve causes a movement along the second curve.

At about this point, Lee suspects that this answer is headed down the wrong path. Think about what might be wrong with Lee's logic, and then read the answer that follows.

Answer: Lee's first step is correct: that is, a drought shifts back the supply curve of wheat and leads to a prediction of a lower equilibrium quantity and a higher equilibrium price. This corresponds to a movement along the original demand curve (D_0), from E_0 to E_1. The rest of Lee's argument is wrong, because it mixes up shifts in supply with quantity supplied, and shifts in demand with quantity demanded. A higher or lower price never shifts the supply curve, as suggested by the shift in supply from S_1 to S_2. Instead, a price change leads to a movement along a given supply curve. Similarly, a higher or lower price never shifts a demand curve, as suggested in the shift from D_0 to D_1. Instead, a price change leads to a movement along a given demand curve. Remember, a change in the price of a good never causes the demand or supply curve for that good to shift.

Think carefully about the timeline of events: What happens first, what happens next? What is cause, what is effect? If you keep the order right, you are more likely to get the analysis correct.

In the four-step analysis of how economic events affect equilibrium price and quantity, the movement from the old to the new equilibrium seems immediate. As a practical matter, however, prices and quantities often do not zoom straight to equilibrium. More realistically, when an economic event causes demand or supply to shift, prices and quantities set off in the general direction of equilibrium. Even as they are moving toward one new equilibrium, a subsequent change in demand or supply often pushes prices toward another equilibrium.

3.4 | Price Ceilings and Price Floors

By the end of this section, you will be able to:

- Explain price controls, price ceilings, and price floors
- Analyze demand and supply as a social adjustment mechanism

To this point in the chapter, we have been assuming that markets are free, that is, they operate with no government intervention. In this section, we will explore the outcomes, both anticipated and otherwise, when government does intervene in a market either to prevent the price of some good or service from rising "too high" or to prevent the price of some good or service from falling "too low".

Economists believe there are a small number of fundamental principles that explain how economic agents respond in different situations. Two of these principles, which we have already introduced, are the laws of demand and supply.

Governments can pass laws affecting market outcomes, but no law can negate these economic principles. Rather, the principles will become apparent in sometimes unexpected ways, which may undermine the intent of the government policy. This is one of the major conclusions of this section.

Controversy sometimes surrounds the prices and quantities established by demand and supply, especially for products that are considered necessities. In some cases, discontent over prices turns into public pressure on politicians, who may then pass legislation to prevent a certain price from climbing "too high" or falling "too low."

The demand and supply model shows how people and firms will react to the incentives that these laws provide to control prices, in ways that will often lead to undesirable consequences. Alternative policy tools can often achieve the desired goals of price control laws, while avoiding at least some of their costs and tradeoffs.

Price Ceilings

Laws that government enact to regulate prices are called **price controls**. Price controls come in two flavors. A **price ceiling** keeps a price from rising above a certain level (the "ceiling"), while a **price floor** keeps a price from falling below a given level (the "floor"). This section uses the demand and supply framework to analyze price ceilings. The next section discusses price floors.

A price ceiling is a legal maximum price that one pays for some good or service. A government imposes price ceilings in order to keep the price of some necessary good or service affordable. For example, in 2005 during Hurricane Katrina, the price of bottled water increased above $5 per gallon. As a result, many people called for price controls on bottled water to prevent the price from rising so high. In this particular case, the government did not impose a price ceiling, but there are other examples of where price ceilings did occur.

In many markets for goods and services, demanders outnumber suppliers. Consumers, who are also potential voters, sometimes unite behind a political proposal to hold down a certain price. In some cities, such as Albany, renters have pressed political leaders to pass rent control laws, a price ceiling that usually works by stating that landlords can raise rents by only a certain maximum percentage each year. Some of the best examples of rent control occur in urban areas such as New York, Washington D.C., or San Francisco.

Rent control becomes a politically hot topic when rents begin to rise rapidly. Everyone needs an affordable place to live. Perhaps a change in tastes makes a certain suburb or town a more popular place to live. Perhaps locally-based businesses expand, bringing higher incomes and more people into the area. Such changes can cause a change in the demand for rental housing, as **Figure 3.21** illustrates. The original equilibrium (E_0) lies at the intersection of supply curve S_0 and demand curve D_0, corresponding to an equilibrium price of $500 and an equilibrium quantity of 15,000 units of rental housing. The effect of greater income or a change in tastes is to shift the demand curve for rental housing to the right, as the data in **Table 3.7** shows and the shift from D_0 to D_1 on the graph. In this market, at the new equilibrium E_1, the price of a rental unit would rise to $600 and the equilibrium quantity would increase to 17,000 units.

Figure 3.21 A Price Ceiling Example—Rent Control The original intersection of demand and supply occurs at E_0. If demand shifts from D_0 to D_1, the new equilibrium would be at E_1—unless a price ceiling prevents the price from rising. If the price is not permitted to rise, the quantity supplied remains at 15,000. However, after the change in demand, the quantity demanded rises to 19,000, resulting in a shortage.

Price	Original Quantity Supplied	Original Quantity Demanded	New Quantity Demanded
$400	12,000	18,000	23,000
$500	15,000	15,000	19,000
$600	17,000	13,000	17,000
$700	19,000	11,000	15,000
$800	20,000	10,000	14,000

Table 3.7 Rent Control

Suppose that a city government passes a rent control law to keep the price at the original equilibrium of $500 for a typical apartment. In **Figure 3.21**, the horizontal line at the price of $500 shows the legally fixed maximum price set by the rent control law. However, the underlying forces that shifted the demand curve to the right are still there. At that price ($500), the quantity supplied remains at the same 15,000 rental units, but the quantity demanded is 19,000 rental units. In other words, the quantity demanded exceeds the quantity supplied, so there is a shortage of rental housing. One of the ironies of price ceilings is that while the price ceiling was intended to help renters, there are actually fewer apartments rented out under the price ceiling (15,000 rental units) than would be the case at the market rent of $600 (17,000 rental units).

Price ceilings do not simply benefit renters at the expense of landlords. Rather, some renters (or potential renters) lose their housing as landlords convert apartments to co-ops and condos. Even when the housing remains in the rental market, landlords tend to spend less on maintenance and on essentials like heating, cooling, hot water, and lighting. The first rule of economics is you do not get something for nothing—everything has an opportunity cost. Thus, if renters obtain “cheaper” housing than the market requires, they tend to also end up with lower quality housing.

Price ceilings are enacted in an attempt to keep prices low for those who need the product. However, when the market price is not allowed to rise to the equilibrium level, quantity demanded exceeds quantity supplied, and thus a shortage occurs. Those who manage to purchase the product at the lower price given by the price ceiling will benefit, but sellers of the product will suffer, along with those who are not able to purchase the product at all. Quality is also likely to deteriorate.

Price Floors

A price floor is the lowest price that one can legally pay for some good or service. Perhaps the best-known example

of a price floor is the minimum wage, which is based on the view that someone working full time should be able to afford a basic standard of living. The federal minimum wage in 2016 was $7.25 per hour, although some states and localities have a higher minimum wage. The federal minimum wage yields an annual income for a single person of $15,080, which is slightly higher than the Federal poverty line of $11,880. As the cost of living rises over time, the Congress periodically raises the federal minimum wage.

Price floors are sometimes called “price supports,” because they support a price by preventing it from falling below a certain level. Around the world, many countries have passed laws to create agricultural price supports. Farm prices and thus farm incomes fluctuate, sometimes widely. Even if, on average, farm incomes are adequate, some years they can be quite low. The purpose of price supports is to prevent these swings.

The most common way price supports work is that the government enters the market and buys up the product, adding to demand to keep prices higher than they otherwise would be. According to the Common Agricultural Policy reform passed in 2013, the European Union (EU) will spend about 60 billion euros per year, or 67 billion dollars per year (with the November 2016 exchange rate), or roughly 38% of the EU budget, on price supports for Europe’s farmers from 2014 to 2020.

Figure 3.22 illustrates the effects of a government program that assures a price above the equilibrium by focusing on the market for wheat in Europe. In the absence of government intervention, the price would adjust so that the quantity supplied would equal the quantity demanded at the equilibrium point E_0, with price P_0 and quantity Q_0. However, policies to keep prices high for farmers keeps the price above what would have been the market equilibrium level—the price Pf shown by the dashed horizontal line in the diagram. The result is a quantity supplied in excess of the quantity demanded (Qd). When quantity supplied exceeds quantity demanded, a surplus exists.

Economists estimate that the high-income areas of the world, including the United States, Europe, and Japan, spend roughly $1 billion per day in supporting their farmers. If the government is willing to purchase the excess supply (or to provide payments for others to purchase it), then farmers will benefit from the price floor, but taxpayers and consumers of food will pay the costs. Agricultural economists and policy makers have offered numerous proposals for reducing farm subsidies. In many countries, however, political support for subsidies for farmers remains strong. This is either because the population views this as supporting the traditional rural way of life or because of industry's lobbying power of the agro-business.

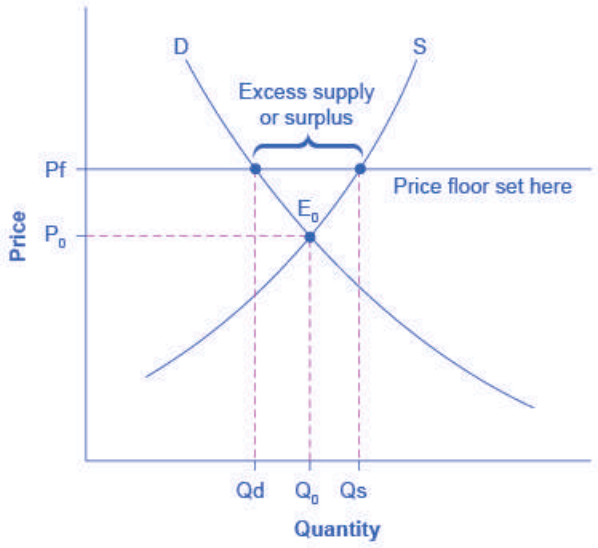


Figure 3.22 European Wheat Prices: A Price Floor Example The intersection of demand (D) and supply (S) would be at the equilibrium point E_0. However, a price floor set at Pf holds the price above E_0 and prevents it from falling. The result of the price floor is that the quantity supplied Qs exceeds the quantity demanded Qd. There is excess supply, also called a surplus.

3.5 | Demand, Supply, and Efficiency

By the end of this section, you will be able to:

- Contrast consumer surplus, producer surplus, and social surplus
- Explain why price floors and price ceilings can be inefficient
- Analyze demand and supply as a social adjustment mechanism

The familiar demand and supply diagram holds within it the concept of economic efficiency. One typical way that economists define efficiency is when it is impossible to improve the situation of one party without imposing a cost on another. Conversely, if a situation is inefficient, it becomes possible to benefit at least one party without imposing costs on others.

Efficiency in the demand and supply model has the same basic meaning: The economy is getting as much benefit as possible from its scarce resources and all the possible gains from trade have been achieved. In other words, the optimal amount of each good and service is produced and consumed.

Consumer Surplus, Producer Surplus, Social Surplus

Consider a market for tablet computers, as **Figure 3.23** shows. The equilibrium price is $80 and the equilibrium quantity is 28 million. To see the benefits to consumers, look at the segment of the demand curve above the equilibrium point and to the left. This portion of the demand curve shows that at least some demanders would have been willing to pay more than $80 for a tablet.

For example, point J shows that if the price were $90, 20 million tablets would be sold. Those consumers who would have been willing to pay $90 for a tablet based on the utility they expect to receive from it, but who were able to pay the equilibrium price of $80, clearly received a benefit beyond what they had to pay. Remember, the demand curve traces consumers' willingness to pay for different quantities. The amount that individuals would have been willing to pay, minus the amount that they actually paid, is called **consumer surplus**. Consumer surplus is the area labeled F—that is, the area above the market price and below the demand curve.

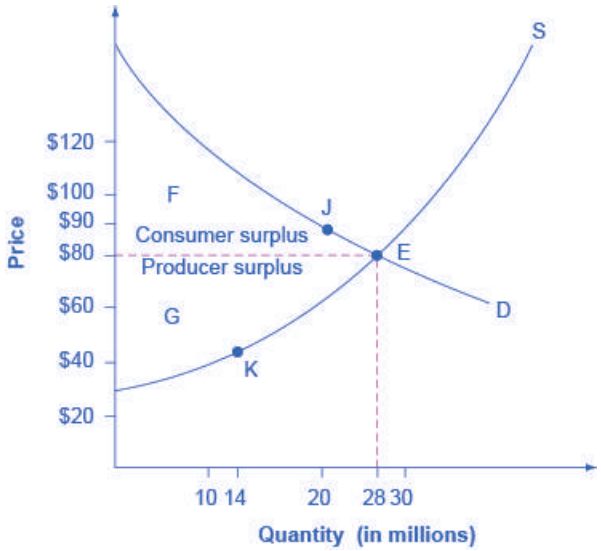


Figure 3.23 Consumer and Producer Surplus The somewhat triangular area labeled by F shows the area of consumer surplus, which shows that the equilibrium price in the market was less than what many of the consumers were willing to pay. Point J on the demand curve shows that, even at the price of $90, consumers would have been willing to purchase a quantity of 20 million. The somewhat triangular area labeled by G shows the area of producer surplus, which shows that the equilibrium price received in the market was more than what many of the producers were willing to accept for their products. For example, point K on the supply curve shows that at a price of $45, firms would have been willing to supply a quantity of 14 million.

The supply curve shows the quantity that firms are willing to supply at each price. For example, point K in **Figure**

3.23 illustrates that, at $45, firms would still have been willing to supply a quantity of 14 million. Those producers who would have been willing to supply the tablets at $45, but who were instead able to charge the equilibrium price of $80, clearly received an extra benefit beyond what they required to supply the product. The amount that a seller is paid for a good minus the seller's actual cost is called **producer surplus**. In Figure 3.23, producer surplus is the area labeled G—that is, the area between the market price and the segment of the supply curve below the equilibrium.

The sum of consumer surplus and producer surplus is **social surplus**, also referred to as **economic surplus** or **total surplus**. In Figure 3.23 we show social surplus as the area F + G. Social surplus is larger at equilibrium quantity and price than it would be at any other quantity. This demonstrates the economic efficiency of the market equilibrium. In addition, at the efficient level of output, it is impossible to produce greater consumer surplus without reducing producer surplus, and it is impossible to produce greater producer surplus without reducing consumer surplus.

Inefficiency of Price Floors and Price Ceilings

The imposition of a price floor or a price ceiling will prevent a market from adjusting to its equilibrium price and quantity, and thus will create an inefficient outcome. However, there is an additional twist here. Along with creating inefficiency, price floors and ceilings will also transfer some consumer surplus to producers, or some producer surplus to consumers.

Imagine that several firms develop a promising but expensive new drug for treating back pain. If this therapy is left to the market, the equilibrium price will be $600 per month and 20,000 people will use the drug, as shown in Figure 3.24 (a). The original level of consumer surplus is T + U and producer surplus is V + W + X. However, the government decides to impose a price ceiling of $400 to make the drug more affordable. At this price ceiling, firms in the market now produce only 15,000.

As a result, two changes occur. First, an inefficient outcome occurs and the total surplus of society is reduced. The loss in social surplus that occurs when the economy produces at an inefficient quantity is called **deadweight loss**. In a very real sense, it is like money thrown away that benefits no one. In Figure 3.24 (a), the deadweight loss is the area U + W. When deadweight loss exists, it is possible for both consumer and producer surplus to be higher, in this case because the price control is blocking some suppliers and demanders from transactions they would both be willing to make.

A second change from the price ceiling is that some of the producer surplus is transferred to consumers. After the price ceiling is imposed, the new consumer surplus is T + V, while the new producer surplus is X. In other words, the price ceiling transfers the area of surplus (V) from producers to consumers. Note that the gain to consumers is less than the loss to producers, which is just another way of seeing the deadweight loss.

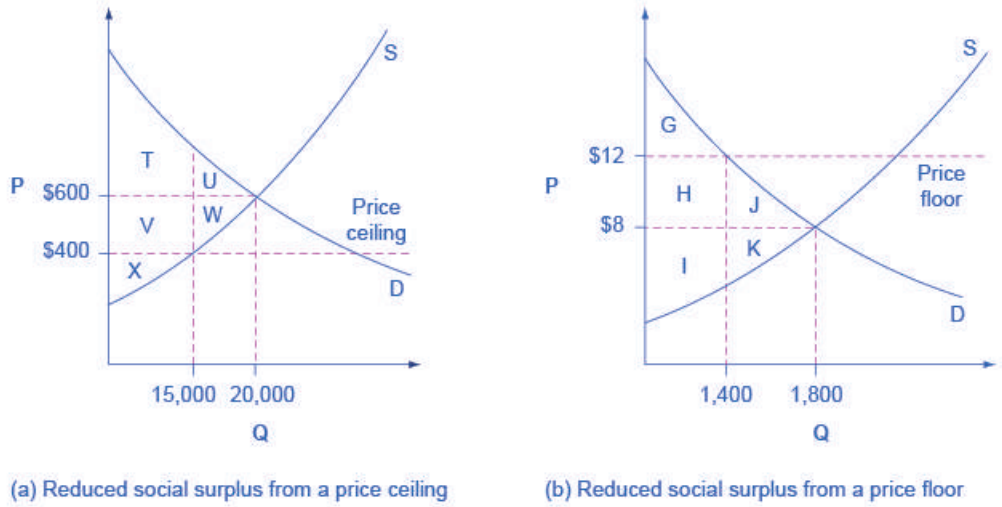


Figure 3.24 Efficiency and Price Floors and Ceilings (a) The original equilibrium price is $600 with a quantity of 20,000. Consumer surplus is T + U, and producer surplus is V + W + X. A price ceiling is imposed at $400, so firms in the market now produce only a quantity of 15,000. As a result, the new consumer surplus is T + V, while the new producer surplus is X. (b) The original equilibrium is $8 at a quantity of 1,800. Consumer surplus is G + H + J, and producer surplus is I + K. A price floor is imposed at $12, which means that quantity demanded falls to 1,400. As a result, the new consumer surplus is G, and the new producer surplus is H + I.

Figure 3.24 (b) shows a price floor example using a string of struggling movie theaters, all in the same city. The current equilibrium is $8 per movie ticket, with 1,800 people attending movies. The original consumer surplus is G + H + J, and producer surplus is I + K. The city government is worried that movie theaters will go out of business, reducing the entertainment options available to citizens, so it decides to impose a price floor of $12 per ticket. As a result, the quantity demanded of movie tickets falls to 1,400. The new consumer surplus is G, and the new producer surplus is H + I. In effect, the price floor causes the area H to be transferred from consumer to producer surplus, but also causes a deadweight loss of J + K.

This analysis shows that a price ceiling, like a law establishing rent controls, will transfer some producer surplus to consumers—which helps to explain why consumers often favor them. Conversely, a price floor like a guarantee that farmers will receive a certain price for their crops will transfer some consumer surplus to producers, which explains why producers often favor them. However, both price floors and price ceilings block some transactions that buyers and sellers would have been willing to make, and creates deadweight loss. Removing such barriers, so that prices and quantities can adjust to their equilibrium level, will increase the economy's social surplus.

Demand and Supply as a Social Adjustment Mechanism

The demand and supply model emphasizes that prices are not set only by demand or only by supply, but by the interaction between the two. In 1890, the famous economist Alfred Marshall wrote that asking whether supply or demand determined a price was like arguing "whether it is the upper or the under blade of a pair of scissors that cuts a piece of paper." The answer is that both blades of the demand and supply scissors are always involved.

The adjustments of equilibrium price and quantity in a market-oriented economy often occur without much government direction or oversight. If the coffee crop in Brazil suffers a terrible frost, then the supply curve of coffee shifts to the left and the price of coffee rises. Some people—call them the coffee addicts—continue to drink coffee and pay the higher price. Others switch to tea or soft drinks. No government commission is needed to figure out how to adjust coffee prices, which companies will be allowed to process the remaining supply, which supermarkets in which cities will get how much coffee to sell, or which consumers will ultimately be allowed to drink the brew. Such adjustments in response to price changes happen all the time in a market economy, often so smoothly and rapidly that we barely notice them.

Think for a moment of all the seasonal foods that are available and inexpensive at certain times of the year, like fresh corn in midsummer, but more expensive at other times of the year. People alter their diets and restaurants alter their menus in response to these fluctuations in prices without fuss or fanfare. For both the U.S. economy and the world

economy as a whole, markets—that is, demand and supply—are the primary social mechanism for answering the basic questions about what is produced, how it is produced, and for whom it is produced.

Bring it Home

Why Can We Not Get Enough of Organic?

Organic food is grown without synthetic pesticides, chemical fertilizers or genetically modified seeds. In recent decades, the demand for organic products has increased dramatically. The Organic Trade Association reported sales increased from $1 billion in 1990 to $35.1 billion in 2013, more than 90% of which were sales of food products.

Why, then, are organic foods more expensive than their conventional counterparts? The answer is a clear application of the theories of supply and demand. As people have learned more about the harmful effects of chemical fertilizers, growth hormones, pesticides and the like from large-scale factory farming, our tastes and preferences for safer, organic foods have increased. This change in tastes has been reinforced by increases in income, which allow people to purchase pricier products, and has made organic foods more mainstream. This has led to an increased demand for organic foods. Graphically, the demand curve has shifted right, and we have moved up the supply curve as producers have responded to the higher prices by supplying a greater quantity.

In addition to the movement along the supply curve, we have also had an increase in the number of farmers converting to organic farming over time. This is represented by a shift to the right of the supply curve. Since both demand and supply have shifted to the right, the resulting equilibrium quantity of organic foods is definitely higher, but the price will only fall when the increase in supply is larger than the increase in demand. We may need more time before we see lower prices in organic foods. Since the production costs of these foods may remain higher than conventional farming, because organic fertilizers and pest management techniques are more expensive, they may never fully catch up with the lower prices of non-organic foods.

As a final, specific example: The Environmental Working Group's "Dirty Dozen" list of fruits and vegetables, which test high for pesticide residue even after washing, was released in April 2013. The inclusion of strawberries on the list has led to an increase in demand for organic strawberries, resulting in both a higher equilibrium price and quantity of sales.

KEY TERMS

ceteris paribus other things being equal

complements goods that are often used together so that consumption of one good tends to enhance consumption of the other

consumer surplus the extra benefit consumers receive from buying a good or service, measured by what the individuals would have been willing to pay minus the amount that they actually paid

deadweight loss the loss in social surplus that occurs when a market produces an inefficient quantity

demand the relationship between price and the quantity demanded of a certain good or service

demand curve a graphic representation of the relationship between price and quantity demanded of a certain good or service, with quantity on the horizontal axis and the price on the vertical axis

demand schedule a table that shows a range of prices for a certain good or service and the quantity demanded at each price

economic surplus see social surplus

equilibrium the situation where quantity demanded is equal to the quantity supplied; the combination of price and quantity where there is no economic pressure from surpluses or shortages that would cause price or quantity to change

equilibrium price the price where quantity demanded is equal to quantity supplied

equilibrium quantity the quantity at which quantity demanded and quantity supplied are equal for a certain price level

excess demand at the existing price, the quantity demanded exceeds the quantity supplied; also called a shortage

excess supply at the existing price, quantity supplied exceeds the quantity demanded; also called a surplus

factors of production the resources such as labor, materials, and machinery that are used to produce goods and services; also called inputs

inferior good a good in which the quantity demanded falls as income rises, and in which quantity demanded rises and income falls

inputs the resources such as labor, materials, and machinery that are used to produce goods and services; also called factors of production

law of demand the common relationship that a higher price leads to a lower quantity demanded of a certain good or service and a lower price leads to a higher quantity demanded, while all other variables are held constant

law of supply the common relationship that a higher price leads to a greater quantity supplied and a lower price leads to a lower quantity supplied, while all other variables are held constant

normal good a good in which the quantity demanded rises as income rises, and in which quantity demanded falls as income falls

price what a buyer pays for a unit of the specific good or service

price ceiling a legal maximum price

price control government laws to regulate prices instead of letting market forces determine prices

price floor a legal minimum price

producer surplus the extra benefit producers receive from selling a good or service, measured by the price the producer actually received minus the price the producer would have been willing to accept

quantity demanded the total number of units of a good or service consumers are willing to purchase at a given price

quantity supplied the total number of units of a good or service producers are willing to sell at a given price

shift in demand when a change in some economic factor (other than price) causes a different quantity to be demanded at every price

shift in supply when a change in some economic factor (other than price) causes a different quantity to be supplied at every price

shortage at the existing price, the quantity demanded exceeds the quantity supplied; also called excess demand

social surplus the sum of consumer surplus and producer surplus

substitute a good that can replace another to some extent, so that greater consumption of one good can mean less of the other

supply the relationship between price and the quantity supplied of a certain good or service

supply curve a line that shows the relationship between price and quantity supplied on a graph, with quantity supplied on the horizontal axis and price on the vertical axis

supply schedule a table that shows a range of prices for a good or service and the quantity supplied at each price

surplus at the existing price, quantity supplied exceeds the quantity demanded; also called excess supply

total surplus see social surplus

KEY CONCEPTS AND SUMMARY

3.1 Demand, Supply, and Equilibrium in Markets for Goods and Services

A demand schedule is a table that shows the quantity demanded at different prices in the market. A demand curve shows the relationship between quantity demanded and price in a given market on a graph. The law of demand states that a higher price typically leads to a lower quantity demanded.

A supply schedule is a table that shows the quantity supplied at different prices in the market. A supply curve shows the relationship between quantity supplied and price on a graph. The law of supply says that a higher price typically leads to a higher quantity supplied.

The equilibrium price and equilibrium quantity occur where the supply and demand curves cross. The equilibrium occurs where the quantity demanded is equal to the quantity supplied. If the price is below the equilibrium level, then the quantity demanded will exceed the quantity supplied. Excess demand or a shortage will exist. If the price is above the equilibrium level, then the quantity supplied will exceed the quantity demanded. Excess supply or a surplus will exist. In either case, economic pressures will push the price toward the equilibrium level.

3.2 Shifts in Demand and Supply for Goods and Services

Economists often use the *ceteris paribus* or “other things being equal” assumption: while examining the economic impact of one event, all other factors remain unchanged for analysis purposes. Factors that can shift the demand curve for goods and services, causing a different quantity to be demanded at any given price, include changes in tastes, population, income, prices of substitute or complement goods, and expectations about future conditions and prices. Factors that can shift the supply curve for goods and services, causing a different quantity to be supplied at any given price, include input prices, natural conditions, changes in technology, and government taxes, regulations, or subsidies.

3.3 Changes in Equilibrium Price and Quantity: The Four-Step Process

When using the supply and demand framework to think about how an event will affect the equilibrium price and quantity, proceed through four steps: (1) sketch a supply and demand diagram to think about what the market looked like before the event; (2) decide whether the event will affect supply or demand; (3) decide whether the effect on supply or demand is negative or positive, and draw the appropriate shifted supply or demand curve; (4) compare the new equilibrium price and quantity to the original ones.

3.4 Price Ceilings and Price Floors

Price ceilings prevent a price from rising above a certain level. When a price ceiling is set below the equilibrium price, quantity demanded will exceed quantity supplied, and excess demand or shortages will result. Price floors prevent a price from falling below a certain level. When a price floor is set above the equilibrium price, quantity supplied will exceed quantity demanded, and excess supply or surpluses will result. Price floors and price ceilings often lead to unintended consequences.

3.5 Demand, Supply, and Efficiency

Consumer surplus is the gap between the price that consumers are willing to pay, based on their preferences, and the market equilibrium price. Producer surplus is the gap between the price for which producers are willing to sell a product, based on their costs, and the market equilibrium price. Social surplus is the sum of consumer surplus and producer surplus. Total surplus is larger at the equilibrium quantity and price than it will be at any other quantity and price. Deadweight loss is loss in total surplus that occurs when the economy produces at an inefficient quantity.

SELF-CHECK QUESTIONS

1. Review **Figure 3.4**. Suppose the price of gasoline is $1.60 per gallon. Is the quantity demanded higher or lower than at the equilibrium price of $1.40 per gallon? What about the quantity supplied? Is there a shortage or a surplus in the market? If so, how much?

2. Why do economists use the *ceteris paribus* assumption?

3. In an analysis of the market for paint, an economist discovers the facts listed below. State whether each of these changes will affect supply or demand, and in what direction.

a. There have recently been some important cost-saving inventions in the technology for making paint.
b. Paint is lasting longer, so that property owners need not repaint as often.
c. Because of severe hailstorms, many people need to repaint now.
d. The hailstorms damaged several factories that make paint, forcing them to close down for several months.

4. Many changes are affecting the market for oil. Predict how each of the following events will affect the equilibrium price and quantity in the market for oil. In each case, state how the event will affect the supply and demand diagram. Create a sketch of the diagram if necessary.

a. Cars are becoming more fuel efficient, and therefore get more miles to the gallon.
b. The winter is exceptionally cold.
c. A major discovery of new oil is made off the coast of Norway.
d. The economies of some major oil-using nations, like Japan, slow down.
e. A war in the Middle East disrupts oil-pumping schedules.
f. Landlords install additional insulation in buildings.
g. The price of solar energy falls dramatically.
h. Chemical companies invent a new, popular kind of plastic made from oil.

5. Let’s think about the market for air travel. From August 2014 to January 2015, the price of jet fuel increased roughly 47%. Using the four-step analysis, how do you think this fuel price increase affected the equilibrium price and quantity of air travel?

6. A tariff is a tax on imported goods. Suppose the U.S. government cuts the tariff on imported flat screen televisions. Using the four-step analysis, how do you think the tariff reduction will affect the equilibrium price and quantity of flat screen TVs?

7. What is the effect of a price ceiling on the quantity demanded of the product? What is the effect of a price ceiling on the quantity supplied? Why exactly does a price ceiling cause a shortage?

8. Does a price ceiling change the equilibrium price?

9. What would be the impact of imposing a price floor below the equilibrium price?

10. Does a price ceiling increase or decrease the number of transactions in a market? Why? What about a price floor?

11. If a price floor benefits producers, why does a price floor reduce social surplus?

REVIEW QUESTIONS

12. What determines the level of prices in a market?

13. What does a downward-sloping demand curve mean about how buyers in a market will react to a higher price?

14. Will demand curves have the same exact shape in all markets? If not, how will they differ?

15. Will supply curves have the same shape in all markets? If not, how will they differ?

16. What is the relationship between quantity demanded and quantity supplied at equilibrium? What is the relationship when there is a shortage? What is the relationship when there is a surplus?

17. How can you locate the equilibrium point on a demand and supply graph?

18. If the price is above the equilibrium level, would you predict a surplus or a shortage? If the price is below the equilibrium level, would you predict a surplus or a shortage? Why?

19. When the price is above the equilibrium, explain how market forces move the market price to equilibrium. Do the same when the price is below the equilibrium.

20. What is the difference between the demand and the quantity demanded of a product, say milk? Explain in words and show the difference on a graph with a demand curve for milk.

21. What is the difference between the supply and the quantity supplied of a product, say milk? Explain in words and show the difference on a graph with the supply curve for milk.

22. When analyzing a market, how do economists deal with the problem that many factors that affect the market are changing at the same time?

23. Name some factors that can cause a shift in the demand curve in markets for goods and services.

24. Name some factors that can cause a shift in the supply curve in markets for goods and services.

25. How does one analyze a market where both demand and supply shift?

26. What causes a movement along the demand curve? What causes a movement along the supply curve?

27. Does a price ceiling attempt to make a price higher or lower?

28. How does a price ceiling set below the equilibrium level affect quantity demanded and quantity supplied?

29. Does a price floor attempt to make a price higher or lower?

30. How does a price floor set above the equilibrium level affect quantity demanded and quantity supplied?

31. What is consumer surplus? How is it illustrated on a demand and supply diagram?

32. What is producer surplus? How is it illustrated on a demand and supply diagram?

33. What is total surplus? How is it illustrated on a demand and supply diagram?

34. What is the relationship between total surplus and economic efficiency?

35. What is deadweight loss?

CRITICAL THINKING QUESTIONS

36. Review **Figure 3.4**. Suppose the government decided that, since gasoline is a necessity, its price should be legally capped at $1.30 per gallon. What do you anticipate would be the outcome in the gasoline market?

37. Explain why the following statement is false: "In the goods market, no buyer would be willing to pay more than the equilibrium price."

38. Explain why the following statement is false: "In the goods market, no seller would be willing to sell for less than the equilibrium price."

39. Consider the demand for hamburgers. If the price of a substitute good (for example, hot dogs) increases and the price of a complement good (for example, hamburger buns) increases, can you tell for sure what will happen to the demand for hamburgers? Why or why not? Illustrate your answer with a graph.

40. How do you suppose the demographics of an aging population of "Baby Boomers" in the United States will affect the demand for milk? Justify your answer.

41. We know that a change in the price of a product causes a movement along the demand curve. Suppose consumers believe that prices will be rising in the future. How will that affect demand for the product in the present? Can you show this graphically?

42. Suppose there is a soda tax to curb obesity. What should a reduction in the soda tax do to the supply of sodas and to the equilibrium price and quantity? Can you show this graphically? *Hint*: Assume that the soda tax is collected from the sellers.

43. Use the four-step process to analyze the impact of the advent of the iPod (or other portable digital music players) on the equilibrium price and quantity of the Sony Walkman (or other portable audio cassette players).

44. Use the four-step process to analyze the impact of a reduction in tariffs on imports of iPods on the equilibrium price and quantity of Sony Walkman-type products.

45. Suppose both of these events took place at the same time. Combine your analyses of the impacts of the iPod and the tariff reduction to determine the likely impact on the equilibrium price and quantity of Sony Walkman-type products. Show your answer graphically.

46. Most government policy decisions have winners and losers. What are the effects of raising the minimum wage? It is more complex than simply producers lose and workers gain. Who are the winners and who are the losers, and what exactly do they win and lose? To what extent does the policy change achieve its goals?

47. Agricultural price supports result in governments holding large inventories of agricultural products. Why do you think the government cannot simply give the products away to poor people?

48. Can you propose a policy that would induce the market to supply more rental housing units?

49. What term would an economist use to describe what happens when a shopper gets a "good deal" on a product?

50. Explain why voluntary transactions improve social welfare.

51. Why would a free market never operate at a quantity greater than the equilibrium quantity? *Hint:* What would be required for a transaction to occur at that quantity?

PROBLEMS

52. Review **Figure 3.4** again. Suppose the price of gasoline is $1.00. Will the quantity demanded be lower or higher than at the equilibrium price of $1.40 per gallon? Will the quantity supplied be lower or higher? Is there a shortage or a surplus in the market? If so, of how much?

53. Table 3.8 shows information on the demand and supply for bicycles, where the quantities of bicycles are measured in thousands.

Price	Qd	Qs
$120	50	36
$150	40	40
$180	32	48
$210	28	56
$240	24	70

Table 3.8

a. What is the quantity demanded and the quantity supplied at a price of $210?
b. At what price is the quantity supplied equal to 48,000?
c. Graph the demand and supply curve for bicycles. How can you determine the equilibrium price and quantity from the graph? How can you determine the equilibrium price and quantity from the table? What are the equilibrium price and equilibrium quantity?
d. If the price was $120, what would the quantities demanded and supplied be? Would a shortage or surplus exist? If so, how large would the shortage or surplus be?

54. The computer market in recent years has seen many more computers sell at much lower prices. What shift in demand or supply is most likely to explain this outcome? Sketch a demand and supply diagram and explain your reasoning for each.

a. A rise in demand
b. A fall in demand
c. A rise in supply
d. A fall in supply

55. Table 3.9 illustrates the market's demand and supply for cheddar cheese. Graph the data and find the equilibrium. Next, create a table showing the change in quantity demanded or quantity supplied, and a graph of the new equilibrium, in each of the following situations:

a. The price of milk, a key input for cheese production, rises, so that the supply decreases by 80 pounds at every price.
b. A new study says that eating cheese is good for your health, so that demand increases by 20% at every price.

Price per Pound	Qd	Qs
$3.00	750	540
$3.20	700	600
$3.40	**650**	**650**
$3.60	620	700
$3.80	600	720
$4.00	590	730

Table 3.9

56. Table 3.10 shows the supply and demand for movie tickets in a city. Graph demand and supply and identify the equilibrium. Then calculate in a table and graph the effect of the following two changes.

a. Three new nightclubs open. They offer decent bands and have no cover charge, but make their money by selling food and drink. As a result, demand for movie tickets falls by six units at every price.
b. The city eliminates a tax that it placed on all local entertainment businesses. The result is that the quantity supplied of movies at any given price increases by 10%.

Price per Pound	Qd	Qs
$5.00	26	16
$6.00	24	18
$7.00	22	20
$8.00	21	21
$9.00	20	22

Table 3.10

57. A low-income country decides to set a price ceiling on bread so it can make sure that bread is affordable to the poor. Table 3.11 provides the conditions of demand and supply. What are the equilibrium price and equilibrium quantity before the price ceiling? What will the excess demand or the shortage (that is, quantity demanded minus quantity supplied) be if the price ceiling is set at $2.40? At $2.00? At $3.60?

Price	Qd	Qs
$1.60	9,000	5,000
$2.00	8,500	5,500
$2.40	8,000	6,400
$2.80	**7,500**	**7,500**
$3.20	7,000	9,000
$3.60	6,500	11,000
$4.00	6,000	15,000

Table 3.11

4 | Labor and Financial Markets

Figure 4.1 People often think of demand and supply in relation to goods, but labor markets, such as the nursing profession, can also apply to this analysis. (Credit: modification of work by "Fotos GOVBA"/Flickr Creative Commons)

Bring it Home

Baby Boomers Come of Age

The Census Bureau reports that as of 2013, 20% of the U.S. population was over 60 years old, which means that almost 63 million people are reaching an age when they will need increased medical care.

The baby boomer population, the group born between 1946 and 1964, is comprised of approximately 74 million people who have just reached retirement age. As this population grows older, they will be faced with common healthcare issues such as heart conditions, arthritis, and Alzheimer's that may require hospitalization, long-term, or at-home nursing care. Aging baby boomers and advances in life-saving and life-extending technologies will increase the demand for healthcare and nursing. Additionally, the Affordable Care Act, which expands access to healthcare for millions of Americans, has further increase the demand, although with the election of Donald J. Trump, this increase may not be sustained.

According to the Bureau of Labor Statistics, registered nursing jobs are expected to increase by 16% between 2014 and 2024. The median annual wage of $67,490 (in 2015) is also expected to increase. The BLS forecasts that 439,000 new nurses will be in demand by 2022.

These data tell us, as economists, that the market for healthcare professionals, and nurses in particular, will face several challenges. Our study of supply and demand will help us to analyze what might happen in the

labor market for nursing and other healthcare professionals, as we will discuss in the second half of this case at the end of the chapter.

Introduction to Labor and Financial Markets

In this chapter, you will learn about:

- Demand and Supply at Work in Labor Markets
- Demand and Supply in Financial Markets
- The Market System as an Efficient Mechanism for Information

The theories of supply and demand do not apply just to markets for goods. They apply to any market, even markets for things we may not think of as goods and services like labor and financial services. Labor markets are markets for employees or jobs. Financial services markets are markets for saving or borrowing.

When we think about demand and supply curves in goods and services markets, it is easy to picture the demanders and suppliers: businesses produce the products and households buy them. Who are the demanders and suppliers in labor and financial service markets? In labor markets job seekers (individuals) are the suppliers of labor, while firms and other employers who hire labor are the demanders for labor. In financial markets, any individual or firm who saves contributes to the supply of money, and any who borrows (person, firm, or government) contributes to the demand for money.

As a college student, you most likely participate in both labor and financial markets. Employment is a fact of life for most college students: According to the National Center for Educational Statistics, in 2013 40% of full-time college students and 76% of part-time college students were employed. Most college students are also heavily involved in financial markets, primarily as borrowers. Among full-time students, about half take out a loan to help finance their education each year, and those loans average about $6,000 per year. Many students also borrow for other expenses, like purchasing a car. As this chapter will illustrate, we can analyze labor markets and financial markets with the same tools we use to analyze demand and supply in the goods markets.

4.1 | Demand and Supply at Work in Labor Markets

By the end of this section, you will be able to:

- Predict shifts in the demand and supply curves of the labor market
- Explain the impact of new technology on the demand and supply curves of the labor market
- Explain price floors in the labor market such as minimum wage or a living wage

Markets for labor have demand and supply curves, just like markets for goods. The law of demand applies in labor markets this way: A higher **salary** or **wage**—that is, a higher price in the labor market—leads to a decrease in the quantity of labor demanded by employers, while a lower salary or wage leads to an increase in the quantity of labor demanded. The law of supply functions in labor markets, too: A higher price for labor leads to a higher quantity of labor supplied; a lower price leads to a lower quantity supplied.

Equilibrium in the Labor Market

In 2015, about 35,000 registered nurses worked in the Minneapolis-St. Paul-Bloomington, Minnesota-Wisconsin metropolitan area, according to the BLS. They worked for a variety of employers: hospitals, doctors' offices, schools, health clinics, and nursing homes. **Figure 4.2** illustrates how demand and supply determine equilibrium in this labor market. The demand and supply schedules in **Table 4.1** list the quantity supplied and quantity demanded of nurses at different salaries.

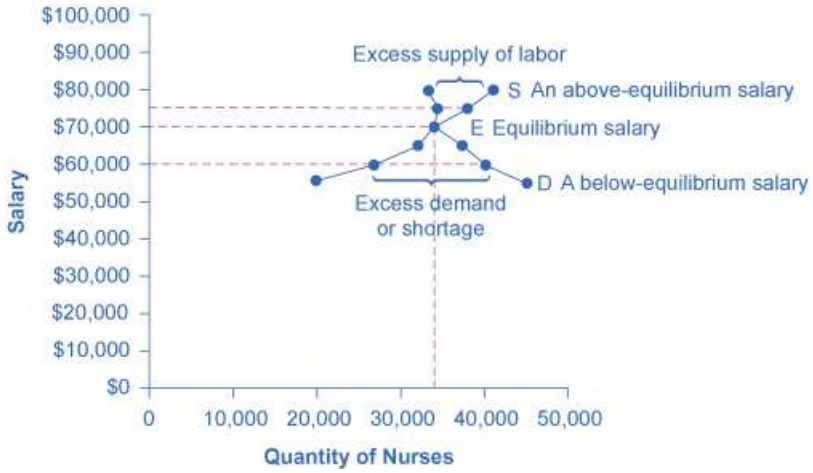


Figure 4.2 Labor Market Example: Demand and Supply for Nurses in Minneapolis-St. Paul-Bloomington The demand curve (D) of those employers who want to hire nurses intersects with the supply curve (S) of those who are qualified and willing to work as nurses at the equilibrium point (E). The equilibrium salary is $70,000 and the equilibrium quantity is 34,000 nurses. At an above-equilibrium salary of $75,000, quantity supplied increases to 38,000, but the quantity of nurses demanded at the higher pay declines to 33,000. At this above-equilibrium salary, an excess supply or surplus of nurses would exist. At a below-equilibrium salary of $60,000, quantity supplied declines to 27,000, while the quantity demanded at the lower wage increases to 40,000 nurses. At this below-equilibrium salary, excess demand or a shortage exists.

Annual Salary	Quantity Demanded	Quantity Supplied
$55,000	45,000	20,000
$60,000	40,000	27,000
$65,000	37,000	31,000
$70,000	34,000	34,000
$75,000	33,000	38,000
$80,000	32,000	41,000

Table 4.1 Demand and Supply of Nurses in Minneapolis-St. Paul-Bloomington

The horizontal axis shows the quantity of nurses hired. In this example we measure labor by number of workers, but another common way to measure the quantity of labor is by the number of hours worked. The vertical axis shows the price for nurses' labor—that is, how much they are paid. In the real world, this "price" would be total labor compensation: salary plus benefits. It is not obvious, but benefits are a significant part (as high as 30 percent) of labor compensation. In this example we measure the price of labor by salary on an annual basis, although in other cases we could measure the price of labor by monthly or weekly pay, or even the wage paid per hour. As the salary for nurses rises, the quantity demanded will fall. Some hospitals and nursing homes may reduce the number of nurses they hire, or they may lay off some of their existing nurses, rather than pay them higher salaries. Employers who face higher nurses' salaries may also try to replace some nursing functions by investing in physical equipment, like computer monitoring and diagnostic systems to monitor patients, or by using lower-paid health care aides to reduce the number of nurses they need.

As the salary for nurses rises, the quantity supplied will rise. If nurses' salaries in Minneapolis-St. Paul-Bloomington are higher than in other cities, more nurses will move to Minneapolis-St. Paul-Bloomington to find jobs, more people will be willing to train as nurses, and those currently trained as nurses will be more likely to pursue nursing as a full-time job. In other words, there will be more nurses looking for jobs in the area.

At **equilibrium**, the quantity supplied and the quantity demanded are equal. Thus, every employer who wants to hire a nurse at this equilibrium wage can find a willing worker, and every nurse who wants to work at this equilibrium salary can find a job. In **Figure 4.2**, the supply curve (S) and demand curve (D) intersect at the equilibrium point (E). The equilibrium quantity of nurses in the Minneapolis-St. Paul-Bloomington area is 34,000, and the equilibrium salary is $70,000 per year. This example simplifies the nursing market by focusing on the "average" nurse. In reality, of course, the market for nurses actually comprises many smaller markets, like markets for nurses with varying degrees of experience and credentials. Many markets contain closely related products that differ in quality. For instance, even a simple product like gasoline comes in regular, premium, and super-premium, each with a different price. Even in such cases, discussing the average price of gasoline, like the average salary for nurses, can still be useful because it reflects what is happening in most of the submarkets.

When the price of labor is not at the equilibrium, economic incentives tend to move salaries toward the equilibrium. For example, if salaries for nurses in Minneapolis-St. Paul-Bloomington were above the equilibrium at $75,000 per year, then 38,000 people want to work as nurses, but employers want to hire only 33,000 nurses. At that above-equilibrium salary, excess supply or a surplus results. In a situation of excess supply in the **labor market**, with many applicants for every job opening, employers will have an incentive to offer lower wages than they otherwise would have. Nurses' salary will move down toward equilibrium.

In contrast, if the salary is below the equilibrium at, say, $60,000 per year, then a situation of excess demand or a shortage arises. In this case, employers encouraged by the relatively lower wage want to hire 40,000 nurses, but only 27,000 individuals want to work as nurses at that salary in Minneapolis-St. Paul-Bloomington. In response to the shortage, some employers will offer higher pay to attract the nurses. Other employers will have to match the higher pay to keep their own employees. The higher salaries will encourage more nurses to train or work in Minneapolis-St. Paul-Bloomington. Again, price and quantity in the labor market will move toward equilibrium.

Shifts in Labor Demand

The demand curve for labor shows the quantity of labor employers wish to hire at any given salary or wage rate, under the *ceteris paribus* assumption. A change in the wage or salary will result in a change in the quantity demanded of labor. If the wage rate increases, employers will want to hire fewer employees. The quantity of labor demanded will decrease, and there will be a movement upward along the demand curve. If the wages and salaries decrease, employers are more likely to hire a greater number of workers. The quantity of labor demanded will increase, resulting in a downward movement along the demand curve.

Shifts in the demand curve for labor occur for many reasons. One key reason is that the demand for labor is based on the demand for the good or service that is produced. For example, the more new automobiles consumers demand, the greater the number of workers automakers will need to hire. Therefore the demand for labor is called a "derived demand." Here are some examples of derived demand for labor:

- The demand for chefs is dependent on the demand for restaurant meals.
- The demand for pharmacists is dependent on the demand for prescription drugs.
- The demand for attorneys is dependent on the demand for legal services.

As the demand for the goods and services increases, the demand for labor will increase, or shift to the right, to meet employers' production requirements. As the demand for the goods and services decreases, the demand for labor will decrease, or shift to the left. **Table 4.2** shows that in addition to the derived demand for labor, demand can also increase or decrease (shift) in response to several factors.

Factors	Results
Demand for Output	When the demand for the good produced (output) increases, both the output price and profitability increase. As a result, producers demand more labor to ramp up production.

Table 4.2 Factors That Can Shift Demand

Factors	Results
Education and Training	A well-trained and educated workforce causes an increase in the demand for that labor by employers. Increased levels of productivity within the workforce will cause the demand for labor to shift to the right. If the workforce is not well-trained or educated, employers will not hire from within that labor pool, since they will need to spend a significant amount of time and money training that workforce. Demand for such will shift to the left.
Technology	Technology changes can act as either substitutes for or complements to labor. When technology acts as a substitute, it replaces the need for the number of workers an employer needs to hire. For example, word processing decreased the number of typists needed in the workplace. This shifted the demand curve for typists left. An increase in the availability of certain technologies may increase the demand for labor. Technology that acts as a complement to labor will increase the demand for certain types of labor, resulting in a rightward shift of the demand curve. For example, the increased use of word processing and other software has increased the demand for information technology professionals who can resolve software and hardware issues related to a firm's network. More and better technology will increase demand for skilled workers who know how to use technology to enhance workplace productivity. Those workers who do not adapt to changes in technology will experience a decrease in demand.
Number of Companies	An increase in the number of companies producing a given product will increase the demand for labor resulting in a shift to the right. A decrease in the number of companies producing a given product will decrease the demand for labor resulting in a shift to the left.
Government Regulations	Complying with government regulations can increase or decrease the demand for labor at any given wage. In the healthcare industry, government rules may require that nurses be hired to carry out certain medical procedures. This will increase the demand for nurses. Less-trained healthcare workers would be prohibited from carrying out these procedures, and the demand for these workers will shift to the left.
Price and Availability of Other Inputs	Labor is not the only input into the production process. For example, a salesperson at a call center needs a telephone and a computer terminal to enter data and record sales. If prices of other inputs fall, production will become more profitable and suppliers will demand more labor to increase production. This will cause a rightward shift in the demand curve for labor. The opposite is also true. Higher prices for other inputs lower demand for labor.

Table 4.2 Factors That Can Shift Demand

Link It Up

Click here (http://openstaxcollege.org/l/Futurework) to read more about "Trends and Challenges for Work in the 21st Century."

Shifts in Labor Supply

The supply of labor is upward-sloping and adheres to the law of supply: The higher the price, the greater the quantity supplied and the lower the price, the less quantity supplied. The supply curve models the tradeoff between supplying labor into the market or using time in leisure activities at every given price level. The higher the wage, the more labor is willing to work and forego leisure activities. **Table 4.3** lists some of the factors that will cause the supply to increase or decrease.

Factors	Results
Number of Workers	An increased number of workers will cause the supply curve to shift to the right. An increased number of workers can be due to several factors, such as immigration, increasing population, an aging population, and changing demographics. Policies that encourage immigration will increase the supply of labor, and vice versa. Population grows when birth rates exceed death rates. This eventually increases supply of labor when the former reach working age. An aging and therefore retiring population will decrease the supply of labor. Another example of changing demographics is more women working outside of the home, which increases the supply of labor.
Required Education	The more required education, the lower the supply. There is a lower supply of PhD mathematicians than of high school mathematics teachers; there is a lower supply of cardiologists than of primary care physicians; and there is a lower supply of physicians than of nurses.
Government Policies	Government policies can also affect the supply of labor for jobs. Alternatively, the government may support rules that set high qualifications for certain jobs: academic training, certificates or licenses, or experience. When these qualifications are made tougher, the number of qualified workers will decrease at any given wage. On the other hand, the government may also subsidize training or even reduce the required level of qualifications. For example, government might offer subsidies for nursing schools or nursing students. Such provisions would shift the supply curve of nurses to the right. In addition, government policies that change the relative desirability of working versus not working also affect the labor supply. These include unemployment benefits, maternity leave, child care benefits, and welfare policy. For example, child care benefits may increase the labor supply of working mothers. Long term unemployment benefits may discourage job searching for unemployed workers. All these policies must therefore be carefully designed to minimize any negative labor supply effects.

Table 4.3 Factors that Can Shift Supply

A change in salary will lead to a movement along labor demand or labor supply curves, but it will not shift those curves. However, other events like those we have outlined here will cause either the demand or the supply of labor to shift, and thus will move the labor market to a new equilibrium salary and quantity.

Technology and Wage Inequality: The Four-Step Process

Economic events can change the equilibrium salary (or wage) and quantity of labor. Consider how the wave of new information technologies, like computer and telecommunications networks, has affected low-skill and high-skill workers in the U.S. economy. From the perspective of employers who demand labor, these new technologies are often a substitute for low-skill laborers like file clerks who used to keep file cabinets full of paper records of transactions. However, the same new technologies are a complement to high-skill workers like managers, who benefit from the technological advances by having the ability to monitor more information, communicate more easily, and juggle a wider array of responsibilities. How will the new technologies affect the wages of high-skill and low-skill workers? For this question, the four-step process of analyzing how shifts in supply or demand affect a market (introduced in **Demand and Supply**) works in this way:

Step 1. What did the markets for low-skill labor and high-skill labor look like before the arrival of the new technologies? In **Figure 4.3** (a) and **Figure 4.3** (b), S_0 is the original supply curve for labor and D_0 is the original demand curve for labor in each market. In each graph, the original point of equilibrium, E_0, occurs at the price W_0 and the quantity Q_0.

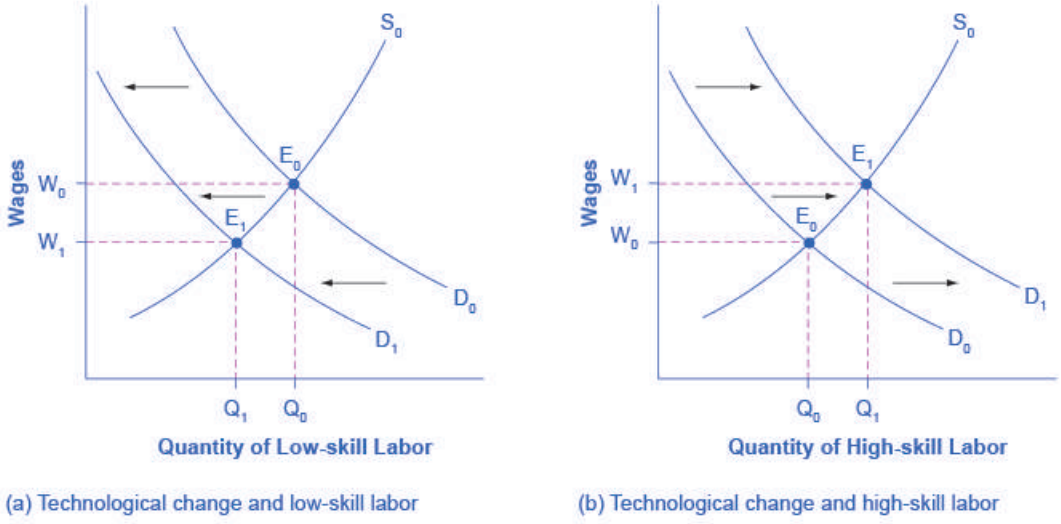


Figure 4.3 Technology and Wages: Applying Demand and Supply (a) The demand for low-skill labor shifts to the left when technology can do the job previously done by these workers. (b) New technologies can also increase the demand for high-skill labor in fields such as information technology and network administration.

Step 2. Does the new technology affect the supply of labor from households or the demand for labor from firms? The technology change described here affects demand for labor by firms that hire workers.

Step 3. Will the new technology increase or decrease demand? Based on the description earlier, as the substitute for low-skill labor becomes available, demand for low-skill labor will shift to the left, from D_0 to D_1. As the technology complement for high-skill labor becomes cheaper, demand for high-skill labor will shift to the right, from D_0 to D_1.

Step 4. The new equilibrium for low-skill labor, shown as point E_1 with price W_1 and quantity Q_1, has a lower wage and quantity hired than the original equilibrium, E_0. The new equilibrium for high-skill labor, shown as point E_1 with price W_1 and quantity Q_1, has a higher wage and quantity hired than the original equilibrium (E_0).

Thus, the demand and supply model predicts that the new computer and communications technologies will raise the pay of high-skill workers but reduce the pay of low-skill workers. From the 1970s to the mid-2000s, the wage gap widened between high-skill and low-skill labor. According to the National Center for Education Statistics, in 1980, for example, a college graduate earned about 30% more than a high school graduate with comparable job experience, but by 2014, a college graduate earned about 66% more than an otherwise comparable high school graduate. Many economists believe that the trend toward greater wage inequality across the U.S. economy is due to improvements in

technology.

Link It Up

Visit this website (http://openstaxcollege.org/l/oldtechjobs) to read about ten tech skills that have lost relevance in today's workforce.

Price Floors in the Labor Market: Living Wages and Minimum Wages

In contrast to goods and services markets, price ceilings are rare in labor markets, because rules that prevent people from earning income are not politically popular. There is one exception: boards of trustees or stockholders, as an example, propose limits on the high incomes of top business executives.

The labor market, however, presents some prominent examples of price floors, which are an attempt to increase the wages of low-paid workers. The U.S. government sets a **minimum wage**, a price floor that makes it illegal for an employer to pay employees less than a certain hourly rate. In mid-2009, the U.S. minimum wage was raised to $7.25 per hour. Local political movements in a number of U.S. cities have pushed for a higher minimum wage, which they call a **living wage**. Promoters of living wage laws maintain that the minimum wage is too low to ensure a reasonable standard of living. They base this conclusion on the calculation that, if you work 40 hours a week at a minimum wage of $7.25 per hour for 50 weeks a year, your annual income is $14,500, which is less than the official U.S. government definition of what it means for a family to be in poverty. (A family with two adults earning minimum wage and two young children will find it more cost efficient for one parent to provide childcare while the other works for income. Thus the family income would be $14,500, which is significantly lower than the federal poverty line for a family of four, which was $24,250 in 2015.)

Supporters of the living wage argue that full-time workers should be assured a high enough wage so that they can afford the essentials of life: food, clothing, shelter, and healthcare. Since Baltimore passed the first living wage law in 1994, several dozen cities enacted similar laws in the late 1990s and the 2000s. The living wage ordinances do not apply to all employers, but they have specified that all employees of the city or employees of firms that the city hires be paid at least a certain wage that is usually a few dollars per hour above the U.S. minimum wage.

Figure 4.4 illustrates the situation of a city considering a living wage law. For simplicity, we assume that there is no federal minimum wage. The wage appears on the vertical axis, because the wage is the price in the labor market. Before the passage of the living wage law, the equilibrium wage is $10 per hour and the city hires 1,200 workers at this wage. However, a group of concerned citizens persuades the city council to enact a living wage law requiring employers to pay no less than $12 per hour. In response to the higher wage, 1,600 workers look for jobs with the city. At this higher wage, the city, as an employer, is willing to hire only 700 workers. At the price floor, the quantity supplied exceeds the quantity demanded, and a surplus of labor exists in this market. For workers who continue to have a job at a higher salary, life has improved. For those who were willing to work at the old wage rate but lost their jobs with the wage increase, life has not improved. **Table 4.4** shows the differences in supply and demand at different wages.

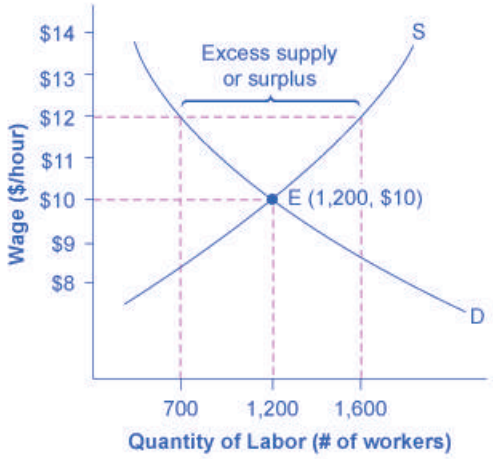


Figure 4.4 A Living Wage: Example of a Price Floor The original equilibrium in this labor market is a wage of $10/hour and a quantity of 1,200 workers, shown at point E. Imposing a wage floor at $12/hour leads to an excess supply of labor. At that wage, the quantity of labor supplied is 1,600 and the quantity of labor demanded is only 700.

Wage	Quantity Labor Demanded	Quantity Labor Supplied
$8/hr	1,900	500
$9/hr	1,500	900
$10/hr	1,200	1,200
$11/hr	900	1,400
$12/hr	700	1,600
$13/hr	500	1,800
$14/hr	400	1,900

Table 4.4 Living Wage: Example of a Price Floor

The Minimum Wage as an Example of a Price Floor

The U.S. minimum wage is a price floor that is set either very close to the equilibrium wage or even slightly below it. About 1% of American workers are actually paid the minimum wage. In other words, the vast majority of the U.S. labor force has its wages determined in the labor market, not as a result of the government price floor. However, for workers with low skills and little experience, like those without a high school diploma or teenagers, the minimum wage is quite important. In many cities, the federal minimum wage is apparently below the market price for unskilled labor, because employers offer more than the minimum wage to checkout clerks and other low-skill workers without any government prodding.

Economists have attempted to estimate how much the minimum wage reduces the quantity demanded of low-skill labor. A typical result of such studies is that a 10% increase in the minimum wage would decrease the hiring of unskilled workers by 1 to 2%, which seems a relatively small reduction. In fact, some studies have even found no effect of a higher minimum wage on employment at certain times and places—although these studies are controversial.

Let's suppose that the minimum wage lies just slightly *below* the equilibrium wage level. Wages could fluctuate according to market forces above this price floor, but they would not be allowed to move beneath the floor. In this situation, the price floor minimum wage is *nonbinding* —that is, the price floor is not determining the market outcome. Even if the minimum wage moves just a little higher, it will still have no effect on the quantity of

employment in the economy, as long as it remains below the equilibrium wage. Even if the government increases minimum wage by enough so that it rises slightly above the equilibrium wage and becomes binding, there will be only a small excess supply gap between the quantity demanded and quantity supplied.

These insights help to explain why U.S. minimum wage laws have historically had only a small impact on employment. Since the minimum wage has typically been set close to the equilibrium wage for low-skill labor and sometimes even below it, it has not had a large effect in creating an excess supply of labor. However, if the minimum wage increased dramatically—say, if it doubled to match the living wages that some U.S. cities have considered—then its impact on reducing the quantity demanded of employment would be far greater. As of 2017, many U.S. states are set to increase their minimum wage to $15 per hour. We will see what happens. The following Clear It Up feature describes in greater detail some of the arguments for and against changes to minimum wage.

Clear It Up

What's the harm in raising the minimum wage?

Because of the law of demand, a higher required wage will reduce the amount of low-skill employment either in terms of employees or in terms of work hours. Although there is controversy over the numbers, let's say for the sake of the argument that a 10% rise in the minimum wage will reduce the employment of low-skill workers by 2%. Does this outcome mean that raising the minimum wage by 10% is bad public policy? Not necessarily.

If 98% of those receiving the minimum wage have a pay increase of 10%, but 2% of those receiving the minimum wage lose their jobs, are the gains for society as a whole greater than the losses? The answer is not clear, because job losses, even for a small group, may cause more pain than modest income gains for others. For one thing, we need to consider which minimum wage workers are losing their jobs. If the 2% of minimum wage workers who lose their jobs are struggling to support families, that is one thing. If those who lose their job are high school students picking up spending money over summer vacation, that is something else.

Another complexity is that many minimum wage workers do not work full-time for an entire year. Imagine a minimum wage worker who holds different part-time jobs for a few months at a time, with bouts of unemployment in between. The worker in this situation receives the 10% raise in the minimum wage when working, but also ends up working 2% fewer hours during the year because the higher minimum wage reduces how much employers want people to work. Overall, this worker's income would rise because the 10% pay raise would more than offset the 2% fewer hours worked.

Of course, these arguments do not prove that raising the minimum wage is necessarily a good idea either. There may well be other, better public policy options for helping low-wage workers. (The **Poverty and Economic Inequality (http://cnx.org/content/m63897/latest/)** chapter discusses some possibilities.) The lesson from this maze of minimum wage arguments is that complex social problems rarely have simple answers. Even those who agree on how a proposed economic policy affects quantity demanded and quantity supplied may still disagree on whether the policy is a good idea.

4.2 | Demand and Supply in Financial Markets

By the end of this section, you will be able to:

- Identify the demanders and suppliers in a financial market
- Explain how interest rates can affect supply and demand
- Analyze the economic effects of U.S. debt in terms of domestic financial markets
- Explain the role of price ceilings and usury laws in the U.S.

United States' households, institutions, and domestic businesses saved almost $1.3 trillion in 2015. Where did that savings go and how was it used? Some of the savings ended up in banks, which in turn loaned the money to individuals or businesses that wanted to borrow money. Some was invested in private companies or loaned to

government agencies that wanted to borrow money to raise funds for purposes like building roads or mass transit. Some firms reinvested their savings in their own businesses.

In this section, we will determine how the demand and supply model links those who wish to supply **financial capital** (i.e., savings) with those who demand financial capital (i.e., borrowing). Those who save money (or make financial investments, which is the same thing), whether individuals or businesses, are on the supply side of the financial market. Those who borrow money are on the demand side of the financial market. For a more detailed treatment of the different kinds of financial investments like bank accounts, stocks and bonds, see the **Financial Markets (http://cnx.org/content/m63860/latest/)** chapter.

Who Demands and Who Supplies in Financial Markets?

In any market, the price is what suppliers receive and what demanders pay. In financial markets, those who supply financial capital through saving expect to receive a rate of return, while those who demand financial capital by receiving funds expect to pay a rate of return. This rate of return can come in a variety of forms, depending on the type of investment.

The simplest example of a rate of return is the **interest rate**. For example, when you supply money into a savings account at a bank, you receive interest on your deposit. The interest the bank pays you as a percent of your deposits is the interest rate. Similarly, if you demand a loan to buy a car or a computer, you will need to pay interest on the money you borrow.

Let’s consider the market for borrowing money with credit cards. In 2015, almost 200 million Americans were cardholders. Credit cards allow you to borrow money from the card's issuer, and pay back the borrowed amount plus interest, although most allow you a period of time in which you can repay the loan without paying interest. A typical credit card interest rate ranges from 12% to 18% per year. In May 2016, Americans had about $943 billion outstanding in credit card debts. About half of U.S. families with credit cards report that they almost always pay the full balance on time, but one-quarter of U.S. families with credit cards say that they “hardly ever” pay off the card in full. In fact, in 2014, 56% of consumers carried an unpaid balance in the last 12 months. Let’s say that, on average, the annual interest rate for credit card borrowing is 15% per year. Thus, Americans pay tens of billions of dollars every year in interest on their credit cards—plus basic fees for the credit card or fees for late payments.

Figure 4.5 illustrates demand and supply in the financial market for credit cards. The horizontal axis of the financial market shows the quantity of money loaned or borrowed in this market. The vertical or price axis shows the rate of return, which in the case of credit card borrowing we can measure with an interest rate. **Table 4.5** shows the quantity of financial capital that consumers demand at various interest rates and the quantity that credit card firms (often banks) are willing to supply.

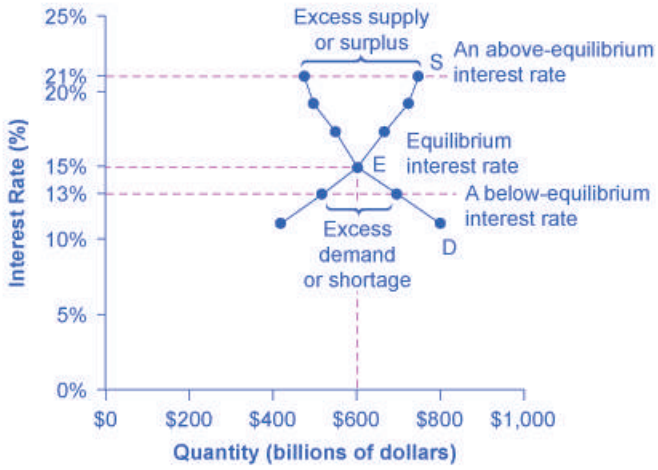


Figure 4.5 Demand and Supply for Borrowing Money with Credit Cards In this market for credit card borrowing, the demand curve (D) for borrowing financial capital intersects the supply curve (S) for lending financial capital at equilibrium E. At the equilibrium, the interest rate (the "price" in this market) is 15% and the quantity of financial capital loaned and borrowed is $600 billion. The equilibrium price is where the quantity demanded and the quantity supplied are equal. At an above-equilibrium interest rate like 21%, the quantity of financial capital supplied would increase to $750 billion, but the quantity demanded would decrease to $480 billion. At a below-equilibrium interest rate like 13%, the quantity of financial capital demanded would increase to $700 billion, but the quantity of financial capital supplied would decrease to $510 billion.

Interest Rate (%)	Quantity of Financial Capital Demanded (Borrowing) ($ billions)	Quantity of Financial Capital Supplied (Lending) ($ billions)
11	$800	$420
13	$700	$510
15	$600	$600
17	$550	$660
19	$500	$720
21	$480	$750

Table 4.5 Demand and Supply for Borrowing Money with Credit Cards

The laws of demand and supply continue to apply in the financial markets. According to the **law of demand**, a higher rate of return (that is, a higher price) will decrease the quantity demanded. As the interest rate rises, consumers will reduce the quantity that they borrow. According to the law of supply, a higher price increases the quantity supplied. Consequently, as the interest rate paid on credit card borrowing rises, more firms will be eager to issue credit cards and to encourage customers to use them. Conversely, if the interest rate on credit cards falls, the quantity of financial capital supplied in the credit card market will decrease and the quantity demanded will fall.

Equilibrium in Financial Markets

In the financial market for credit cards in **Figure 4.5**, the supply curve (S) and the demand curve (D) cross at the equilibrium point (E). The equilibrium occurs at an interest rate of 15%, where the quantity of funds demanded and the quantity supplied are equal at an equilibrium quantity of $600 billion.

If the interest rate (remember, this measures the "price" in the financial market) is above the equilibrium level, then an excess supply, or a surplus, of financial capital will arise in this market. For example, at an interest rate of 21%, the quantity of funds supplied increases to $750 billion, while the quantity demanded decreases to $480 billion. At this

above-equilibrium interest rate, firms are eager to supply loans to credit card borrowers, but relatively few people or businesses wish to borrow. As a result, some credit card firms will lower the interest rates (or other fees) they charge to attract more business. This strategy will push the interest rate down toward the equilibrium level.

If the interest rate is below the equilibrium, then excess demand or a shortage of funds occurs in this market. At an interest rate of 13%, the quantity of funds credit card borrowers demand increases to $700 billion, but the quantity credit card firms are willing to supply is only $510 billion. In this situation, credit card firms will perceive that they are overloaded with eager borrowers and conclude that they have an opportunity to raise interest rates or fees. The interest rate will face economic pressures to creep up toward the equilibrium level.

The FRED database publishes some two dozen measures of interest rates, including interest rates on credit cards, automobile loans, personal loans, mortgage loans, and more. You can find these at the FRED **website (https://openstax.org/l/FRED_stlouis)** .

Shifts in Demand and Supply in Financial Markets

Those who supply financial capital face two broad decisions: how much to save, and how to divide up their savings among different forms of financial investments. We will discuss each of these in turn.

Participants in financial markets must decide when they prefer to consume goods: now or in the future. Economists call this **intertemporal decision making** because it involves decisions across time. Unlike a decision about what to buy from the grocery store, people make investment or savings decisions across a period of time, sometimes a long period.

Most workers save for retirement because their income in the present is greater than their needs, while the opposite will be true once they retire. Thus, they save today and supply financial markets. If their income increases, they save more. If their perceived situation in the future changes, they change the amount of their saving. For example, there is some evidence that Social Security, the program that workers pay into in order to qualify for government checks after retirement, has tended to reduce the quantity of financial capital that workers save. If this is true, Social Security has shifted the supply of financial capital at any interest rate to the left.

By contrast, many college students need money today when their income is low (or nonexistent) to pay their college expenses. As a result, they borrow today and demand from financial markets. Once they graduate and become employed, they will pay back the loans. Individuals borrow money to purchase homes or cars. A business seeks financial investment so that it has the funds to build a factory or invest in a research and development project that will not pay off for five years, ten years, or even more. Thus, when consumers and businesses have greater confidence that they will be able to repay in the future, the quantity demanded of financial capital at any given interest rate will shift to the right.

For example, in the technology boom of the late 1990s, many businesses became extremely confident that investments in new technology would have a high rate of return, and their demand for financial capital shifted to the right. Conversely, during the 2008 and 2009 Great Recession, their demand for financial capital at any given interest rate shifted to the left.

To this point, we have been looking at saving in total. Now let us consider what affects saving in different types of financial investments. In deciding between different forms of financial investments, suppliers of financial capital will have to consider the rates of return and the risks involved. Rate of return is a positive attribute of investments, but risk is a negative. If Investment A becomes more risky, or the return diminishes, then savers will shift their funds to Investment B—and the supply curve of financial capital for Investment A will shift back to the left while the supply curve of capital for Investment B shifts to the right.

The United States as a Global Borrower

In the global economy, trillions of dollars of financial investment cross national borders every year. In the early 2000s, financial investors from foreign countries were investing several hundred billion dollars per year more in the U.S. economy than U.S. financial investors were investing abroad. The following Work It Out deals with one of the macroeconomic concerns for the U.S. economy in recent years.

Work It Out

The Effect of Growing U.S. Debt

Imagine that foreign investors viewed the U.S. economy as a less desirable place to put their money because of fears about the growth of the U.S. public debt. Using the four-step process for analyzing how changes in supply and demand affect equilibrium outcomes, how would increased U.S. public debt affect the equilibrium price and quantity for capital in U.S. financial markets?

Step 1. Draw a diagram showing demand and supply for financial capital that represents the original scenario in which foreign investors are pouring money into the U.S. economy. **Figure 4.6** shows a demand curve, D, and a supply curve, S, where the supply of capital includes the funds arriving from foreign investors. The original equilibrium E_0 occurs at interest rate R_0 and quantity of financial investment Q_0.

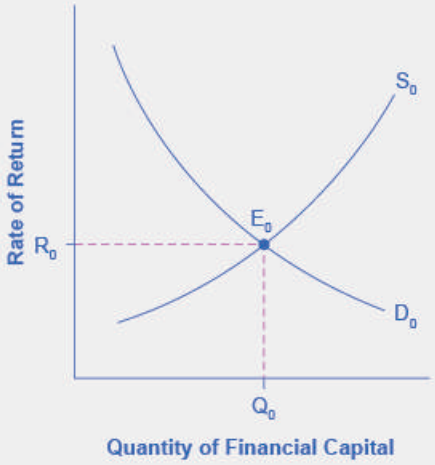


Figure 4.6 The United States as a Global Borrower Before U.S. Debt Uncertainty The graph shows the demand for financial capital from and supply of financial capital into the U.S. financial markets by the foreign sector before the increase in uncertainty regarding U.S. public debt. The original equilibrium (E_0) occurs at an equilibrium rate of return (R_0) and the equilibrium quantity is at Q_0.

Step 2. Will the diminished confidence in the U.S. economy as a place to invest affect demand or supply of financial capital? Yes, it will affect supply. Many foreign investors look to the U.S. financial markets to store their money in safe financial vehicles with low risk and stable returns. Diminished confidence means U.S. financial assets will be seen as more risky.

Step 3. Will supply increase or decrease? When the enthusiasm of foreign investors' for investing their money in the U.S. economy diminishes, the supply of financial capital shifts to the left. **Figure 4.7** shows the supply curve shift from S_0 to S_1.

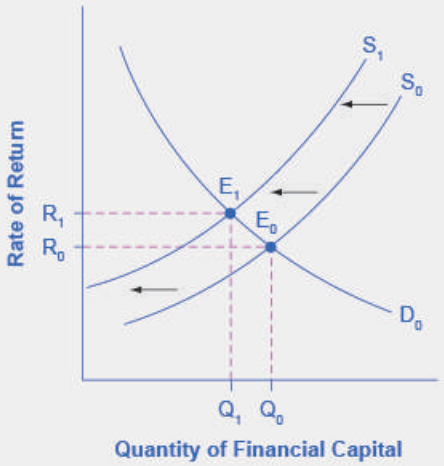


Figure 4.7 The United States as a Global Borrower Before and After U.S. Debt Uncertainty The graph shows the demand for financial capital and supply of financial capital into the U.S. financial markets by the foreign sector before and after the increase in uncertainty regarding U.S. public debt. The original equilibrium (E_0) occurs at an equilibrium rate of return (R_0) and the equilibrium quantity is at Q_0.

Step 4. Thus, foreign investors' diminished enthusiasm leads to a new equilibrium, E_1, which occurs at the higher interest rate, R_1, and the lower quantity of financial investment, Q_1. In short, U.S. borrowers will have to pay more interest on their borrowing.

The economy has experienced an enormous inflow of foreign capital. According to the U.S. Bureau of Economic Analysis, by the third quarter of 2015, U.S. investors had accumulated $23.3 trillion of foreign assets, but foreign investors owned a total of $30.6 trillion of U.S. assets. If foreign investors were to pull their money out of the U.S. economy and invest elsewhere in the world, the result could be a significantly lower quantity of financial investment in the United States, available only at a higher interest rate. This reduced inflow of foreign financial investment could impose hardship on U.S. consumers and firms interested in borrowing.

In a modern, developed economy, financial capital often moves invisibly through electronic transfers between one bank account and another. Yet we can analyze these flows of funds with the same tools of demand and supply as markets for goods or labor.

Price Ceilings in Financial Markets: Usury Laws

As we noted earlier, about 200 million Americans own credit cards, and their interest payments and fees total tens of billions of dollars each year. It is little wonder that political pressures sometimes arise for setting limits on the interest rates or fees that credit card companies charge. The firms that issue credit cards, including banks, oil companies, phone companies, and retail stores, respond that the higher interest rates are necessary to cover the losses created by those who borrow on their credit cards and who do not repay on time or at all. These companies also point out that cardholders can avoid paying interest if they pay their bills on time.

Consider the credit card market as **Figure 4.8** illustrators. In this financial market, the vertical axis shows the interest rate (which is the price in the financial market). Demanders in the credit card market are households and businesses. Suppliers are the companies that issue credit cards. This figure does not use specific numbers, which would be hypothetical in any case, but instead focuses on the underlying economic relationships. Imagine a law imposes a price ceiling that holds the interest rate charged on credit cards at the rate Rc, which lies below the interest rate R_0 that would otherwise have prevailed in the market. The horizontal dashed line at interest rate R_c in **Figure 4.8** shows the price ceiling. The demand and supply model predicts that at the lower price ceiling interest rate, the quantity demanded of credit card debt will increase from its original level of Q_0 to Qd; however, the quantity supplied of credit card debt will decrease from the original Q_0 to Qs. At the price ceiling (Rc), quantity demanded will exceed quantity supplied. Consequently, a number of people who want to have credit cards and are willing to pay the prevailing interest rate will find that companies are unwilling to issue cards to them. The result will be a credit shortage.

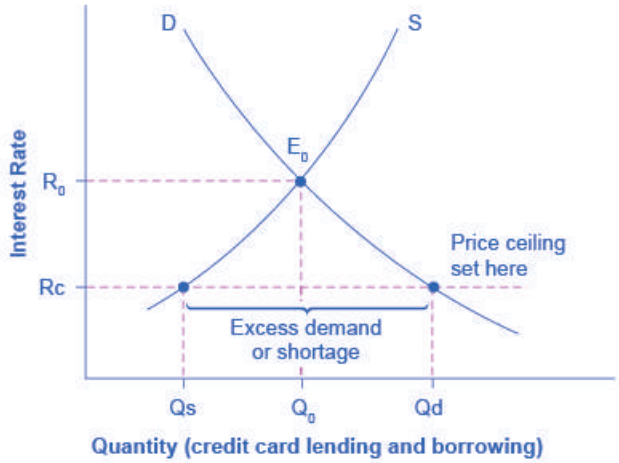


Figure 4.8 Credit Card Interest Rates: Another Price Ceiling Example The original intersection of demand D and supply S occurs at equilibrium E_0. However, a price ceiling is set at the interest rate Rc, below the equilibrium interest rate R_0, and so the interest rate cannot adjust upward to the equilibrium. At the price ceiling, the quantity demanded, Qd, exceeds the quantity supplied, Qs. There is excess demand, also called a shortage.

Many states do have **usury laws**, which impose an upper limit on the interest rate that lenders can charge. However, in many cases these upper limits are well above the market interest rate. For example, if the interest rate is not allowed to rise above 30% per year, it can still fluctuate below that level according to market forces. A price ceiling that is set at a relatively high level is nonbinding, and it will have no practical effect unless the equilibrium price soars high enough to exceed the price ceiling.

4.3 | The Market System as an Efficient Mechanism for Information

By the end of this section, you will be able to:

- Apply demand and supply models to analyze prices and quantities
- Explain the effects of price controls on the equilibrium of prices and quantities

Prices exist in markets for goods and services, for labor, and for financial capital. In all of these markets, prices serve as a remarkable social mechanism for collecting, combining, and transmitting information that is relevant to the market—namely, the relationship between demand and supply—and then serving as messengers to convey that information to buyers and sellers. In a market-oriented economy, no government agency or guiding intelligence oversees the set of responses and interconnections that result from a change in price. Instead, each consumer reacts according to that person's preferences and budget set, and each profit-seeking producer reacts to the impact on its expected profits. The following Clear It Up feature examines the **demand and supply models**.

Clear It Up

Why are demand and supply curves important?

The demand and supply model is the second fundamental diagram for this course. (The opportunity set model that we introduced in the **Choice in a World of Scarcity** chapter was the first.) Just as it would be foolish to try to learn the arithmetic of long division by memorizing every possible combination of numbers that can be divided by each other, it would be foolish to try to memorize every specific example of demand and supply in this chapter, this textbook, or this course. Demand and supply is not primarily a list of examples. It is a

model to analyze prices and quantities. Even though demand and supply diagrams have many labels, they are fundamentally the same in their logic. Your goal should be to understand the underlying model so you can use it to analyze *any* market.

Figure 4.9 displays a generic demand and supply curve. The horizontal axis shows the different measures of quantity: a quantity of a good or service, or a quantity of labor for a given job, or a quantity of financial capital. The vertical axis shows a measure of price: the price of a good or service, the wage in the labor market, or the rate of return (like the interest rate) in the financial market.

The demand and supply model can explain the existing levels of prices, wages, and rates of return. To carry out such an analysis, think about the quantity that will be demanded at each price and the quantity that will be supplied at each price—that is, think about the shape of the demand and supply curves—and how these forces will combine to produce equilibrium.

We can also use demand and supply to explain how economic events will cause changes in prices, wages, and rates of return. There are only four possibilities: the change in any single event may cause the demand curve to shift right or to shift left, or it may cause the supply curve to shift right or to shift left. The key to analyzing the effect of an economic event on equilibrium prices and quantities is to determine which of these four possibilities occurred. The way to do this correctly is to think back to the list of factors that shift the demand and supply curves. Note that if more than one variable is changing at the same time, the overall impact will depend on the degree of the shifts. When there are multiple variables, economists isolate each change and analyze it independently.

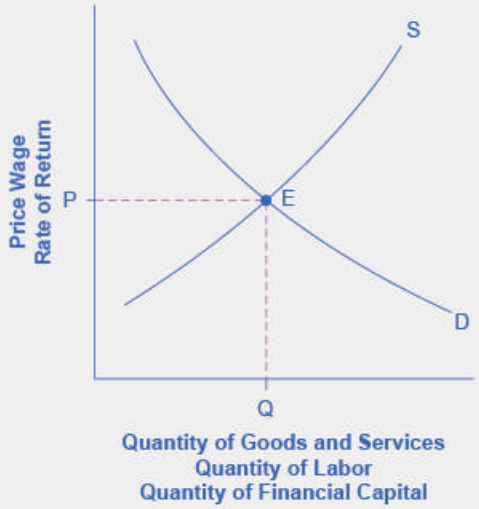


Figure 4.9 Demand and Supply Curves The figure displays a generic demand and supply curve. The horizontal axis shows the different measures of quantity: a quantity of a good or service, a quantity of labor for a given job, or a quantity of financial capital. The vertical axis shows a measure of price: the price of a good or service, the wage in the labor market, or the rate of return (like the interest rate) in the financial market. We can use the demand and supply curves explain how economic events will cause changes in prices, wages, and rates of return.

An increase in the price of some product signals consumers that there is a shortage; therefore, they may want to economize on buying this product. For example, if you are thinking about taking a plane trip to Hawaii, but the ticket turns out to be expensive during the week you intend to go, you might consider other weeks when the ticket might be cheaper. The price could be high because you were planning to travel during a holiday when demand for traveling is high. Maybe the cost of an input like jet fuel increased or the airline has raised the price temporarily to see how many people are willing to pay it. Perhaps all of these factors are present at the same time. You do not need to analyze the market and break down the price change into its underlying factors. You just have to look at the ticket price and decide whether and when to fly.

In the same way, price changes provide useful information to producers. Imagine the situation of a farmer who grows

oats and learns that the price of oats has risen. The higher price could be due to an increase in demand caused by a new scientific study proclaiming that eating oats is especially healthful. Perhaps the price of a substitute grain, like corn, has risen, and people have responded by buying more oats. The oat farmer does not need to know the details. The farmer only needs to know that the price of oats has risen and that it will be profitable to expand production as a result.

The actions of individual consumers and producers as they react to prices overlap and interlock in markets for goods, labor, and financial capital. A change in any single market is transmitted through these multiple interconnections to other markets. The vision of the role of flexible prices helping markets to reach equilibrium and linking different markets together helps to explain why price controls can be so counterproductive. Price controls are government laws that serve to regulate prices rather than allow the various markets to determine prices. There is an old proverb: "Don't kill the messenger." In ancient times, messengers carried information between distant cities and kingdoms. When they brought bad news, there was an emotional impulse to kill the messenger. However, killing the messenger did not kill the bad news. Moreover, killing the messenger had an undesirable side effect: Other messengers would refuse to bring news to that city or kingdom, depriving its citizens of vital information.

Those who seek price controls are trying to kill the messenger—or at least to stifle an unwelcome message that prices are bringing about the equilibrium level of price and quantity. However, price controls do nothing to affect the underlying forces of demand and supply, and this can have serious repercussions. During China's "Great Leap Forward" in the late 1950s, the government kept food prices artificially low, with the result that 30 to 40 million people died of starvation because the low prices depressed farm production. This was communist party leader Mao Zedong's social and economic campaign to rapidly transform the country from an agrarian economy to a socialist society through rapid industrialization and collectivization. Changes in demand and supply will continue to reveal themselves through consumers' and producers' behavior. Immobilizing the price messenger through price controls will deprive everyone in the economy of critical information. Without this information, it becomes difficult for everyone—buyers and sellers alike—to react in a flexible and appropriate manner as changes occur throughout the economy.

Bring it Home

Baby Boomers Come of Age

The theory of supply and demand can explain what happens in the labor markets and suggests that the demand for nurses will increase as healthcare needs of baby boomers increase, as **Figure 4.10** shows. The impact of that increase will result in an average salary higher than the $67,490 earned in 2015 referenced in the first part of this case. The new equilibrium (E_1) will be at the new equilibrium price (Pe_1).Equilibrium quantity will also increase from Qe_0 to Qe_1.

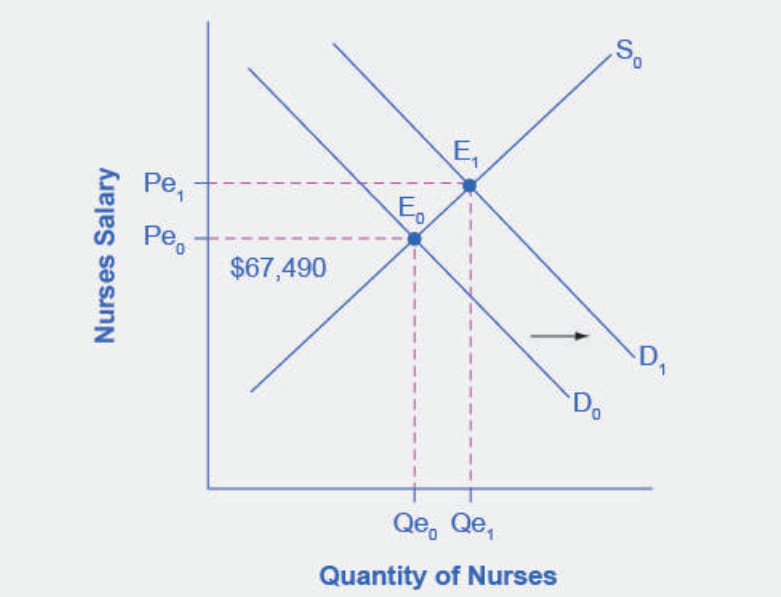


Figure 4.10 Impact of Increasing Demand for Nurses 2014-2024 In 2014, the median salary for nurses was $67,490. As demand for services increases, the demand curve shifts to the right (from D_0 to D_1) and the equilibrium quantity of nurses increases from Qe_0 to Qe_1. The equilibrium salary increases from Pe_0 to Pe_1.

Suppose that as the demand for nurses increases, the supply shrinks due to an increasing number of nurses entering retirement and increases in the tuition of nursing degrees. The leftward shift of the supply curve in **Figure 4.11** captures the impact of a decreasing supply of nurses. The shifts in the two curves result in higher salaries for nurses, but the overall impact in the quantity of nurses is uncertain, as it depends on the relative shifts of supply and demand.

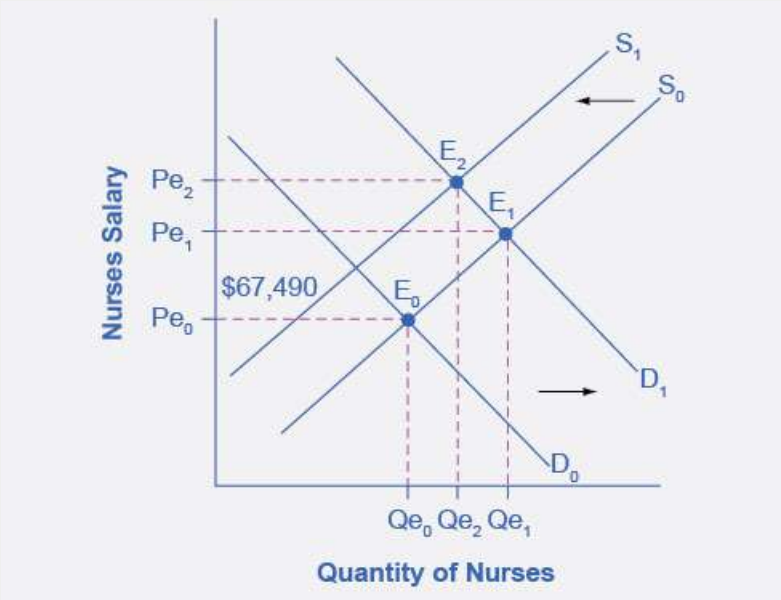


Figure 4.11 Impact of Decreasing Supply of Nurses between 2014 and 2024 The increase in demand for nurses shown in Figure 4.10 leads to both higher prices and higher quantities demanded. As nurses retire from the work force, the supply of nurses decreases, causing a leftward shift in the supply curve and higher salaries for nurses at Pe_2. The net effect on the equilibrium quantity of nurses is uncertain, which in this representation is less than Qe_1, but more than the initial Qe_0.

While we do not know if the number of nurses will increase or decrease relative to their initial employment, we know they will have higher salaries.

KEY TERMS

interest rate the "price" of borrowing in the financial market; a rate of return on an investment

minimum wage a price floor that makes it illegal for an employer to pay employees less than a certain hourly rate

usury laws laws that impose an upper limit on the interest rate that lenders can charge

KEY CONCEPTS AND SUMMARY

4.1 Demand and Supply at Work in Labor Markets

In the labor market, households are on the supply side of the market and firms are on the demand side. In the market for financial capital, households and firms can be on either side of the market: they are suppliers of financial capital when they save or make financial investments, and demanders of financial capital when they borrow or receive financial investments.

In the demand and supply analysis of labor markets, we can measure the price by the annual salary or hourly wage received. We can measure the quantity of labor various ways, like number of workers or the number of hours worked.

Factors that can shift the demand curve for labor include: a change in the quantity demanded of the product that the labor produces; a change in the production process that uses more or less labor; and a change in government policy that affects the quantity of labor that firms wish to hire at a given wage. Demand can also increase or decrease (shift) in response to: workers' level of education and training, technology, the number of companies, and availability and price of other inputs.

The main factors that can shift the supply curve for labor are: how desirable a job appears to workers relative to the alternatives, government policy that either restricts or encourages the quantity of workers trained for the job, the number of workers in the economy, and required education.

4.2 Demand and Supply in Financial Markets

In the demand and supply analysis of financial markets, the "price" is the rate of return or the interest rate received. We measure the quantity by the money that flows from those who supply financial capital to those who demand it.

Two factors can shift the supply of financial capital to a certain investment: if people want to alter their existing levels of consumption, and if the riskiness or return on one investment changes relative to other investments. Factors that can shift demand for capital include business confidence and consumer confidence in the future—since financial investments received in the present are typically repaid in the future.

4.3 The Market System as an Efficient Mechanism for Information

The market price system provides a highly efficient mechanism for disseminating information about relative scarcities of goods, services, labor, and financial capital. Market participants do not need to know why prices have changed, only that the changes require them to revisit previous decisions they made about supply and demand. Price controls hide information about the true scarcity of products and thereby cause misallocation of resources.

SELF-CHECK QUESTIONS

1. In the labor market, what causes a movement along the demand curve? What causes a shift in the demand curve?

2. In the labor market, what causes a movement along the supply curve? What causes a shift in the supply curve?

3. Why is a living wage considered a price floor? Does imposing a living wage have the same outcome as a minimum wage?

4. In the financial market, what causes a movement along the demand curve? What causes a shift in the demand curve?

5. In the financial market, what causes a movement along the supply curve? What causes a shift in the supply curve?

6. If a usury law limits interest rates to no more than 35%, what would the likely impact be on the amount of loans made and interest rates paid?

7. Which of the following changes in the financial market will lead to a decline in interest rates:
a. a rise in demand
b. a fall in demand
c. a rise in supply
d. a fall in supply

8. Which of the following changes in the financial market will lead to an increase in the quantity of loans made and received:
a. a rise in demand
b. a fall in demand
c. a rise in supply
d. a fall in supply

9. Identify the most accurate statement. A price floor will have the largest effect if it is set:
a. substantially above the equilibrium price
b. slightly above the equilibrium price
c. slightly below the equilibrium price
d. substantially below the equilibrium price

Sketch all four of these possibilities on a demand and supply diagram to illustrate your answer.

10. A price ceiling will have the largest effect:
a. substantially below the equilibrium price
b. slightly below the equilibrium price
c. substantially above the equilibrium price
d. slightly above the equilibrium price

Sketch all four of these possibilities on a demand and supply diagram to illustrate your answer.

11. Select the correct answer. A price floor will usually shift:
a. demand
b. supply
c. both
d. neither

Illustrate your answer with a diagram.

12. Select the correct answer. A price ceiling will usually shift:
a. demand
b. supply
c. both
d. neither

REVIEW QUESTIONS

13. What is the “price” commonly called in the labor market?

14. Are households demanders or suppliers in the goods market? Are firms demanders or suppliers in the goods market? What about the labor market and the financial market?

15. Name some factors that can cause a shift in the demand curve in labor markets.

16. Name some factors that can cause a shift in the supply curve in labor markets.

17. How do economists define equilibrium in financial markets?

18. What would be a sign of a shortage in financial markets?

19. Would usury laws help or hinder resolution of a shortage in financial markets?

20. Whether the product market or the labor market, what happens to the equilibrium price and quantity for each of the four possibilities: increase in demand, decrease in demand, increase in supply, and decrease in supply.

CRITICAL THINKING QUESTIONS

21. Other than the demand for labor, what would be another example of a "derived demand?"

22. Suppose that a 5% increase in the minimum wage causes a 5% reduction in employment. How would this affect employers and how would it affect workers? In your opinion, would this be a good policy?

23. Under what circumstances would a minimum wage be a nonbinding price floor? Under what circumstances would a living wage be a binding price floor?

24. Suppose the U.S. economy began to grow more rapidly than other countries in the world. What would be the likely impact on U.S. financial markets as part of the global economy?

25. If the government imposed a federal interest rate ceiling of 20% on all loans, who would gain and who would lose?

26. Why are the factors that shift the demand for a product different from the factors that shift the demand for labor? Why are the factors that shift the supply of a product different from those that shift the supply of labor?

27. During a discussion several years ago on building a pipeline to Alaska to carry natural gas, the U.S. Senate passed a bill stipulating that there should be a guaranteed minimum price for the natural gas that would flow through the pipeline. The thinking behind the bill was that if private firms had a guaranteed price for their natural gas, they would be more willing to drill for gas and to pay to build the pipeline.

a. Using the demand and supply framework, predict the effects of this price floor on the price, quantity demanded, and quantity supplied.
b. With the enactment of this price floor for natural gas, what are some of the likely unintended consequences in the market?
c. Suggest some policies other than the price floor that the government can pursue if it wishes to encourage drilling for natural gas and for a new pipeline in Alaska.

PROBLEMS

28. Identify each of the following as involving either demand or supply. Draw a circular flow diagram and label the flows A through F. (Some choices can be on both sides of the goods market.)

a. Households in the labor market
b. Firms in the goods market
c. Firms in the financial market
d. Households in the goods market
e. Firms in the labor market
f. Households in the financial market

29. Predict how each of the following events will raise or lower the equilibrium wage and quantity of oil workers in Texas. In each case, sketch a demand and supply diagram to illustrate your answer.

a. The price of oil rises.
b. New oil-drilling equipment is invented that is cheap and requires few workers to run.
c. Several major companies that do not drill oil open factories in Texas, offering many well-paid jobs outside the oil industry.
d. Government imposes costly new regulations to make oil-drilling a safer job.

30. Predict how each of the following economic changes will affect the equilibrium price and quantity in the financial market for home loans. Sketch a demand and supply diagram to support your answers.

a. The number of people at the most common ages for home-buying increases.
b. People gain confidence that the economy is growing and that their jobs are secure.
c. Banks that have made home loans find that a larger number of people than they expected are not repaying those loans.
d. Because of a threat of a war, people become uncertain about their economic future.
e. The overall level of saving in the economy diminishes.
f. The federal government changes its bank regulations in a way that makes it cheaper and easier for banks to make home loans.

31. Table 4.6 shows the amount of savings and borrowing in a market for loans to purchase homes, measured in millions of dollars, at various interest rates. What is the equilibrium interest rate and quantity in the capital financial market? How can you tell? Now, imagine that because of a shift in the perceptions of foreign investors, the supply curve shifts so that there will be $10 million less supplied at every interest rate. Calculate the new equilibrium interest rate and quantity, and explain why the direction of the interest rate shift makes intuitive sense.

Interest Rate	Qs	Qd
5%	130	170
6%	135	150
7%	140	140
8%	145	135
9%	150	125
10%	155	110

Table 4.6

32. Imagine that to preserve the traditional way of life in small fishing villages, a government decides to impose a price floor that will guarantee all fishermen a certain price for their catch.

a. Using the demand and supply framework, predict the effects on the price, quantity demanded, and quantity supplied.
b. With the enactment of this price floor for fish, what are some of the likely unintended consequences in the market?
c. Suggest some policies other than the price floor to make it possible for small fishing villages to continue.

33. What happens to the price and the quantity bought and sold in the cocoa market if countries producing cocoa experience a drought and a new study is released demonstrating the health benefits of cocoa? Illustrate your answer with a demand and supply graph.

5 | Elasticity

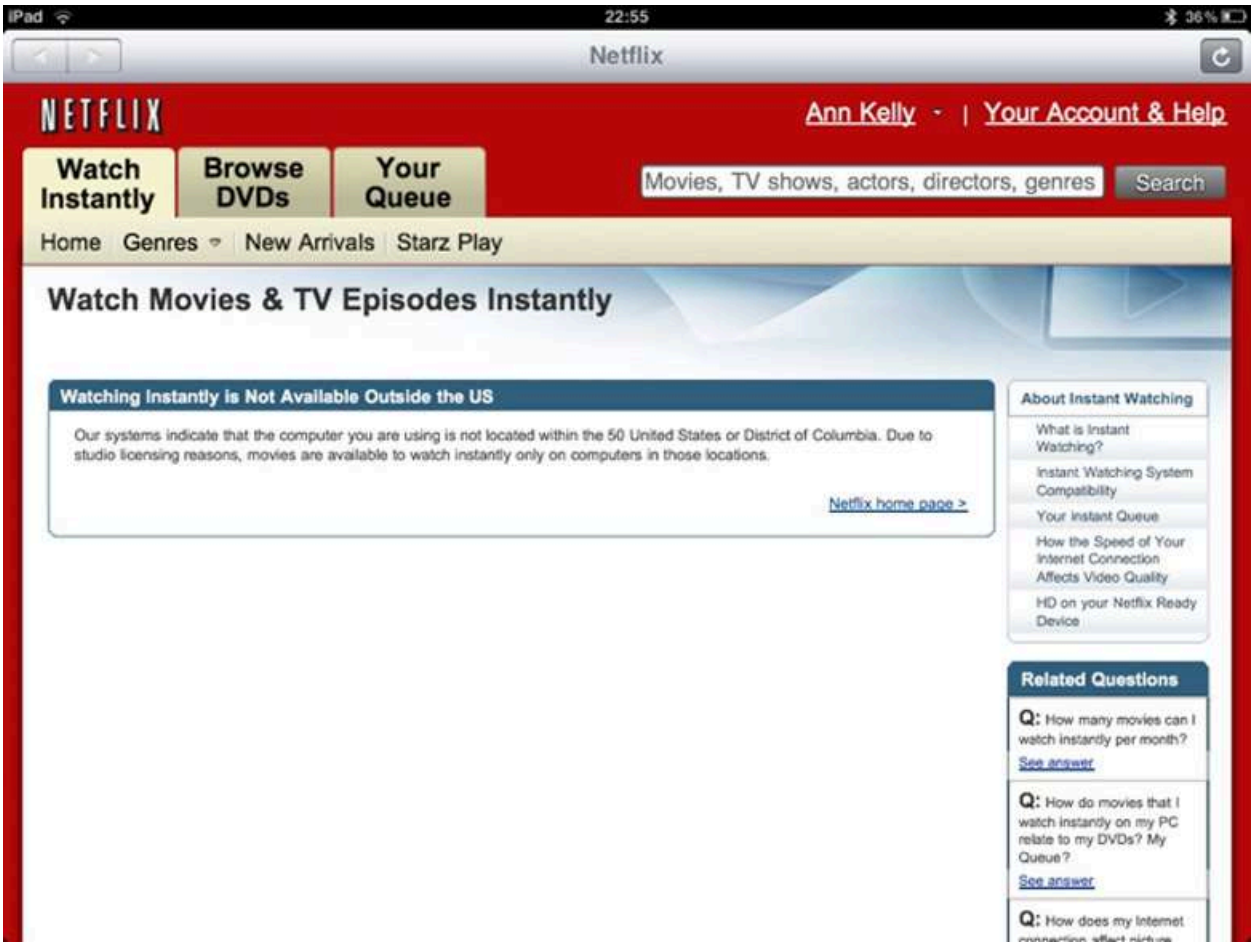


Figure 5.1 Netflix On-Demand Media Netflix, Inc. is an American provider of on-demand Internet streaming media to many countries around the world, including the United States, and of flat rate DVD-by-mail in the United States. (Credit: modification of work by Traci Lawson/Flickr Creative Commons)

Bring it Home

That Will Be How Much?

Imagine going to your favorite coffee shop and having the waiter inform you the pricing has changed. Instead of $3 for a cup of coffee, you will now be charged $2 for coffee, $1 for creamer, and $1 for your choice of sweetener. If you pay your usual $3 for a cup of coffee, you must choose between creamer and sweetener. If you want both, you now face an extra charge of $1. Sound absurd? Well, that is similar to the situation Netflix customers found themselves in—they faced a 60% price hike to retain the same service in 2011.

In early 2011, Netflix consumers paid about $10 a month for a package consisting of streaming video and DVD rentals. In July 2011, the company announced a packaging change. Customers wishing to retain both streaming video and DVD rental would be charged $15.98 per month, a price increase of about 60%. In 2014, Netflix also raised its streaming video subscription price from $7.99 to $8.99 per month for new U.S. customers. The company also changed its policy of 4K streaming content from $9.00 to $12.00 per month that year.

How would customers of the 18-year-old firm react? Would they abandon Netflix? Would the ease of access to other venues make a difference in how consumers responded to the Netflix price change? We will explore the answers to those questions in this chapter, which focuses on the change in quantity with respect to a change in price, a concept economists call elasticity.

Introduction to Elasticity

In this chapter, you will learn about:

- Price Elasticity of Demand and Price Elasticity of Supply
- Polar Cases of Elasticity and Constant Elasticity
- Elasticity and Pricing
- Elasticity in Areas Other Than Price

Anyone who has studied economics knows the law of demand: a higher price will lead to a lower quantity demanded. What you may not know is how much lower the quantity demanded will be. Similarly, the law of supply states that a higher price will lead to a higher quantity supplied. The question is: How much higher? This chapter will explain how to answer these questions and why they are critically important in the real world.

To find answers to these questions, we need to understand the concept of elasticity. **Elasticity** is an economics concept that measures responsiveness of one variable to changes in another variable. Suppose you drop two items from a second-floor balcony. The first item is a tennis ball. The second item is a brick. Which will bounce higher? Obviously, the tennis ball. We would say that the tennis ball has greater elasticity.

Consider an economic example. Cigarette taxes are an example of a "sin tax," a tax on something that is bad for you, like alcohol. Governments tax cigarettes at the state and national levels. State taxes range from a low of 17 cents per pack in Missouri to $4.35 per pack in New York. The average state cigarette tax is $1.69 per pack. The 2014 federal tax rate on cigarettes was $1.01 per pack, but in 2015 the Obama Administration proposed raising the federal tax nearly a dollar to $1.95 per pack. The key question is: How much would cigarette purchases decline?

Taxes on cigarettes serve two purposes: to raise tax revenue for government and to discourage cigarette consumption. However, if a higher cigarette tax discourages consumption considerably, meaning a greatly reduced quantity of cigarette sales, then the cigarette tax on each pack will not raise much revenue for the government. Alternatively, a higher cigarette tax that does not discourage consumption by much will actually raise more tax revenue for the government. Thus, when a government agency tries to calculate the effects of altering its cigarette tax, it must analyze how much the tax affects the quantity of cigarettes consumed. This issue reaches beyond governments and taxes. Every firm faces a similar issue. When a firm considers raising the sales price, it must consider how much a price increase will reduce the quantity demanded of what it sells. Conversely, when a firm puts its products on sale, it must expect (or hope) that the lower price will lead to a significantly higher quantity demanded.

5.1 | Price Elasticity of Demand and Price Elasticity of Supply

By the end of this section, you will be able to:

- Calculate the price elasticity of demand
- Calculate the price elasticity of supply

Both the demand and supply curve show the relationship between price and the number of units demanded or supplied. **Price elasticity** is the ratio between the percentage change in the quantity demanded (Qd) or supplied (Qs) and the corresponding percent change in price. The **price elasticity of demand** is the percentage change in the quantity *demanded* of a good or service divided by the percentage change in the price. The **price elasticity of supply** is the percentage change in quantity *supplied* divided by the percentage change in price.

We can usefully divide elasticities into three broad categories: elastic, inelastic, and unitary. An **elastic demand** or **elastic supply** is one in which the elasticity is greater than one, indicating a high responsiveness to changes in price. Elasticities that are less than one indicate low responsiveness to price changes and correspond to **inelastic demand** or **inelastic supply**. **Unitary elasticities** indicate proportional responsiveness of either demand or supply, as **Table 5.1** summarizes.

If . . .	Then . . .	And It Is Called . . .
% change in quantity > % change in price	$\frac{\%\text{ change in quantity}}{\%\text{ change in price}} > 1$	Elastic
% change in quantity = % change in price	$\frac{\%\text{ change in quantity}}{\%\text{ change in price}} = 1$	Unitary
% change in quantity < % change in price	$\frac{\%\text{ change in quantity}}{\%\text{ change in price}} < 1$	Inelastic

Table 5.1 Elastic, Inelastic, and Unitary: Three Cases of Elasticity

Link It Up

Before we delve into the details of elasticity, enjoy this article (http://openstaxcollege.org/l/Super_Bowl) on elasticity and ticket prices at the Super Bowl.

To calculate elasticity along a demand or supply curve economists use the average percent change in both quantity and price. This is called the Midpoint Method for Elasticity, and is represented in the following equations:

$$\%\text{ change in quantity} = \frac{Q_2 - Q_1}{(Q_2 + Q_1)/2} \times 100$$

$$\%\text{ change in price} = \frac{P_2 - P_1}{(P_2 + P_1)/2} \times 100$$

The advantage of the Midpoint Method is that one obtains the same elasticity between two price points whether there is a price increase or decrease. This is because the formula uses the same base (average quantity and average price) for both cases.

Calculating Price Elasticity of Demand

Let's calculate the elasticity between points A and B and between points G and H as **Figure 5.2** shows.

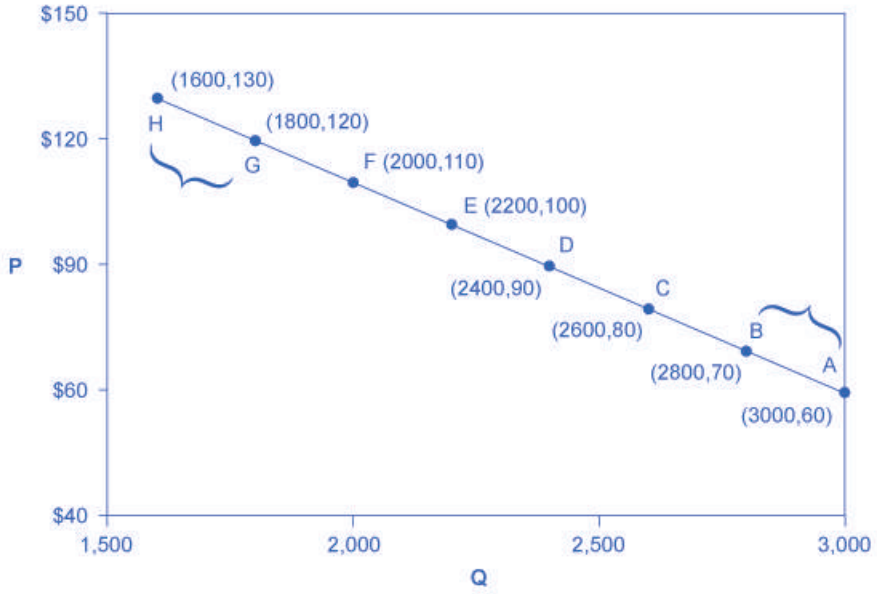


Figure 5.2 Calculating the Price Elasticity of Demand We calculate the price elasticity of demand as the percentage change in quantity divided by the percentage change in price.

First, apply the formula to calculate the elasticity as price decreases from \$70 at point B to \$60 at point A:

$$\begin{aligned}
\%\text{ change in quantity} &= \frac{3{,}000 - 2{,}800}{(3{,}000 + 2{,}800)/2} \times 100 \\
&= \frac{200}{2{,}900} \times 100 \\
&= 6.9 \\
\%\text{ change in price} &= \frac{60 - 70}{(60 + 70)/2} \times 100 \\
&= \frac{-10}{65} \times 100 \\
&= -15.4 \\
\text{Price Elasticity of Demand} &= \frac{6.9\%}{-15.4\%} \\
&= 0.45
\end{aligned}$$

Therefore, the elasticity of demand between these two points is $\frac{6.9\%}{-15.4\%}$ which is 0.45, an amount smaller than one, showing that the demand is inelastic in this interval. Price elasticities of demand are *always* negative since price and quantity demanded always move in opposite directions (on the demand curve). By convention, we always talk about elasticities as positive numbers. Mathematically, we take the absolute value of the result. We will ignore this detail from now on, while remembering to interpret elasticities as positive numbers.

This means that, along the demand curve between point B and A, if the price changes by 1%, the quantity demanded will change by 0.45%. A change in the price will result in a smaller percentage change in the quantity demanded. For example, a 10% *increase* in the price will result in only a 4.5% *decrease* in quantity demanded. A 10% *decrease* in the price will result in only a 4.5% *increase* in the quantity demanded. Price elasticities of demand are negative numbers indicating that the demand curve is downward sloping, but we read them as absolute values. The following Work It Out feature will walk you through calculating the price elasticity of demand.

Work It Out

Finding the Price Elasticity of Demand

Calculate the price elasticity of demand using the data in **Figure 5.2** for an increase in price from G to H. Has the elasticity increased or decreased?

Step 1. We know that:

$$\text{Price Elasticity of Demand} = \frac{\%\text{ change in quantity}}{\%\text{ change in price}}$$

Step 2. From the Midpoint Formula we know that:

$$\%\text{ change in quantity} = \frac{Q_2 - Q_1}{(Q_2 + Q_1)/2} \times 100$$

$$\%\text{ change in price} = \frac{P_2 - P_1}{(P_2 + P_1)/2} \times 100$$

Step 3. So we can use the values provided in the figure in each equation:

$$\begin{aligned} \%\text{ change in quantity} &= \frac{1{,}600 - 1{,}800}{(1{,}600 + 1{,}800)/2} \times 100 \\ &= \frac{-200}{1{,}700} \times 100 \\ &= -11.76 \\ \%\text{ change in price} &= \frac{130 - 120}{(130 + 120)/2} \times 100 \\ &= \frac{10}{125} \times 100 \\ &= 8.0 \end{aligned}$$

Step 4. Then, we can use those values to determine the price elasticity of demand:

$$\begin{aligned} \text{Price Elasticity of Demand} &= \frac{\%\text{ change in quantity}}{\%\text{ change in price}} \\ &= \frac{-11.76}{8} \\ &= 1.47 \end{aligned}$$

Therefore, the elasticity of demand from G to is H 1.47. The magnitude of the elasticity has increased (in absolute value) as we moved up along the demand curve from points A to B. Recall that the elasticity between these two points was 0.45. Demand was inelastic between points A and B and elastic between points G and H. This shows us that price elasticity of demand changes at different points along a straight-line demand curve.

Calculating the Price Elasticity of Supply

Assume that an apartment rents for $650 per month and at that price the landlord rents 10,000 units are rented as **Figure 5.3** shows. When the price increases to $700 per month, the landlord supplies 13,000 units into the market. By what percentage does apartment supply increase? What is the price sensitivity?

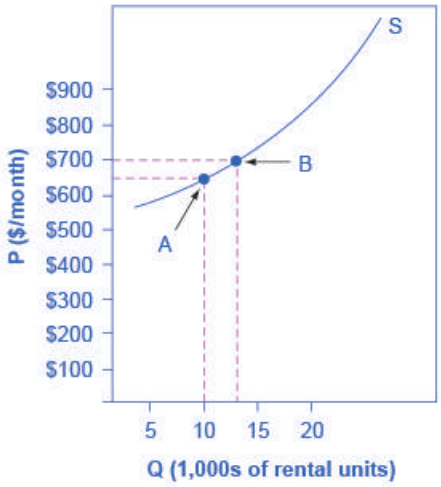


Figure 5.3 Price Elasticity of Supply We calculate the price elasticity of supply as the percentage change in quantity divided by the percentage change in price.

Using the Midpoint Method,

$$\begin{aligned} \%\text{ change in quantity} &= \frac{13{,}000 - 10{,}000}{(13{,}000 + 10{,}000)/2} \times 100 \\ &= \frac{3{,}000}{11{,}500} \times 100 \\ &= 26.1 \\ \%\text{ change in price} &= \frac{\$700 - \$650}{(\$700 + \$650)/2} \times 100 \\ &= \frac{50}{675} \times 100 \\ &= 7.4 \\ \text{Price Elasticity of Supply} &= \frac{26.1\%}{7.4\%} \\ &= 3.53 \end{aligned}$$

Again, as with the elasticity of demand, the elasticity of supply is not followed by any units. Elasticity is a ratio of one percentage change to another percentage change—nothing more—and we read it as an absolute value. In this case, a 1% rise in price causes an increase in quantity supplied of 3.5%. The greater than one elasticity of supply means that the percentage change in quantity supplied will be greater than a one percent price change. If you're starting to wonder if the concept of slope fits into this calculation, read the following Clear It Up box.

Clear It Up

Is the elasticity the slope?

It is a common mistake to confuse the slope of either the supply or demand curve with its elasticity. The slope is the rate of change in units along the curve, or the rise/run (change in y over the change in x). For example, in **Figure 5.2**, at each point shown on the demand curve, price drops by $10 and the number of units demanded increases by 200 compared to the point to its left. The slope is –10/200 along the entire demand curve and does not change. The price elasticity, however, changes along the curve. Elasticity between points A and B was 0.45 and increased to 1.47 between points G and H. Elasticity is the *percentage* change, which is a different calculation from the slope and has a different meaning.

When we are at the upper end of a demand curve, where price is high and the quantity demanded is low, a small change in the quantity demanded, even in, say, one unit, is pretty big in percentage terms. A change

in price of, say, a dollar, is going to be much less important in percentage terms than it would have been at the bottom of the demand curve. Likewise, at the bottom of the demand curve, that one unit change when the quantity demanded is high will be small as a percentage.

Thus, at one end of the demand curve, where we have a large percentage change in quantity demanded over a small percentage change in price, the elasticity value would be high, or demand would be relatively elastic. Even with the same change in the price and the same change in the quantity demanded, at the other end of the demand curve the quantity is much higher, and the price is much lower, so the percentage change in quantity demanded is smaller and the percentage change in price is much higher. That means at the bottom of the curve we'd have a small numerator over a large denominator, so the elasticity measure would be much lower, or inelastic.

As we move along the demand curve, the values for quantity and price go up or down, depending on which way we are moving, so the percentages for, say, a $1 difference in price or a one unit difference in quantity, will change as well, which means the ratios of those percentages and hence the elasticity will change.

5.2 | Polar Cases of Elasticity and Constant Elasticity

By the end of this section, you will be able to:

- Differentiate between infinite and zero elasticity
- Analyze graphs in order to classify elasticity as constant unitary, infinite, or zero

There are two extreme cases of elasticity: when elasticity equals zero and when it is infinite. A third case is that of constant unitary elasticity. We will describe each case. **Infinite elasticity** or **perfect elasticity** refers to the extreme case where either the quantity demanded (Qd) or supplied (Qs) changes by an infinite amount in response to any change in price at all. In both cases, the supply and the demand curve are horizontal as **Figure 5.4** shows. While perfectly elastic supply curves are for the most part unrealistic, goods with readily available inputs and whose production can easily expand will feature highly elastic supply curves. Examples include pizza, bread, books, and pencils. Similarly, perfectly elastic demand is an extreme example. However, luxury goods, items that take a large share of individuals' income, and goods with many substitutes are likely to have highly elastic demand curves. Examples of such goods are Caribbean cruises and sports vehicles.

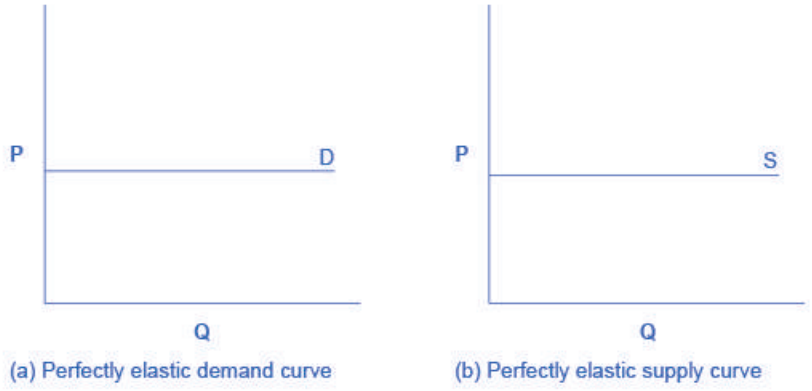


Figure 5.4 Infinite Elasticity The horizontal lines show that an infinite quantity will be demanded or supplied at a specific price. This illustrates the cases of a perfectly (or infinitely) elastic demand curve and supply curve. The quantity supplied or demanded is extremely responsive to price changes, moving from zero for prices close to P to infinite when prices reach P.

Zero elasticity or **perfect inelasticity**, as **Figure 5.5** depicts, refers to the extreme case in which a percentage change in price, no matter how large, results in zero change in quantity. While a perfectly inelastic supply is an extreme example, goods with limited supply of inputs are likely to feature highly inelastic supply curves. Examples include diamond rings or housing in prime locations such as apartments facing Central Park in New York City. Similarly,

while perfectly inelastic demand is an extreme case, necessities with no close substitutes are likely to have highly inelastic demand curves. This is the case of life-saving drugs and gasoline.

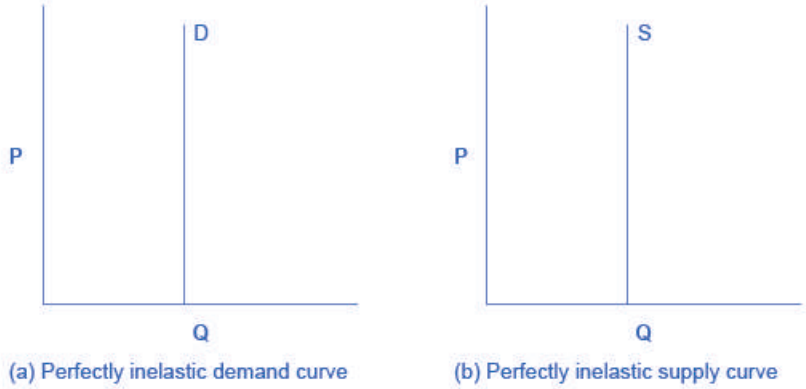


Figure 5.5 Zero Elasticity The vertical supply curve and vertical demand curve show that there will be zero percentage change in quantity (a) demanded or (b) supplied, regardless of the price.

Constant unitary elasticity, in either a supply or demand curve, occurs when a price change of one percent results in a quantity change of one percent. **Figure 5.6** shows a demand curve with constant unit elasticity. Constant unitary elasticity, in either a supply or demand curve, occurs when a price change of one percent results in a quantity change of one percent. Figure 5.6 shows a demand curve with constant unit elasticity. Using the midpoint method, you can calculate that between points A and B on the demand curve, the price changes by 28.6% and quantity demanded also changes by 28.6%. Hence, the elasticity equals 1. Between points B and C, price again changes by 28.6% as does quantity, while between points C and D the corresponding percentage changes are 22.2% for both price and quantity. In each case, then, the percentage change in price equals the percentage change in quantity, and consequently elasticity equals 1. Notice that in absolute value, the declines in price, as you step down the demand curve, are not identical. Instead, the price falls by $2.00 from A to B, by a smaller amount of $1.50 from B to C, and by a still smaller amount of $0.90 from C to D. As a result, a demand curve with constant unitary elasticity moves from a steeper slope on the left and a flatter slope on the right—and a curved shape overall. Notice that in absolute value, the declines in price, as you step down the demand curve, are not identical. Instead, the price falls by $23 from A to B, by a smaller amount of $1.50 from B to C, and by a still smaller amount of $.90 from C to D. As a result, a demand curve with constant unitary elasticity has a steeper slope on the left and a flatter slope on the right—and a curved shape overall.

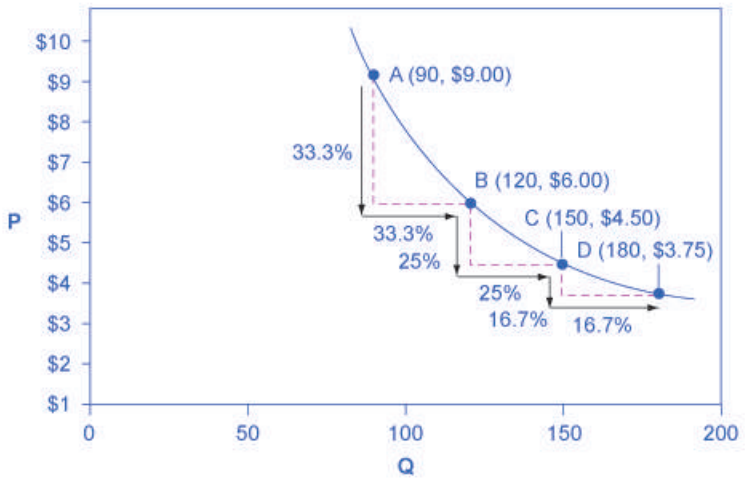


Figure 5.6 A Constant Unitary Elasticity Demand Curve A demand curve with constant unitary elasticity will be a curved line. Notice how price and quantity demanded change by an identical percentage amount between each pair of points on the demand curve.

Unlike the demand curve with unitary elasticity, the supply curve with unitary elasticity is represented by a straight

line, and that line goes through the origin. In each pair of points on the supply curve there is an equal difference in quantity of 30. However, in percentage value, using the midpoint method, the steps are decreasing as one moves from left to right, from 28.6% to 22.2% to 18.2%, because the quantity points in each percentage calculation are getting increasingly larger, which expands the denominator in the elasticity calculation of the percentage change in quantity.

Consider the price changes moving up the supply curve in **Figure 5.7**. From points D to E to F and to G on the supply curve, each step of $1.50 is the same in absolute value. However, if we measure the price changes in percentage change terms, using the midpoint method, they are also decreasing, from 28.6% to 22.2% to 18.2%, because the original price points in each percentage calculation are getting increasingly larger in value, increasing the denominator in the calculation of the percentage change in price. Along the constant unitary elasticity supply curve, the percentage quantity increases on the horizontal axis exactly match the percentage price increases on the vertical axis—so this supply curve has a constant unitary elasticity at all points.

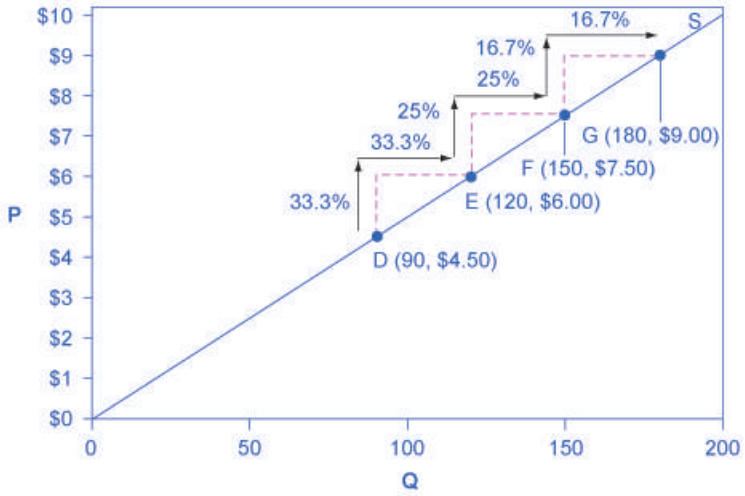


Figure 5.7 A Constant Unitary Elasticity Supply Curve A constant unitary elasticity supply curve is a straight line reaching up from the origin. Between each pair of points, the percentage increase in quantity supplied is the same as the percentage increase in price.

5.3 | Elasticity and Pricing

By the end of this section, you will be able to:

- Analyze how price elasticities impact revenue
- Evaluate how elasticity can cause shifts in demand and supply
- Predict how the long-run and short-run impacts of elasticity affect equilibrium
- Explain how the elasticity of demand and supply determine the incidence of a tax on buyers and sellers

Studying elasticities is useful for a number of reasons, pricing being most important. Let's explore how elasticity relates to revenue and pricing, both in the long and short run. First, let's look at the elasticities of some common goods and services.

Table 5.2 shows a selection of demand elasticities for different goods and services drawn from a variety of different studies by economists, listed in order of increasing elasticity.

Goods and Services	Elasticity of Price
Housing	0.12
Transatlantic air travel (economy class)	0.12
Rail transit (rush hour)	0.15
Electricity	0.20
Taxi cabs	0.22
Gasoline	0.35
Transatlantic air travel (first class)	0.40
Wine	0.55
Beef	0.59
Transatlantic air travel (business class)	0.62
Kitchen and household appliances	0.63
Cable TV (basic rural)	0.69
Chicken	0.64
Soft drinks	0.70
Beer	0.80
New vehicle	0.87
Rail transit (off-peak)	1.00
Computer	1.44
Cable TV (basic urban)	1.51
Cable TV (premium)	1.77
Restaurant meals	2.27

Table 5.2 Some Selected Elasticities of Demand

Note that demand for necessities such as housing and electricity is inelastic, while items that are not necessities such as restaurant meals are more price-sensitive. If the price of a restaurant meal increases by 10%, the quantity demanded will decrease by 22.7%. A 10% increase in the price of housing will cause only a slight decrease of 1.2% in the quantity of housing demanded.

Link It Up

Read this article (http://openstaxcollege.org/l/Movietickets) for an example of price elasticity that may have affected you.

Does Raising Price Bring in More Revenue?

Imagine that a band on tour is playing in an indoor arena with 15,000 seats. To keep this example simple, assume that the band keeps all the money from ticket sales. Assume further that the band pays the costs for its appearance, but that these costs, like travel, and setting up the stage, are the same regardless of how many people are in the audience. Finally, assume that all the tickets have the same price. (The same insights apply if ticket prices are more expensive for some seats than for others, but the calculations become more complicated.) The band knows that it faces a downward-sloping demand curve; that is, if the band raises the ticket price and, it will sell fewer seats. How should the band set the ticket price to generate the most total revenue, which in this example, because costs are fixed, will also mean the highest profits for the band? Should the band sell more tickets at a lower price or fewer tickets at a higher price?

The key concept in thinking about collecting the most revenue is the price elasticity of demand. Total revenue is price times the quantity of tickets sold. Imagine that the band starts off thinking about a certain price, which will result in the sale of a certain quantity of tickets. The three possibilities are in **Table 5.3**. If demand is elastic at that price level, then the band should cut the price, because the percentage drop in price will result in an even larger percentage increase in the quantity sold—thus raising total revenue. However, if demand is inelastic at that original quantity level, then the band should raise the ticket price, because a certain percentage increase in price will result in a smaller percentage decrease in the quantity sold—and total revenue will rise. If demand has a unitary elasticity at that quantity, then an equal percentage change in quantity will offset a moderate percentage change in the price—so the band will earn the same revenue whether it (moderately) increases or decreases the ticket price.

If Demand Is . . .	Then . . .	Therefore . . .
Elastic	% change in Qd > % change in P	A given % rise in P will be more than offset by a larger % fall in Q so that total revenue (P × Q) falls.
Unitary	% change in Qd = % change in P	A given % rise in P will be exactly offset by an equal % fall in Q so that total revenue (P × Q) is unchanged.
Inelastic	% change in Qd < % change in P	A given % rise in P will cause a smaller % fall in Q so that total revenue (P × Q) rises.

Table 5.3 Will the Band Earn More Revenue by Changing Ticket Prices?

What if the band keeps cutting price, because demand is elastic, until it reaches a level where it sells all 15,000 seats in the available arena? If demand remains elastic at that quantity, the band might try to move to a bigger arena, so that it could slash ticket prices further and see a larger percentage increase in the quantity of tickets sold. However, if the 15,000-seat arena is all that is available or if a larger arena would add substantially to costs, then this option may not work.

Conversely, a few bands are so famous, or have such fanatical followings, that demand for tickets may be inelastic right up to the point where the arena is full. These bands can, if they wish, keep raising the ticket price. Ironically,

some of the most popular bands could make more revenue by setting prices so high that the arena is not full—but those who buy the tickets would have to pay very high prices. However, bands sometimes choose to sell tickets for less than the absolute maximum they might be able to charge, often in the hope that fans will feel happier and spend more on recordings, T-shirts, and other paraphernalia.

Can Businesses Pass Costs on to Consumers?

Most businesses face a day-to-day struggle to figure out ways to produce at a lower cost, as one pathway to their goal of earning higher profits. However, in some cases, the price of a key input over which the firm has no control may rise. For example, many chemical companies use petroleum as a key input, but they have no control over the world market price for crude oil. Coffee shops use coffee as a key input, but they have no control over the world market price of coffee. If the cost of a key input rises, can the firm pass those higher costs along to consumers in the form of higher prices? Conversely, if new and less expensive ways of producing are invented, can the firm keep the benefits in the form of higher profits, or will the market pressure them to pass the gains along to consumers in the form of lower prices? The price elasticity of demand plays a key role in answering these questions.

Imagine that as a consumer of legal pharmaceutical products, you read a newspaper story that a technological breakthrough in the production of aspirin has occurred, so that every aspirin factory can now produce aspirin more cheaply. What does this discovery mean to you? **Figure 5.8** illustrates two possibilities. In **Figure 5.8** (a), the demand curve is highly inelastic. In this case, a technological breakthrough that shifts supply to the right, from S_0 to S_1, so that the equilibrium shifts from E_0 to E_1, creates a substantially lower price for the product with relatively little impact on the quantity sold. In **Figure 5.8** (b), the demand curve is highly elastic. In this case, the technological breakthrough leads to a much greater quantity sold in the market at very close to the original price. Consumers benefit more, in general, when the demand curve is more inelastic because the shift in the supply results in a much lower price for consumers.

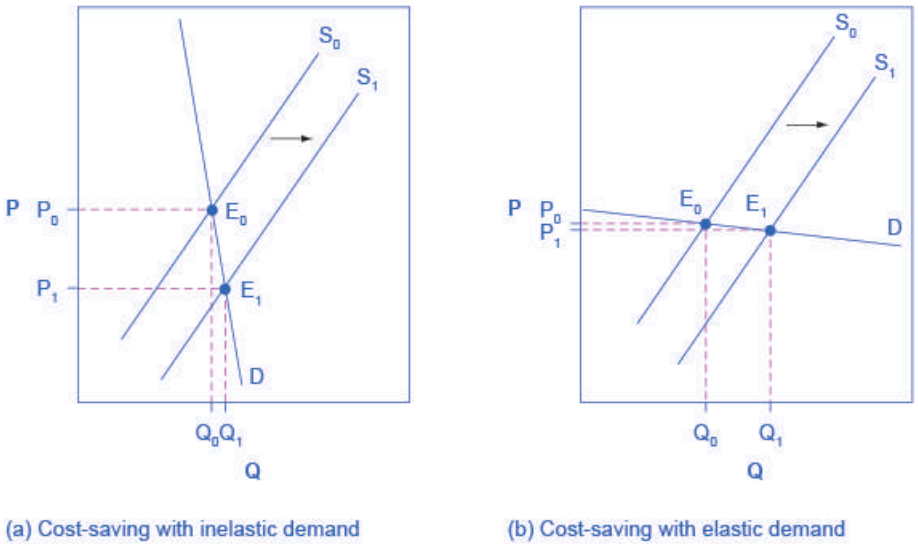


Figure 5.8 Passing along Cost Savings to Consumers Cost-saving gains cause supply to shift out to the right from S_0 to S_1; that is, at any given price, firms will be willing to supply a greater quantity. If demand is inelastic, as in (a), the result of this cost-saving technological improvement will be substantially lower prices. If demand is elastic, as in (b), the result will be only slightly lower prices. Consumers benefit in either case, from a greater quantity at a lower price, but the benefit is greater when demand is inelastic, as in (a).

Aspirin producers may find themselves in a nasty bind here. The situation in **Figure 5.8**, with extremely inelastic demand, means that a new invention may cause the price to drop dramatically while quantity changes little. As a result, the new production technology can lead to a drop in the revenue that firms earn from aspirin sales. However, if strong competition exists between aspirin producer, each producer may have little choice but to search for and implement any breakthrough that allows it to reduce production costs. After all, if one firm decides not to implement such a cost-saving technology, other firms that do can drive them out of business.

Since demand for food is generally inelastic, farmers may often face the situation in **Figure 5.8** (a). That is, a surge in

production leads to a severe drop in price that can actually decrease the total revenue that farmers receive. Conversely, poor weather or other conditions that cause a terrible year for farm production can sharply raise prices so that the total revenue that the farmer receives increases. The Clear It Up box discusses how these issues relate to coffee.

Clear It Up

How do coffee prices fluctuate?

Coffee is an international crop. The top five coffee-exporting nations are Brazil, Vietnam, Colombia, Indonesia, and Ethiopia. In these nations and others, 20 million families depend on selling coffee beans as their main source of income. These families are exposed to enormous risk, because the world price of coffee bounces up and down. For example, in 1993, the world price of coffee was about 50 cents per pound. In 1995 it was four times as high, at $2 per pound. By 1997 it had fallen by half to $1.00 per pound. In 1998 it leaped back up to $2 per pound. By 2001 it had fallen back to 46 cents a pound. By early 2011 it rose to about $2.31 per pound. By the end of 2012, the price had fallen back to about $1.31 per pound.

The reason for these price fluctuations lies in a combination of inelastic demand and shifts in supply. The elasticity of coffee demand is only about 0.3; that is, a 10% rise in the price of coffee leads to a decline of about 3% in the quantity of coffee consumed. When a major frost hit the Brazilian coffee crop in 1994, coffee supply shifted to the left with an inelastic demand curve, leading to much higher prices. Conversely, when Vietnam entered the world coffee market as a major producer in the late 1990s, the supply curve shifted out to the right. With a highly inelastic demand curve, coffee prices fell dramatically. **Figure 5.8** (a) illustrates this situation.

Elasticity also reveals whether firms can pass higher costs that they incur on to consumers. Addictive substances, for which demand is inelastic, are products for which producers can pass higher costs on to consumers. For example, the demand for cigarettes is relatively inelastic among regular smokers who are somewhat addicted. Economic research suggests that increasing cigarette prices by 10% leads to about a 3% reduction in the quantity of cigarettes that adults smoke, so the elasticity of demand for cigarettes is 0.3. If society increases taxes on companies that produce cigarettes, the result will be, as in **Figure 5.9** (a), that the supply curve shifts from S_0 to S_1. However, as the equilibrium moves from E_0 to E_1, governments mainly pass along these taxes to consumers in the form of higher prices. These higher taxes on cigarettes will raise tax revenue for the government, but they will not much affect the quantity of smoking.

If the goal is to reduce the quantity of cigarettes demanded, we must achieve it by shifting this inelastic demand back to the left, perhaps with public programs to discourage cigarette use or to help people to quit. For example, anti-smoking advertising campaigns have shown some ability to reduce smoking. However, if cigarette demand were more elastic, as in **Figure 5.9** (b), then an increase in taxes that shifts supply from S_0 to S_1 and equilibrium from E_0 to E_1 would reduce the quantity of cigarettes smoked substantially. Youth smoking seems to be more elastic than adult smoking—that is, the quantity of youth smoking will fall by a greater percentage than the quantity of adult smoking in response to a given percentage increase in price.

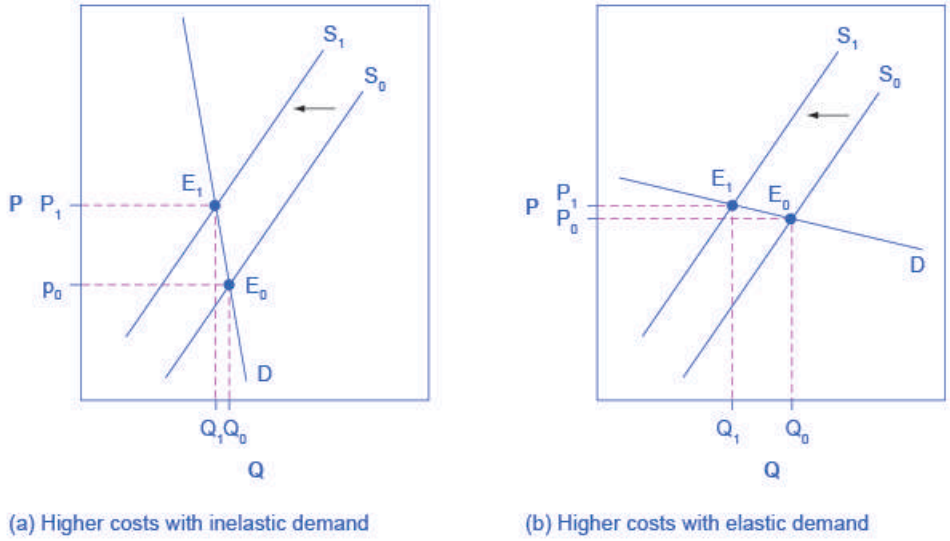


Figure 5.9 Passing along Higher Costs to Consumers Higher costs, like a higher tax on cigarette companies for the example we gave in the text, lead supply to shift to the left. This shift is identical in (a) and (b). However, in (a), where demand is inelastic, companies largely can pass the cost increase along to consumers in the form of higher prices, without much of a decline in equilibrium quantity. In (b), demand is elastic, so the shift in supply results primarily in a lower equilibrium quantity. Consumers suffer in either case, but in (a), they suffer from paying a higher price for the same quantity, while in (b), they suffer from buying a lower quantity (and presumably needing to shift their consumption elsewhere).

Elasticity and Tax Incidence

The example of cigarette taxes demonstrated that because demand is inelastic, taxes are not effective at reducing the equilibrium quantity of smoking, and they mainly pass along to consumers in the form of higher prices. The analysis, or manner, of how a tax burden is divided between consumers and producers is called **tax incidence**. Typically, the tax incidence, or burden, falls both on the consumers and producers of the taxed good. However, if one wants to predict which group will bear most of the burden, all one needs to do is examine the elasticity of demand and supply. In the tobacco example, the tax burden falls on the most inelastic side of the market.

If demand is more inelastic than supply, consumers bear most of the tax burden, and if supply is more inelastic than demand, sellers bear most of the tax burden.

The intuition for this is simple. When the demand is inelastic, consumers are not very responsive to price changes, and the quantity demanded reduces only modestly when the tax is introduced. In the case of smoking, the demand is inelastic because consumers are addicted to the product. The government can then pass the tax burden along to consumers in the form of higher prices, without much of a decline in the equilibrium quantity.

Similarly, when a government introduces a tax in a market with an inelastic supply, such as, for example, beachfront hotels, and sellers have no alternative than to accept lower prices for their business, taxes do not greatly affect the equilibrium quantity. The tax burden now passes on to the sellers. If the supply was elastic and sellers had the possibility of reorganizing their businesses to avoid supplying the taxed good, the tax burden on the sellers would be much smaller. The tax would result in a much lower quantity sold instead of lower prices received. **Figure 5.10** illustrates this relationship between the tax incidence and elasticity of demand and supply.

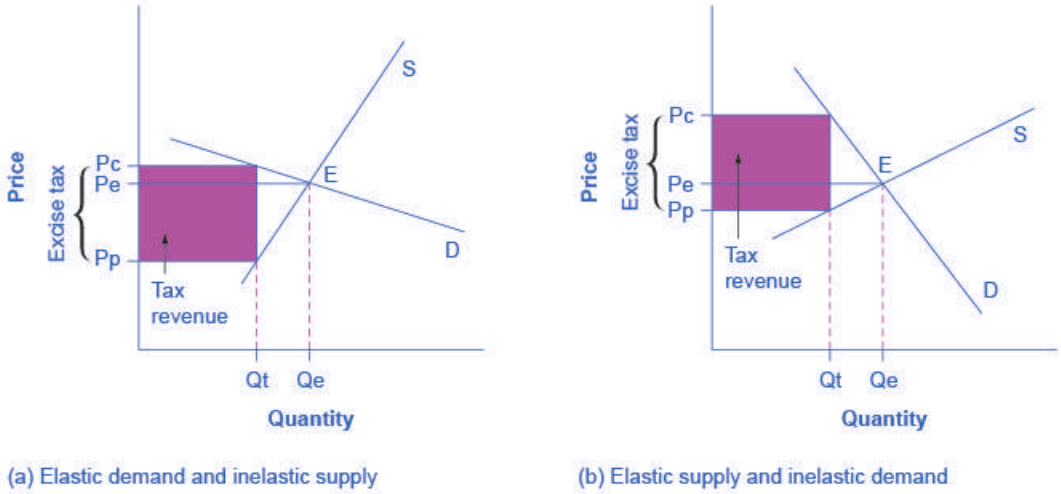


Figure 5.10 Elasticity and Tax Incidence An excise tax introduces a wedge between the price paid by consumers (Pc) and the price received by producers (Pp). The vertical distance between Pc and Pp is the amount of the tax per unit. Pe is the equilibrium price prior to introduction of the tax. (a) When the demand is more elastic than supply, the tax incidence on consumers Pc – Pe is lower than the tax incidence on producers Pe – Pp. (b) When the supply is more elastic than demand, the tax incidence on consumers Pc – Pe is larger than the tax incidence on producers Pe – Pp. The more elastic the demand and supply curves, the lower the tax revenue.

In **Figure 5.10** (a), the supply is inelastic and the demand is elastic, such as in the example of beachfront hotels. While consumers may have other vacation choices, sellers can't easily move their businesses. By introducing a tax, the government essentially creates a wedge between the price paid by consumers Pc and the price received by producers Pp. In other words, of the total price paid by consumers, part is retained by the sellers and part is paid to the government in the form of a tax. The distance between Pc and Pp is the tax rate. The new market price is Pc, but sellers receive only Pp per unit sold, as they pay Pc-Pp to the government. Since we can view a tax as raising the costs of production, this could also be represented by a leftward shift of the supply curve, where the new supply curve would intercept the demand at the new quantity Qt. For simplicity, **Figure 5.10** omits the shift in the supply curve.

The tax revenue is given by the shaded area, which we obtain by multiplying the tax per unit by the total quantity sold Qt. The tax incidence on the consumers is given by the difference between the price paid Pc and the initial equilibrium price Pe. The tax incidence on the sellers is given by the difference between the initial equilibrium price Pe and the price they receive after the tax is introduced Pp. In **Figure 5.10** (a), the tax burden falls disproportionately on the sellers, and a larger proportion of the tax revenue (the shaded area) is due to the resulting lower price received by the sellers than by the resulting higher prices paid by the buyers. **Figure 5.10** (b) describes the example of the tobacco excise tax where the supply is more elastic than demand. The tax incidence now falls disproportionately on consumers, as shown by the large difference between the price they pay, Pc, and the initial equilibrium price, Pe. Sellers receive a lower price than before the tax, but this difference is much smaller than the change in consumers' price. From this analysis one can also predict whether a tax is likely to create a large revenue or not. The more elastic the demand curve, the more likely that consumers will reduce quantity instead of paying higher prices. The more elastic the supply curve, the more likely that sellers will reduce the quantity sold, instead of taking lower prices. In a market where both the demand and supply are very elastic, the imposition of an excise tax generates low revenue.

Some believe that excise taxes hurt mainly the specific industries they target. For example, the medical device excise tax, in effect since 2013, has been controversial for it can delay industry profitability and therefore hamper start-ups and medical innovation. However, whether the tax burden falls mostly on the medical device industry or on the patients depends simply on the elasticity of demand and supply.

Long-Run vs. Short-Run Impact

Elasticities are often lower in the short run than in the long run. On the demand side of the market, it can sometimes be difficult to change Qd in the short run, but easier in the long run. Consumption of energy is a clear example. In the short run, it is not easy for a person to make substantial changes in energy consumption. Maybe you can carpool to work sometimes or adjust your home thermostat by a few degrees if the cost of energy rises, but that is about all.

However, in the long run you can purchase a car that gets more miles to the gallon, choose a job that is closer to where you live, buy more energy-efficient home appliances, or install more insulation in your home. As a result, the elasticity of demand for energy is somewhat inelastic in the short run, but much more elastic in the long run.

Figure 5.11 is an example, based roughly on historical experience, for the responsiveness of Qd to price changes. In 1973, the price of crude oil was $12 per barrel and total consumption in the U.S. economy was 17 million barrels per day. That year, the nations who were members of the Organization of Petroleum Exporting Countries (OPEC) cut off oil exports to the United States for six months because the Arab members of OPEC disagreed with the U.S. support for Israel. OPEC did not bring exports back to their earlier levels until 1975—a policy that we can interpret as a shift of the supply curve to the left in the U.S. petroleum market. **Figure 5.11** (a) and **Figure 5.11** (b) show the same original equilibrium point and the same identical shift of a supply curve to the left from S_0 to S_1.

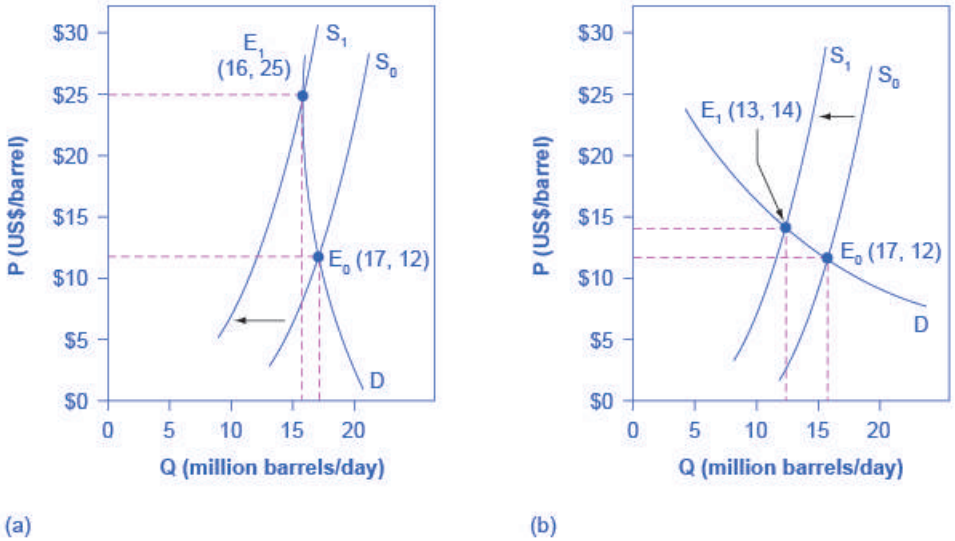


Figure 5.11 How a Shift in Supply Can Affect Price or Quantity The intersection (E_0) between demand curve D and supply curve S_0 is the same in both (a) and (b). The shift of supply to the left from S_0 to S_1 is identical in both (a) and (b). The new equilibrium (E_1) has a higher price and a lower quantity than the original equilibrium (E_0) in both (a) and (b). However, the shape of the demand curve D is different in (a) and (b), being more elastic in (b) than in (a). As a result, the shift in supply can result either in a new equilibrium with a much higher price and an only slightly smaller quantity, as in (a), with more inelastic demand, or in a new equilibrium with only a small increase in price and a relatively larger reduction in quantity, as in (b), with more elastic demand.

Figure 5.11 (a) shows inelastic demand for oil in the short run similar to that which existed for the United States in 1973. In **Figure 5.11** (a), the new equilibrium (E_1) occurs at a price of $25 per barrel, roughly double the price before the OPEC shock, and an equilibrium quantity of 16 million barrels per day. **Figure 5.11** (b) shows what the outcome would have been if the U.S. demand for oil had been more elastic, a result more likely over the long term. This alternative equilibrium (E_1) would have resulted in a smaller price increase to $14 per barrel and larger reduction in equilibrium quantity to 13 million barrels per day. In 1983, for example, U.S. petroleum consumption was 15.3 million barrels a day, which was lower than in 1973 or 1975. U.S. petroleum consumption was down even though the U.S. economy was about one-fourth larger in 1983 than it had been in 1973. The primary reason for the lower quantity was that higher energy prices spurred conservation efforts, and after a decade of home insulation, more fuel-efficient cars, more efficient appliances and machinery, and other fuel-conserving choices, the demand curve for energy had become more elastic.

On the supply side of markets, producers of goods and services typically find it easier to expand production in the long term of several years rather than in the short run of a few months. After all, in the short run it can be costly or difficult to build a new factory, hire many new workers, or open new stores. However, over a few years, all of these are possible.

In most markets for goods and services, prices bounce up and down more than quantities in the short run, but quantities often move more than prices in the long run. The underlying reason for this pattern is that supply and demand are often inelastic in the short run, so that shifts in either demand or supply can cause a relatively greater

change in prices. However, since supply and demand are more elastic in the long run, the long-run movements in prices are more muted, while quantity adjusts more easily in the long run.

5.4 | Elasticity in Areas Other Than Price

By the end of this section, you will be able to:

- Calculate the income elasticity of demand and the cross-price elasticity of demand
- Calculate the elasticity in labor and financial capital markets through an understanding of the elasticity of labor supply and the elasticity of savings
- Apply concepts of price elasticity to real-world situations

The basic idea of elasticity—how a percentage change in one variable causes a percentage change in another variable—does not just apply to the responsiveness quantity supplied and quantity demanded to changes in the price of a product. Recall that quantity demanded (Qd) depends on income, tastes and preferences, the prices of related goods, and so on, as well as price. Similarly, quantity supplied (Qs) depends on factors such as the cost of production, as well as price. We can measure elasticity for any determinant of quantity supplied and quantity demanded, not just the price.

Income Elasticity of Demand

The income elasticity of demand is the percentage change in quantity demanded divided by the percentage change in income.

$$\text{Income elasticity of demand} = \frac{\%\text{ change in quantity demanded}}{\%\text{ change in income}}$$

For most products, most of the time, the income elasticity of demand is positive: that is, a rise in income will cause an increase in the quantity demanded. This pattern is common enough that we refer to these goods as normal goods. However, for a few goods, an increase in income means that one might purchase less of the good. For example, those with a higher income might buy fewer hamburgers, because they are buying more steak instead, or those with a higher income might buy less cheap wine and more imported beer. When the income elasticity of demand is negative, we call the good an inferior good.

We introduced the concepts of normal and inferior goods in **Demand and Supply**. A higher level of income causes a demand curve to shift to the right for a normal good, which means that the income elasticity of demand is positive. How far the demand shifts depends on the income elasticity of demand. A higher income elasticity means a larger shift. However, for an inferior good, that is, when the income elasticity of demand is negative, a higher level of income would cause the demand curve for that good to shift to the left. Again, how much it shifts depends on how large the (negative) income elasticity is.

Cross-Price Elasticity of Demand

A change in the price of one good can shift the quantity demanded for another good. If the two goods are complements, like bread and peanut butter, then a drop in the price of one good will lead to an increase in the quantity demanded of the other good. However, if the two goods are substitutes, like plane tickets and train tickets, then a drop in the price of one good will cause people to substitute toward that good, and to reduce consumption of the other good. Cheaper plane tickets lead to fewer train tickets, and vice versa.

The **cross-price elasticity of demand** puts some meat on the bones of these ideas. The term “cross-price” refers to the idea that the price of one good is affecting the quantity demanded of a different good. Specifically, the cross-price elasticity of demand is the percentage change in the quantity of good A that is demanded as a result of a percentage change in the price of good B.

$$\text{Cross-price elasticity of demand} = \frac{\%\text{ change in Qd of good A}}{\%\text{ change in price of good B}}$$

Substitute goods have positive cross-price elasticities of demand: if good A is a substitute for good B, like coffee and tea, then a higher price for B will mean a greater quantity consumed of A. Complement goods have negative cross-price elasticities: if good A is a complement for good B, like coffee and sugar, then a higher price for B will mean a

lower quantity consumed of A.

Elasticity in Labor and Financial Capital Markets

The concept of elasticity applies to any market, not just markets for goods and services. In the labor market, for example, the **wage elasticity of labor supply**—that is, the percentage change in hours worked divided by the percentage change in wages—will reflect the shape of the labor supply curve. Specifically:

$$\text{Elasticity of labor supply} = \frac{\%\text{ change in quantity of labor supplied}}{\%\text{ change in wage}}$$

The wage elasticity of labor supply for teenage workers is generally fairly elastic: that is, a certain percentage change in wages will lead to a larger percentage change in the quantity of hours worked. Conversely, the wage elasticity of labor supply for adult workers in their thirties and forties is fairly inelastic. When wages move up or down by a certain percentage amount, the quantity of hours that adults in their prime earning years are willing to supply changes but by a lesser percentage amount.

In markets for financial capital, the **elasticity of savings**—that is, the percentage change in the quantity of savings divided by the percentage change in interest rates—will describe the shape of the supply curve for financial capital. That is:

$$\text{Elasticity of savings} = \frac{\%\text{ change in quantity of financial savings}}{\%\text{ change in interest rate}}$$

Sometimes laws are proposed that seek to increase the quantity of savings by offering tax breaks so that the return on savings is higher. Such a policy will have a comparatively large impact on increasing the quantity saved if the supply curve for financial capital is elastic, because then a given percentage increase in the return to savings will cause a higher percentage increase in the quantity of savings. However, if the supply curve for financial capital is highly inelastic, then a percentage increase in the return to savings will cause only a small increase in the quantity of savings. The evidence on the supply curve of financial capital is controversial but, at least in the short run, the elasticity of savings with respect to the interest rate appears fairly inelastic.

Expanding the Concept of Elasticity

The elasticity concept does not even need to relate to a typical supply or demand curve at all. For example, imagine that you are studying whether the Internal Revenue Service should spend more money on auditing tax returns. We can frame the question in terms of the elasticity of tax collections with respect to spending on tax enforcement; that is, what is the percentage change in tax collections derived from a given percentage change in spending on tax enforcement?

With all of the elasticity concepts that we have just described, some of which are in **Table 5.4**, the possibility of confusion arises. When you hear the phrases “elasticity of demand” or “elasticity of supply,” they refer to the elasticity with respect to price. Sometimes, either to be extremely clear or because economists are discussing a wide variety of elasticities, we will call the elasticity of demand or the demand elasticity the price elasticity of demand or the “elasticity of demand with respect to price.” Similarly, economists sometimes use the term elasticity of supply or the supply elasticity, to avoid any possibility of confusion, the price elasticity of supply or “the elasticity of supply with respect to price.” However, in whatever context, the idea of elasticity always refers to percentage change in one variable, almost always a price or money variable, and how it causes a percentage change in another variable, typically a quantity variable of some kind.

Income elasticity of demand	$= \frac{\%\text{ change in Qd}}{\%\text{ change in income}}$
Cross-price elasticity of demand	$= \frac{\%\text{ change in Qd of good A}}{\%\text{ change in price of good B}}$

Table 5.4 Formulas for Calculating Elasticity

Wage elasticity of labor supply	$= \frac{\%\text{ change in quantity of labor supplied}}{\%\text{ change in wage}}$
Wage elasticity of labor demand	$= \frac{\%\text{ change in quantity of labor demanded}}{\%\text{ change in wage}}$
Interest rate elasticity of savings	$= \frac{\%\text{ change in quantity of savings}}{\%\text{ change in interest rate}}$
Interest rate elasticity of borrowing	$= \frac{\%\text{ change in quantity of borrowing}}{\%\text{ change in interest rate}}$

Table 5.4 Formulas for Calculating Elasticity

Bring it Home

That Will Be How Much?

How did the 60% price increase in 2011 end up for Netflix? It has been a very bumpy ride.

Before the price increase, there were about 24.6 million U.S. subscribers. After the price increase, 810,000 infuriated U.S. consumers canceled their Netflix subscriptions, dropping the total number of subscribers to 23.79 million. Fast forward to June 2013, when there were 36 million streaming Netflix subscribers in the United States. This was an increase of 11.4 million subscribers since the price increase—an average per quarter growth of about 1.6 million. This growth is less than the 2 million per quarter increases Netflix experienced in the fourth quarter of 2010 and the first quarter of 2011.

During the first year after the price increase, the firm's stock price (a measure of future expectations for the firm) fell from about $33.60 per share per share to just under $7.80. By the end of 2016, however, the stock price was at $123 per share. Today, Netflix has more than 86 million subscribers million subscribers in fifty countries.

What happened? Obviously, Netflix company officials understood the law of demand. Company officials reported, when announcing the price increase, this could result in the loss of about 600,000 existing subscribers. Using the elasticity of demand formula, it is easy to see company officials expected an inelastic response:

$$\begin{aligned} &= \frac{-600{,}000/[(24\text{ million} + 24.6\text{ million})/2]}{\$6/[(\$10 + \$16)/2]} \\ &= \frac{-600{,}000/24.3\text{ million}}{\$6/\$13} \\ &= \frac{-0.025}{0.46} \\ &= -0.05 \end{aligned}$$

In addition, Netflix officials had anticipated the price increase would have little impact on attracting new customers. Netflix anticipated adding up to 1.29 million new subscribers in the third quarter of 2011. It is true this was slower growth than the firm had experienced—about 2 million per quarter.

Why was the estimate of customers leaving so far off? In the more than two decades since Netflix had been founded, there was an increase in the number of close, but not perfect, substitutes. Consumers now had choices ranging from Vudu, Amazon Prime, Hulu, and Redbox, to retail stores. Jaime Weinman reported in *Maclean's* that Redbox kiosks are "a five-minute drive for less from 68 percent of Americans, and it seems that many people still find a five-minute drive more convenient than loading up a movie online." It seems that in 2012, many consumers still preferred a physical DVD disk over streaming video.

What missteps did the Netflix management make? In addition to misjudging the elasticity of demand, by failing

to account for close substitutes, it seems they may have also misjudged customers' preferences and tastes. Yet, as the population increases, the preference for streaming video may overtake physical DVD disks. Netflix, the source of numerous late night talk show laughs and jabs in 2011, may yet have the last laugh.

KEY TERMS

constant unitary elasticity when a given percent price change in price leads to an equal percentage change in quantity demanded or supplied

cross-price elasticity of demand the percentage change in the quantity of good A that is demanded as a result of a percentage change in good B

elastic demand when the elasticity of demand is greater than one, indicating a high responsiveness of quantity demanded or supplied to changes in price

elastic supply when the elasticity of either supply is greater than one, indicating a high responsiveness of quantity demanded or supplied to changes in price

elasticity an economics concept that measures responsiveness of one variable to changes in another variable

elasticity of savings the percentage change in the quantity of savings divided by the percentage change in interest rates

inelastic demand when the elasticity of demand is less than one, indicating that a 1 percent increase in price paid by the consumer leads to less than a 1 percent change in purchases (and vice versa); this indicates a low responsiveness by consumers to price changes

inelastic supply when the elasticity of supply is less than one, indicating that a 1 percent increase in price paid to the firm will result in a less than 1 percent increase in production by the firm; this indicates a low responsiveness of the firm to price increases (and vice versa if prices drop)

infinite elasticity the extremely elastic situation of demand or supply where quantity changes by an infinite amount in response to any change in price; horizontal in appearance

perfect elasticity see infinite elasticity

perfect inelasticity see zero elasticity

price elasticity the relationship between the percent change in price resulting in a corresponding percentage change in the quantity demanded or supplied

price elasticity of demand percentage change in the quantity *demanded* of a good or service divided the percentage change in price

price elasticity of supply percentage change in the quantity *supplied* divided by the percentage change in price

tax incidence manner in which the tax burden is divided between buyers and sellers

unitary elasticity when the calculated elasticity is equal to one indicating that a change in the price of the good or service results in a proportional change in the quantity demanded or supplied

wage elasticity of labor supply the percentage change in hours worked divided by the percentage change in wages

zero inelasticity the highly inelastic case of demand or supply in which a percentage change in price, no matter how large, results in zero change in the quantity; vertical in appearance

KEY CONCEPTS AND SUMMARY

5.1 Price Elasticity of Demand and Price Elasticity of Supply

Price elasticity measures the responsiveness of the quantity demanded or supplied of a good to a change in its price. We compute it as the percentage change in quantity demanded (or supplied) divided by the percentage change in price. We can describe elasticity as elastic (or very responsive), unit elastic, or inelastic (not very responsive). Elastic demand or supply curves indicate that quantity demanded or supplied respond to price changes in a greater than proportional manner. An inelastic demand or supply curve is one where a given percentage change in price will cause a smaller percentage change in quantity demanded or supplied. A unitary elasticity means that a given percentage change in price leads to an equal percentage change in quantity demanded or supplied.

5.2 Polar Cases of Elasticity and Constant Elasticity

Infinite or perfect elasticity refers to the extreme case where either the quantity demanded or supplied changes by an infinite amount in response to any change in price at all. Zero elasticity refers to the extreme case in which a percentage change in price, no matter how large, results in zero change in quantity. Constant unitary elasticity in either a supply or demand curve refers to a situation where a price change of one percent results in a quantity change of one percent.

5.3 Elasticity and Pricing

In the market for goods and services, quantity supplied and quantity demanded are often relatively slow to react to changes in price in the short run, but react more substantially in the long run. As a result, demand and supply often (but not always) tend to be relatively inelastic in the short run and relatively elastic in the long run. A tax incidence depends on the relative price elasticity of supply and demand. When supply is more elastic than demand, buyers bear most of the tax burden, and when demand is more elastic than supply, producers bear most of the cost of the tax. Tax revenue is larger the more inelastic the demand and supply are.

5.4 Elasticity in Areas Other Than Price

Elasticity is a general term, that reflects responsiveness. It refers to the change of one variable divided by the percentage change of a related variable that we can apply to many economic connections. For instance, the income elasticity of demand is the percentage change in quantity demanded divided by the percentage change in income. The cross-price elasticity of demand is the percentage change in the quantity demanded of a good divided by the percentage change in the price of another good. Elasticity applies in labor markets and financial capital markets just as it does in markets for goods and services. The wage elasticity of labor supply is the percentage change in the quantity of hours supplied divided by the percentage change in the wage. The elasticity of savings with respect to interest rates is the percentage change in the quantity of savings divided by the percentage change in interest rates.

SELF-CHECK QUESTIONS

1. From the data in **Table 5.5** about demand for smart phones, calculate the price elasticity of demand from: point B to point C, point D to point E, and point G to point H. Classify the elasticity at each point as elastic, inelastic, or unit elastic.

Points	P	Q
A	60	3,000
B	70	2,800
C	80	2,600
D	90	2,400
E	100	2,200
F	110	2,000
G	120	1,800
H	130	1,600

Table 5.5

2. From the data in **Table 5.6** about supply of alarm clocks, calculate the price elasticity of supply from: point J to point K, point L to point M, and point N to point P. Classify the elasticity at each point as elastic, inelastic, or unit elastic.

Point	Price	Quantity Supplied
J	$8	50
K	$9	70
L	$10	80
M	$11	88
N	$12	95
P	$13	100

Table 5.6

3. Why is the demand curve with constant unitary elasticity concave?

4. Why is the supply curve with constant unitary elasticity a straight line?

5. The federal government decides to require that automobile manufacturers install new anti-pollution equipment that costs $2,000 per car. Under what conditions can carmakers pass almost all of this cost along to car buyers? Under what conditions can carmakers pass very little of this cost along to car buyers?

6. Suppose you are in charge of sales at a pharmaceutical company, and your firm has a new drug that causes bald men to grow hair. Assume that the company wants to earn as much revenue as possible from this drug. If the elasticity of demand for your company's product at the current price is 1.4, would you advise the company to raise the price, lower the price, or to keep the price the same? What if the elasticity were 0.6? What if it were 1? Explain your answer.

7. What would the gasoline price elasticity of supply mean to UPS or FedEx?

8. The average annual income rises from $25,000 to $38,000, and the quantity of bread consumed in a year by the average person falls from 30 loaves to 22 loaves. What is the income elasticity of bread consumption? Is bread a normal or an inferior good?

9. Suppose the cross-price elasticity of apples with respect to the price of oranges is 0.4, and the price of oranges falls by 3%. What will happen to the demand for apples?

REVIEW QUESTIONS

10. What is the formula for calculating elasticity?

11. What is the price elasticity of demand? Can you explain it in your own words?

12. What is the price elasticity of supply? Can you explain it in your own words?

13. Describe the general appearance of a demand or a supply curve with zero elasticity.

14. Describe the general appearance of a demand or a supply curve with infinite elasticity.

15. If demand is elastic, will shifts in supply have a larger effect on equilibrium quantity or on price?

16. If demand is inelastic, will shifts in supply have a larger effect on equilibrium price or on quantity?

17. If supply is elastic, will shifts in demand have a larger effect on equilibrium quantity or on price?

18. If supply is inelastic, will shifts in demand have a larger effect on equilibrium price or on quantity?

19. Would you usually expect elasticity of demand or supply to be higher in the short run or in the long run? Why?

20. Under which circumstances does the tax burden fall entirely on consumers?

21. What is the formula for the income elasticity of demand?

22. What is the formula for the cross-price elasticity of demand?

23. What is the formula for the wage elasticity of labor supply?

24. What is the formula for elasticity of savings with respect to interest rates?

CRITICAL THINKING QUESTIONS

25. Transatlantic air travel in business class has an estimated elasticity of demand of 0.62, while transatlantic air travel in economy class has an estimated price elasticity of 0.12. Why do you think this is the case?

26. What is the relationship between price elasticity and position on the demand curve? For example, as you move up the demand curve to higher prices and lower quantities, what happens to the measured elasticity? How would you explain that?

27. Can you think of an industry (or product) with near infinite elasticity of supply in the short term? That is, what is an industry that could increase Qs almost without limit in response to an increase in the price?

28. Would you expect supply to play a more significant role in determining the price of a basic necessity like food or a luxury like perfume? Explain. *Hint*: Think about how the price elasticity of demand will differ between necessities and luxuries.

29. A city has built a bridge over a river and it decides to charge a toll to everyone who crosses. For one year, the city charges a variety of different tolls and records information on how many drivers cross the bridge. The city thus gathers information about elasticity of demand. If the city wishes to raise as much revenue as possible from the tolls, where will the city decide to charge a toll: in the inelastic portion of the demand curve, the elastic portion of the demand curve, or the unit elastic portion? Explain.

30. In a market where the supply curve is perfectly inelastic, how does an excise tax affect the price paid by consumers and the quantity bought and sold?

31. Economists define normal goods as having a positive income elasticity. We can divide normal goods into two types: Those whose income elasticity is less than one and those whose income elasticity is greater than one. Think about products that would fall into each category. Can you come up with a name for each category?

32. Suppose you could buy shoes one at a time, rather than in pairs. What do you predict the cross-price elasticity for left shoes and right shoes would be?

PROBLEMS

33. The equation for a demand curve is $P = 48 – 3Q$. What is the elasticity in moving from a quantity of 5 to a quantity of 6?

34. The equation for a demand curve is $P = 2/Q$. What is the elasticity of demand as price falls from 5 to 4? What is the elasticity of demand as the price falls from 9 to 8? Would you expect these answers to be the same?

35. The equation for a supply curve is $4P = Q$. What is the elasticity of supply as price rises from 3 to 4? What is the elasticity of supply as the price rises from 7 to 8? Would you expect these answers to be the same?

36. The equation for a supply curve is $P = 3Q – 8$. What is the elasticity in moving from a price of 4 to a price of 7?

37. The supply of paintings by Leonardo Da Vinci, who painted the *Mona Lisa* and *The Last Supper* and died in 1519, is highly inelastic. Sketch a supply and demand diagram, paying attention to the appropriate elasticities, to illustrate that demand for these paintings will determine the price.

38. Say that a certain stadium for professional football has 70,000 seats. What is the shape of the supply curve for tickets to football games at that stadium? Explain.

39. When someone's kidneys fail, the person needs to have medical treatment with a dialysis machine (unless or until they receive a kidney transplant) or they will die. Sketch a supply and demand diagram, paying attention to the appropriate elasticities, to illustrate that the supply of such dialysis machines will primarily determine the price.

40. Assume that the supply of low-skilled workers is fairly elastic, but the employers' demand for such workers is fairly inelastic. If the policy goal is to expand employment for low-skilled workers, is it better to focus on policy tools to shift the supply of unskilled labor or on tools to shift the demand for unskilled labor? What if the policy goal is to raise wages for this group? Explain your answers with supply and demand diagrams.

6 | The Macroeconomic Perspective

Figure 6.1 The Great Depression At times, such as when many people having trouble making ends meet, it is easy to tell how the economy is doing. This photograph shows people lined up during the Great Depression, waiting for relief checks. At other times, when some are doing well and others are not, it is more difficult to ascertain how the economy of a country is doing. (Credit: modification of work by the U.S. Library of Congress/Wikimedia Commons)

Bring it Home

How is the Economy Doing? How Does One Tell?

The 1990s were boom years for the U.S. economy. Beginning in the late 2000s, from 2007 to 2014 economic performance in in the U.S. was poor. What causes the economy to expand or contract? Why do businesses fail when they are making all the right decisions? Why do workers lose their jobs when they are hardworking and productive? Are bad economic times a failure of the market system? Are they a failure of the government? These are all questions of macroeconomics, which we will begin to address in this chapter. We will not be able to answer all of these questions here, but we will start with the basics: How is the economy doing? How can we tell?

The macro economy includes all buying and selling, all production and consumption; everything that goes on in every market in the economy. How can we get a handle on that? The answer begins more than 80 years ago, during the Great Depression. President Franklin D. Roosevelt and his economic advisers knew things were bad—but how could they express and measure just how bad it was? An economist named Simon Kuznets, who later won the Nobel Prize for his work, came up with a way to track what the entire economy is producing. In this chapter, you will learn how the government constructs GDP, how we use it, and why it is so important.

Introduction to the Macroeconomic Perspective

In this chapter, you will learn about:

- Measuring the Size of the Economy: Gross Domestic Product
- Adjusting Nominal Values to Real Values
- Tracking Real GDP over Time
- Comparing GDP among Countries
- How Well GDP Measures the Well-Being of Society

Macroeconomics focuses on the economy as a whole (or on whole economies as they interact). What causes recessions? What makes unemployment stay high when recessions are supposed to be over? Why do some countries grow faster than others? Why do some countries have higher standards of living than others? These are all questions that macroeconomics addresses. Macroeconomics involves adding up the economic activity of all households and all businesses in all markets to obtain the overall demand and supply in the economy. However, when we do that, something curious happens. It is not unusual that what results at the macro level is different from the sum of the microeconomic parts. What seems sensible from a microeconomic point of view can have unexpected or counterproductive results at the macroeconomic level. Imagine that you are sitting at an event with a large audience, like a live concert or a basketball game. A few people decide that they want a better view, and so they stand up. However, when these people stand up, they block the view for other people, and the others need to stand up as well if they wish to see. Eventually, nearly everyone is standing up, and as a result, no one can see much better than before. The rational decision of some individuals at the micro level—to stand up for a better view—ended up as self-defeating at the macro level. This is not macroeconomics, but it is an apt analogy.

Macroeconomics is a rather massive subject. How are we going to tackle it? **Figure 6.2** illustrates the structure we will use. We will study macroeconomics from three different perspectives:

1. What are the macroeconomic goals? (Macroeconomics as a discipline does not have goals, but we do have goals for the macro economy.)
2. What are the frameworks economists can use to analyze the macroeconomy?
3. Finally, what are the policy tools governments can use to manage the macroeconomy?

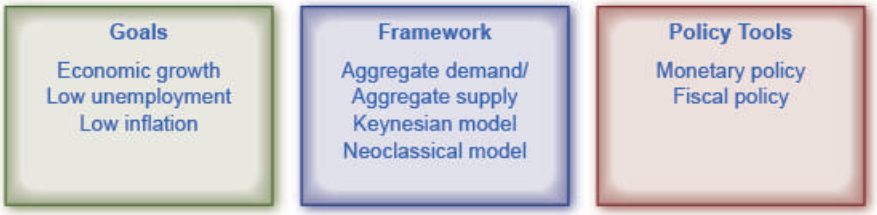


Figure 6.2 Macroeconomic Goals, Framework, and Policies This chart shows what macroeconomics is about. The box on the left indicates a consensus of what are the most important goals for the macro economy, the middle box lists the frameworks economists use to analyze macroeconomic changes (such as inflation or recession), and the box on the right indicates the two tools the federal government uses to influence the macro economy.

Goals

In thinking about the macroeconomy's overall health, it is useful to consider three primary goals: economic growth, low unemployment, and low inflation.

- Economic growth ultimately determines the prevailing standard of living in a country. Economists measure growth by the percentage change in real (inflation-adjusted) gross domestic product. A growth rate of more than 3% is considered good.
- Unemployment, as measured by the unemployment rate, is the percentage of people in the labor force who do not have a job. When people lack jobs, the economy is wasting a precious resource-labor, and the result is

lower goods and services produced. Unemployment, however, is more than a statistic—it represents people's livelihoods. While measured unemployment is unlikely to ever be zero, economists consider a measured unemployment rate of 5% or less low (good).

- Inflation is a sustained increase in the overall level of prices, and is measured by the consumer price index. If many people face a situation where the prices that they pay for food, shelter, and healthcare are rising much faster than the wages they receive for their labor, there will be widespread unhappiness as their standard of living declines. For that reason, low inflation—an inflation rate of 1–2%—is a major goal.

Frameworks

As you learn in the micro part of this book, principal tools that economists use are theories and models (see **Welcome to Economics!** for more on this). In microeconomics, we used the theories of supply and demand. In macroeconomics, we use the theories of aggregate demand (AD) and aggregate supply (AS). This book presents two perspectives on macroeconomics: the Neoclassical perspective and the Keynesian perspective, each of which has its own version of AD and AS. Between the two perspectives, you will obtain a good understanding of what drives the macroeconomy.

Policy Tools

National governments have two tools for influencing the macroeconomy. The first is monetary policy, which involves managing the money supply and interest rates. The second is fiscal policy, which involves changes in government spending/purchases and taxes.

We will explain each of the items in **Figure 6.2** in detail in one or more other chapters. As you learn these things, you will discover that the goals and the policy tools are in the news almost every day.

6.1 | Measuring the Size of the Economy: Gross Domestic Product

By the end of this section, you will be able to:

- Identify the components of GDP on the demand side and on the supply side
- Evaluate how economists measure gross domestic product (GDP)
- Contrast and calculate GDP, net exports, and net national product

Macroeconomics is an empirical subject, so the first step toward understanding it is to measure the economy.

How large is the U.S. economy? Economists typically measure the size of a nation's overall economy by its **gross domestic product (GDP)**, which is the value of all final goods and services produced within a country in a given year. Measuring GDP involves counting the production of millions of different goods and services—smart phones, cars, music downloads, computers, steel, bananas, college educations, and all other new goods and services that a country produced in the current year—and summing them into a total dollar value. This task is straightforward: take the quantity of everything produced, multiply it by the price at which each product sold, and add up the total. In 2016, the U.S. GDP totaled $18.6 trillion, the largest GDP in the world.

Each of the market transactions that enter into GDP must involve both a buyer and a seller. We can measure an economy's GDP either by the total dollar value of what consumers purchase in the economy, or by the total dollar value of what is the country produces. There is even a third way, as we will explain later.

GDP Measured by Components of Demand

Who buys all of this production? We can divide this demand into four main parts: consumer spending (consumption), business spending (investment), government spending on goods and services, and spending on net exports. (See the following Clear It Up feature to understand what we mean by investment.) **Table 6.1** shows how these four components added up to the GDP in 2016. **Figure 6.4** (a) shows the levels of consumption, investment, and government purchases over time, expressed as a percentage of GDP, while **Figure 6.4** (b) shows the levels of exports and imports as a percentage of GDP over time. A few patterns about each of these components are worth noticing. **Table 6.1** shows the components of GDP from the demand side.

	Components of GDP on the Demand Side (in trillions of dollars)	Percentage of Total
Consumption	$12.8	68.8%
Investment	$3.0	16.1%
Government	$3.3	17.7%
Exports	$2.2	11.8%
Imports	–$2.7	–14.5%
Total GDP	**$18.6**	**100%**

Table 6.1 Components of U.S. GDP in 2016: From the Demand Side (Source: http://bea.gov/iTable/index_nipa.cfm)

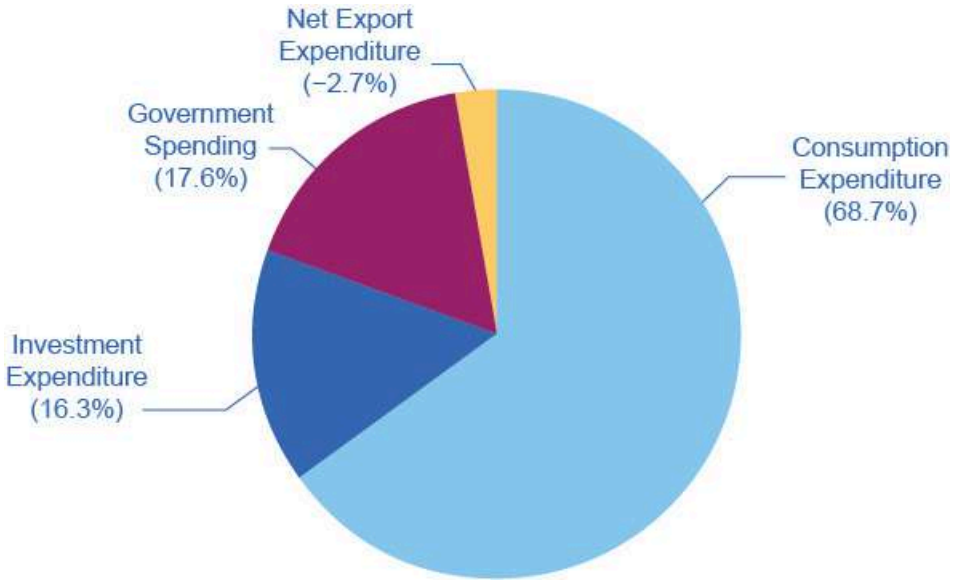


Figure 6.3 Percentage of Components of U.S. GDP on the Demand Side Consumption makes up over half of the demand side components of the GDP. (Source: http://bea.gov/iTable/index_nipa.cfm)

Clear It Up

What does the word “investment” mean?

What do economists mean by investment, or business spending? In calculating GDP, investment does not refer to purchasing stocks and bonds or trading financial assets. It refers to purchasing new capital goods, that is, new commercial real estate (such as buildings, factories, and stores) and equipment, residential housing construction, and inventories. Inventories that manufacturers produce this year are included in this year’s GDP—even if they are not yet sold. From the accountant’s perspective, it is as if the firm invested in its own inventories. Business investment in 2016 was $3 trillion, according to the Bureau of Economic Analysis.

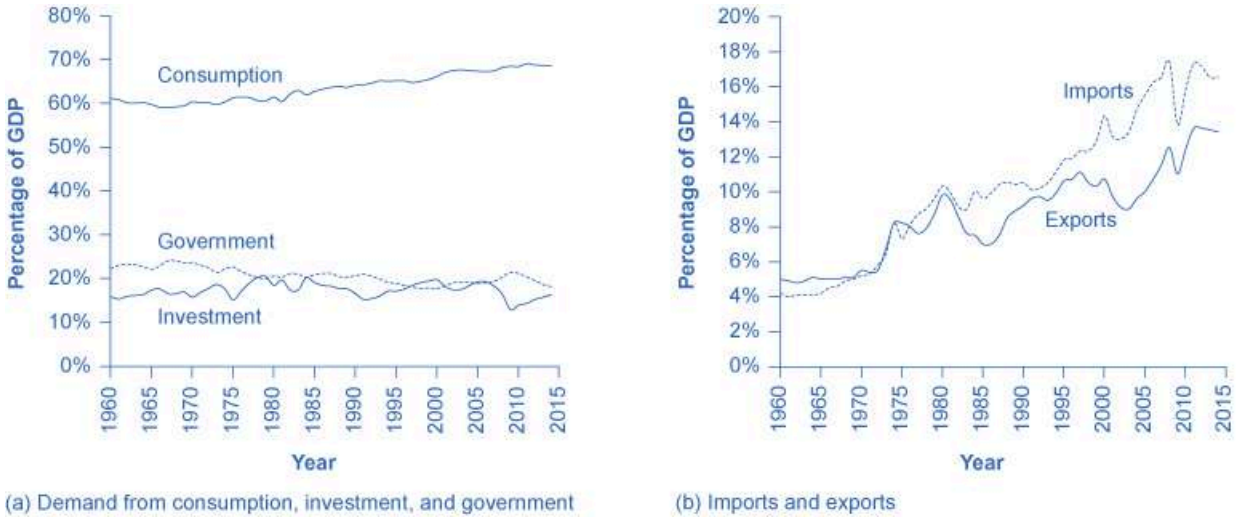


Figure 6.4 Components of GDP on the Demand Side (a) Consumption is about two-thirds of GDP, and it has been on a slight upward trend over time. Business investment hovers around 15% of GDP, but it fluctuates more than consumption. Government spending on goods and services is slightly under 20% of GDP and has declined modestly over time. (b) Exports are added to total demand for goods and services, while imports are subtracted from total demand. If exports exceed imports, as in most of the 1960s and 1970s in the U.S. economy, a trade surplus exists. If imports exceed exports, as in recent years, then a trade deficit exists. (Source: http://bea.gov/iTable/index_nipa.cfm)

Consumption expenditure by households is the largest component of GDP, accounting for about two-thirds of the GDP in any year. This tells us that consumers' spending decisions are a major driver of the economy. However, consumer spending is a gentle elephant: when viewed over time, it does not jump around too much, and has increased modestly from about 60% of GDP in the 1960s and 1970s.

Investment expenditure refers to purchases of physical plant and equipment, primarily by businesses. If Starbucks builds a new store, or Amazon buys robots, they count these expenditures under business investment. Investment demand is far smaller than consumption demand, typically accounting for only about 15–18% of GDP, but it is very important for the economy because this is where jobs are created. However, it fluctuates more noticeably than consumption. Business investment is volatile. New technology or a new product can spur business investment, but then confidence can drop and business investment can pull back sharply.

If you have noticed any of the infrastructure projects (new bridges, highways, airports) launched during the 2009 recession, you have seen how important government spending can be for the economy. Government expenditure in the United States is close to 20% of GDP, and includes spending by all three levels of government: federal, state, and local. The only part of government spending counted in demand is government purchases of goods or services produced in the economy. Examples include the government buying a new fighter jet for the Air Force (federal government spending), building a new highway (state government spending), or a new school (local government spending). A significant portion of government budgets consists of transfer payments, like unemployment benefits, veteran's benefits, and Social Security payments to retirees. The government excludes these payments from GDP because it does not receive a new good or service in return or exchange. Instead they are transfers of income from taxpayers to others. If you are curious about the awesome undertaking of adding up GDP, read the following Clear It Up feature.

Clear It Up

How do statisticians measure GDP?

Government economists at the Bureau of Economic Analysis (BEA), within the U.S. Department of Commerce, piece together estimates of GDP from a variety of sources.

Once every five years, in the second and seventh year of each decade, the Bureau of the Census carries

out a detailed census of businesses throughout the United States. In between, the Census Bureau carries out a monthly survey of retail sales. The government adjusts these figures with foreign trade data to account for exports that are produced in the United States and sold abroad and for imports that are produced abroad and sold here. Once every ten years, the Census Bureau conducts a comprehensive survey of housing and residential finance. Together, these sources provide the main basis for figuring out what is produced for consumers.

For investment, the Census Bureau carries out a monthly survey of construction and an annual survey of expenditures on physical capital equipment.

For what the federal government purchases, the statisticians rely on the U.S. Department of the Treasury. An annual Census of Governments gathers information on state and local governments. Because the government spends a considerable amount at all levels hiring people to provide services, it also tracks a large portion of spending through payroll records that state governments and the Social Security Administration collect.

With regard to foreign trade, the Census Bureau compiles a monthly record of all import and export documents. Additional surveys cover transportation and travel, and make adjustments for financial services that are produced in the United States for foreign customers.

Many other sources contribute to GDP estimates. Information on energy comes from the U.S. Department of Transportation and Department of Energy. The Agency for Health Care Research and Quality collects information on healthcare. Surveys of landlords find out about rental income. The Department of Agriculture collects statistics on farming.

All these bits and pieces of information arrive in different forms, at different time intervals. The BEA melds them together to produce GDP estimates on a quarterly basis (every three months). The BEA then "annualizes" these numbers by multiplying by four. As more information comes in, the BEA updates and revises these estimates. BEA releases the GDP “advance” estimate for a certain quarter one month after a quarter. The “preliminary” estimate comes out one month after that. The BEA publishes the “final” estimate one month later, but it is not actually final. In July, the BEA releases roughly updated estimates for the previous calendar year. Then, once every five years, after it has processed all the results of the latest detailed five-year business census, the BEA revises all of the past GDP estimates according to the newest methods and data, going all the way back to 1929.

Link It Up

Visit this website (http://openstaxcollege.org/l/beafaq) to read FAQs on the BEA site. You can even email your own questions!

When thinking about the demand for domestically produced goods in a global economy, it is important to count spending on exports—domestically produced goods that a country sells abroad. Similarly, we must also subtract spending on imports—goods that a country produces in other countries that residents of this country purchase. The GDP net export component is equal to the dollar value of exports (X) minus the dollar value of imports (M), (X – M). We call the gap between exports and imports the **trade balance**. If a country’s exports are larger than its imports, then a country has a **trade surplus**. In the United States, exports typically exceeded imports in the 1960s and 1970s, as **Figure 6.4**(b) shows.

Since the early 1980s, imports have typically exceeded exports, and so the United States has experienced a **trade**

deficit in most years. The trade deficit grew quite large in the late 1990s and in the mid-2000s. **Figure 6.4** (b) also shows that imports and exports have both risen substantially in recent decades, even after the declines during the Great Recession between 2008 and 2009. As we noted before, if exports and imports are equal, foreign trade has no effect on total GDP. However, even if exports and imports are balanced overall, foreign trade might still have powerful effects on particular industries and workers by causing nations to shift workers and physical capital investment toward one industry rather than another.

Based on these four components of demand, we can measure GDP as:

$$\text{GDP} = \text{Consumption} + \text{Investment} + \text{Government} + \text{Trade balance}$$
$$\text{GDP} = C + I + G + (X - M)$$

Understanding how to measure GDP is important for analyzing connections in the macro economy and for thinking about macroeconomic policy tools.

GDP Measured by What is Produced

Everything that we purchase somebody must first produce. **Table 6.2** breaks down what a country produces into five categories: **durable goods**, **nondurable goods**, **services**, **structures**, and the change in **inventories**. Before going into detail about these categories, notice that total GDP measured according to what is produced is exactly the same as the GDP measured by looking at the five components of demand. **Figure 6.5** provides a visual representation of this information.

	Components of GDP on the Supply Side (in trillions of dollars)	Percentage of Total
Goods		
Durable goods	$3.0	16.1%
Nondurable goods	$2.5	13.4%
Services	$11.6	62.4%
Structures	$1.5	8.1%
Change in inventories	$0.0	0.0%
Total GDP	**$18.6**	**100%**

Table 6.2 Components of U.S. GDP on the Production Side, 2016 (Source: http://bea.gov/iTable/index_nipa.cfm)

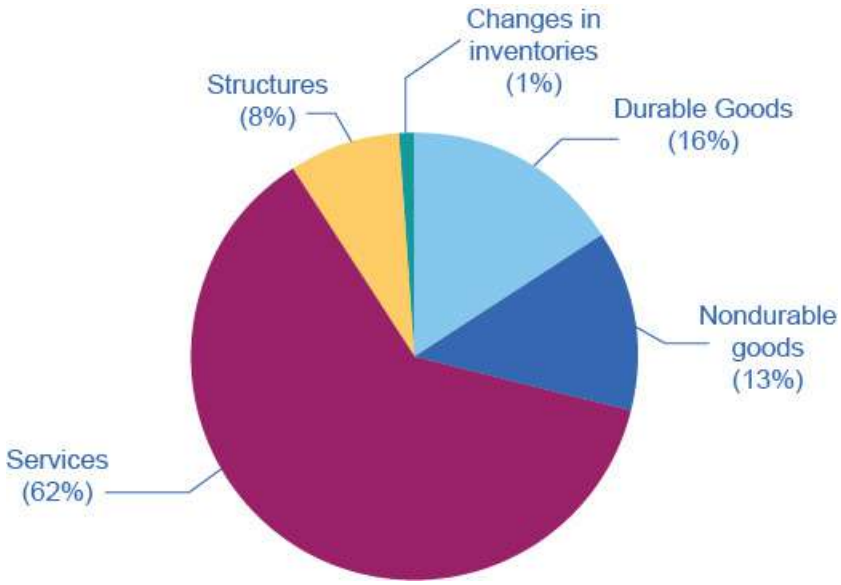


Figure 6.5 Percentage of Components of GDP on the Production Side Services make up over 60 percent of the production side components of GDP in the United States.

Since every market transaction must have both a buyer and a seller, GDP must be the same whether measured by what is demanded or by what is produced. **Figure 6.6** shows these components of what is produced, expressed as a percentage of GDP, since 1960.

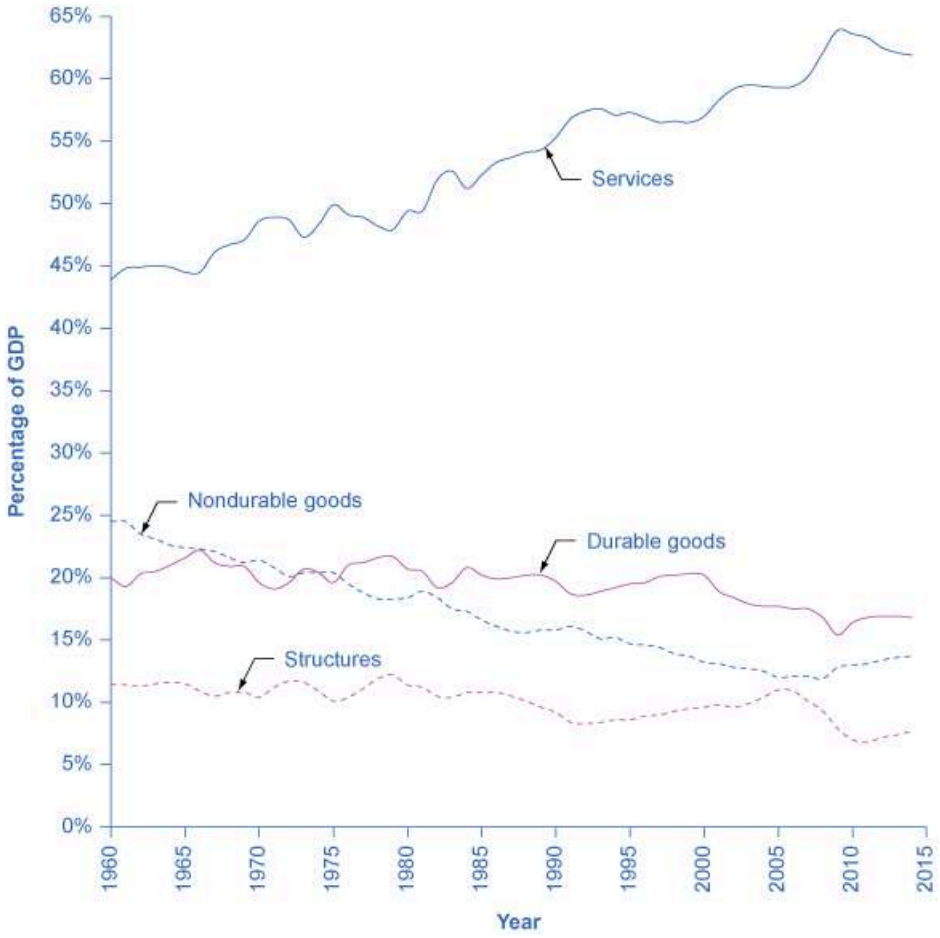


Figure 6.6 Types of Production Services are the largest single component of total supply, representing over 60 percent of GDP, up from about 45 perent in the early 1960s. Durable and nondurable goods constitute the manufacturing sector, and they have declined from 45 percent of GDP in 1960 to about 30 percent in 2016. Nondurable goods used to be larger than durable goods, but in recent years, nondurable goods have been dropping to below the share of durable goods, which is less than 20% of GDP. Structures hover around 10% of GDP. We do not show here the change in inventories, the final component of aggregate supply. It is typically less than 1% of GDP.

In thinking about what is produced in the economy, many non-economists immediately focus on solid, long-lasting goods, like cars and computers. By far the largest part of GDP, however, is services. Moreover, services have been a growing share of GDP over time. A detailed breakdown of the leading service industries would include healthcare, education, and legal and financial services. It has been decades since most of the U.S. economy involved making solid objects. Instead, the most common jobs in a modern economy involve a worker looking at pieces of paper or a computer screen; meeting with co-workers, customers, or suppliers; or making phone calls.

Even within the overall category of goods, long-lasting durable goods like cars and refrigerators are about the same share of the economy as short-lived nondurable goods like food and clothing. The category of structures includes everything from homes, to office buildings, shopping malls, and factories. Inventories is a small category that refers to the goods that one business has produced but has not yet sold to consumers, and are still sitting in warehouses and on shelves. The amount of inventories sitting on shelves tends to decline if business is better than expected, or to rise if business is worse than expected.

Another Way to Measure GDP: The National Income Approach

GDP is a measure of what is produced in a nation. The primary way GDP is estimated is with the Expenditure Approach we discussed above, but there is another way. Everything a firm produces, when sold, becomes revenues to the firm. Businesses use revenues to pay their bills: Wages and salaries for labor, interest and dividends for capital, rent for land, profit to the entrepreneur, etc. So adding up all the income produced in a year provides a second way of measuring GDP. This is why the terms GDP and **national income** are sometimes used interchangeably. The total

value of a nation's output is equal to the total value of a nation's income.

The Problem of Double Counting

We define GDP as the current value of all final goods and services produced in a nation in a year. What are final goods? They are goods at the furthest stage of production at the end of a year. Statisticians who calculate GDP must avoid the mistake of **double counting**, in which they count output more than once as it travels through the production stages. For example, imagine what would happen if government statisticians first counted the value of tires that a tire manufacturer produces, and then counted the value of a new truck that an automaker sold that contains those tires. In this example, the statisticians would have counted the value of the tires twice-because the truck's price includes the value of the tires.

To avoid this problem, which would overstate the size of the economy considerably, government statisticians count just the value of **final goods and services** in the chain of production that are sold for consumption, investment, government, and trade purposes. Statisticians exclude intermediate **intermediate goods**, which are goods that go into producing other goods, from GDP calculations. From the example above, they will only count the Ford truck's value. The value of what businesses provide to other businesses is captured in the final products at the end of the production chain.

The concept of GDP is fairly straightforward: it is just the dollar value of all final goods and services produced in the economy in a year. In our decentralized, market-oriented economy, actually calculating the more than $18 trillion-dollar U.S. GDP—along with how it is changing every few months—is a full-time job for a brigade of government statisticians.

What is Counted in GDP	What is not included in GDP
Consumption	Intermediate goods
Business investment	Transfer payments and non-market activities
Government spending on goods and services	Used goods
Net exports	Illegal goods

Table 6.3 Counting GDP

Notice the items that are not counted into GDP, as **Table 6.3** outlines. The sales of used goods are not included because they were produced in a previous year and are part of that year's GDP. The entire underground economy of services paid "under the table" and illegal sales should be counted, but is not, because it is impossible to track these sales. In Friedrich Schneider's recent study of shadow economies, he estimated the underground economy in the United States to be 6.6% of GDP, or close to $2 trillion dollars in 2013 alone. Transfer payments, such as payment by the government to individuals, are not included, because they do not represent production. Also, production of some goods—such as home production as when you make your breakfast—is not counted because these goods are not sold in the marketplace.

Link It Up

Visit this website (http://openstaxcollege.org/l/undergroundecon) to read about the "New Underground Economy."

Other Ways to Measure the Economy

Besides GDP, there are several different but closely related ways of measuring the size of the economy. We mentioned above that we can think of GDP as total production and as total purchases. We can also think of it as total income since anything one produces and sells yields income.

One of the closest cousins of GDP is the **gross national product (GNP)**. GDP includes only what country produces within its borders. GNP adds what domestic businesses and labor abroad produces, and subtracts any payments that foreign labor and businesses located in the United States send home to other countries. In other words, GNP is based more on what a country's citizens and firms produce, wherever they are located, and GDP is based on what happens within a certain county's geographic boundaries. For the United States, the gap between GDP and GNP is relatively small; in recent years, only about 0.2%. For small nations, which may have a substantial share of their population working abroad and sending money back home, the difference can be substantial.

We calculate **net national product (NNP)** by taking GNP and then subtracting the value of how much physical capital is worn out, or reduced in value because of aging, over the course of a year. The process by which capital ages and loses value is called **depreciation**. We can further subdivide NNP into **national income**, which includes all income to businesses and individuals, and personal income, which includes only income to people.

For practical purposes, it is not vital to memorize these definitions. However, it is important to be aware that these differences exist and to know what statistic you are examining, so that you do not accidentally compare, say, GDP in one year or for one country with GNP or NNP in another year or another country. To get an idea of how these calculations work, follow the steps in the following Work It Out feature.

Work It Out

Calculating GDP, Net Exports, and NNP

Based on the information in Table 6.4:

a. What is the value of GDP?
b. What is the value of net exports?
c. What is the value of NNP?

Government purchases	$120 billion
Depreciation	$40 billion
Consumption	$400 billion
Business Investment	$60 billion
Exports	$100 billion

Table 6.4

Imports	$120 billion
Income receipts from rest of the world	$10 billion
Income payments to rest of the world	$8 billion

Table 6.4

Step 1. To calculate GDP use the following formula:

$$\begin{aligned} \text{GDP} &= \text{Consumption} + \text{Investment} + \text{Government spending} + (\text{Exports} - \text{Imports}) \\ &= C + I + G + (X - M) \\ &= \$400 + \$60 + \$120 + (\$100 - \$120) \\ &= \$560 \text{ billion} \end{aligned}$$

Step 2. To calculate net exports, subtract imports from exports.

$$\begin{aligned} \text{Net exports} &= X - M \\ &= \$100 - \$120 \\ &= -\$20 \text{ billion} \end{aligned}$$

Step 3. To calculate NNP, use the following formula:

$$\begin{aligned} \text{NNP} &= \text{GDP} + \text{Income receipts from the rest of the world} \\ &\quad - \text{Income payments to the rest of the world} - \text{Depreciation} \\ &= \$560 + \$10 - \$8 - \$40 \\ &= \$522 \text{ billion} \end{aligned}$$

6.2 | Adjusting Nominal Values to Real Values

By the end of this section, you will be able to:

- Contrast nominal GDP and real GDP
- Explain GDP deflator
- Calculate real GDP based on nominal GDP values

When examining economic statistics, there is a crucial distinction worth emphasizing. The distinction is between nominal and real measurements, which refer to whether or not inflation has distorted a given statistic. Looking at economic statistics without considering inflation is like looking through a pair of binoculars and trying to guess how close something is: unless you know how strong the lenses are, you cannot guess the distance very accurately. Similarly, if you do not know the inflation rate, it is difficult to figure out if a rise in GDP is due mainly to a rise in the overall level of prices or to a rise in quantities of goods produced. The **nominal value** of any economic statistic means that we measure the statistic in terms of actual prices that exist at the time. The **real value** refers to the same statistic after it has been adjusted for inflation. Generally, it is the real value that is more important.

Converting Nominal to Real GDP

Table 6.5 shows U.S. GDP at five-year intervals since 1960 in nominal dollars; that is, GDP measured using the actual market prices prevailing in each stated year. Figure 6.7 also reflects this data in a graph.

Year	Nominal GDP (billions of dollars)	GDP Deflator (2005 = 100)
1960	543.3	19.0
1965	743.7	20.3
1970	1,075.9	24.8
1975	1,688.9	34.1
1980	2,862.5	48.3
1985	4,346.7	62.3
1990	5,979.6	72.7
1995	7,664.0	81.7
2000	10,289.7	89.0
2005	13,095.4	100.0
2010	14,958.3	110.0

Table 6.5 U.S. Nominal GDP and the GDP Deflator (Source: www.bea.gov)

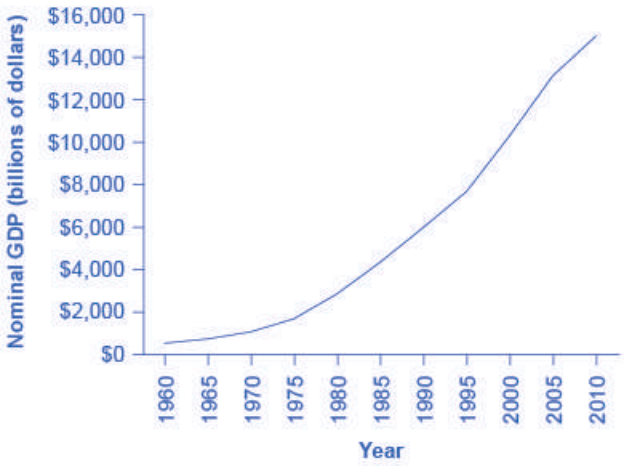


Figure 6.7 U.S. Nominal GDP, 1960–2010 Nominal GDP values have risen exponentially from 1960 through 2010, according to the BEA.

If an unwary analyst compared nominal GDP in 1960 to nominal GDP in 2010, it might appear that national output had risen by a factor of more than twenty-seven over this time (that is, GDP of $14,958 billion in 2010 divided by GDP of $543 billion in 1960 = 27.5). This conclusion would be highly misleading. Recall that we define nominal GDP as the quantity of every good or service produced multiplied by the price at which it was sold, summed up for all goods and services. In order to see how much production has actually increased, we need to extract the effects of higher prices on nominal GDP. We can easily accomplish this using the GDP deflator.

The GDP deflator is a price index measuring the average prices of all goods and services included in the economy. We explore price indices in detail and how we compute them in **Inflation**, but this definition will do in the context of this chapter. **Table 6.5** provides the GDP deflator data and **Figure 6.8** shows it graphically.

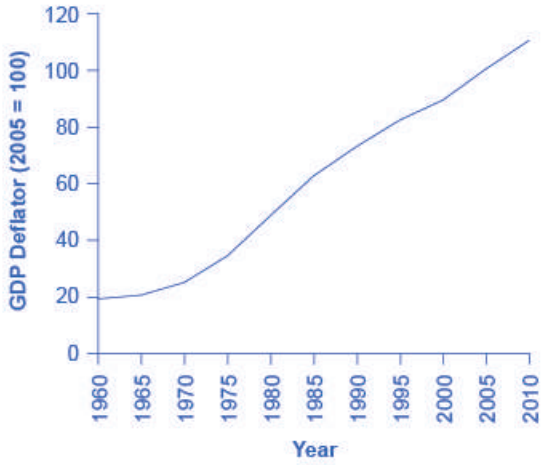


Figure 6.8 U.S. GDP Deflator, 1960–2010 Much like nominal GDP, the GDP deflator has risen exponentially from 1960 through 2010. (Source: BEA)

Figure 6.8 shows that the price level has risen dramatically since 1960. The price level in 2010 was almost six times higher than in 1960 (the deflator for 2010 was 110 versus a level of 19 in 1960). Clearly, much of the growth in nominal GDP was due to inflation, not an actual change in the quantity of goods and services produced, in other words, not in real GDP. Recall that nominal GDP can rise for two reasons: an increase in output, and/or an increase in prices. What is needed is to extract the increase in prices from nominal GDP so as to measure only changes in output. After all, the dollars used to measure nominal GDP in 1960 are worth more than the inflated dollars of 1990—and the price index tells exactly how much more. This adjustment is easy to do if you understand that nominal measurements are in value terms, where

$$\begin{aligned} \text{Value} &= \text{Price} \times \text{Quantity} \\ &\text{or} \\ \text{Nominal GDP} &= \text{GDP Deflator} \times \text{Real GDP} \end{aligned}$$

Let's look at an example at the micro level. Suppose the t-shirt company, Coolshirts, sells 10 t-shirts at a price of $9 each.

$$\begin{aligned} \text{Coolshirt's nominal revenue from sales} &= \text{Price} \times \text{Quantity} \\ &= \$9 \times 10 \\ &= \$90 \end{aligned}$$

Then,

$$\begin{aligned} \text{Coolshirt's real income} &= \frac{\text{Nominal revenue}}{\text{Price}} \\ &= \frac{\$90}{\$9} \\ &= 10 \end{aligned}$$

In other words, when we compute "real" measurements we are trying to obtain actual quantities, in this case, 10 t-shirts.

With GDP, it is just a tiny bit more complicated. We start with the same formula as above:

$$\text{Real GDP} = \frac{\text{Nominal GDP}}{\text{Price Index}}$$

For reasons that we will explain in more detail below, mathematically, a price index is a two-digit decimal number like 1.00 or 0.85 or 1.25. Because some people have trouble working with decimals, when the price index is published, it has traditionally been multiplied by 100 to get integer numbers like 100, 85, or 125. What this means is that when we "deflate" nominal figures to get real figures (by dividing the nominal by the price index). We also need to remember to divide the published price index by 100 to make the math work. Thus, the formula becomes:

$$\text{Real GDP} = \frac{\text{Nominal GDP}}{\text{Price Index} / 100}$$

Now read the following Work It Out feature for more practice calculating real GDP.

Work It Out

Computing GDP

It is possible to use the data in **Table 6.5** to compute real GDP.

Step 1. Look at **Table 6.5**, to see that, in 1960, nominal GDP was $543.3 billion and the price index (GDP deflator) was 19.0.

Step 2. To calculate the real GDP in 1960, use the formula:

$$\begin{aligned} \text{Real GDP} &= \frac{\text{Nominal GDP}}{\text{Price Index} / 100} \\ &= \frac{\$543.3 \text{ billion}}{19 / 100} \\ &= \$2{,}859.5 \text{ billion} \end{aligned}$$

We'll do this in two parts to make it clear. First adjust the price index: 19 divided by 100 = 0.19. Then divide into nominal GDP: $543.3 billion / 0.19 = $2,859.5 billion.

Step 3. Use the same formula to calculate the real GDP in 1965.

$$\begin{aligned} \text{Real GDP} &= \frac{\text{Nominal GDP}}{\text{Price Index} / 100} \\ &= \frac{\$743.7 \text{ billion}}{20.3 / 100} \\ &= \$3{,}663.5 \text{ billion} \end{aligned}$$

Step 4. Continue using this formula to calculate all of the real GDP values from 1960 through 2010. The calculations and the results are in **Table 6.6**.

Year	Nominal GDP (billions of dollars)	GDP Deflator (2005 = 100)	Calculations	Real GDP (billions of 2005 dollars)
1960	543.3	19.0	543.3 / (19.0/100)	2859.5
1965	743.7	20.3	743.7 / (20.3/100)	3663.5
1970	1075.9	24.8	1,075.9 / (24.8/100)	4338.3
1975	1688.9	34.1	1,688.9 / (34.1/100)	4952.8
1980	2862.5	48.3	2,862.5 / (48.3/100)	5926.5
1985	4346.7	62.3	4,346.7 / (62.3/100)	6977.0
1990	5979.6	72.7	5,979.6 / (72.7/100)	8225.0

Table 6.6 Converting Nominal to Real GDP (Source: Bureau of Economic Analysis, www.bea.gov)

Year	Nominal GDP (billions of dollars)	GDP Deflator (2005 = 100)	Calculations	Real GDP (billions of 2005 dollars)
1995	7664.0	82.0	7,664 / (82.0/100)	9346.3
2000	10289.7	89.0	10,289.7 / (89.0/100)	11561.5
2005	13095.4	100.0	13,095.4 / (100.0/100)	13095.4
2010	14958.3	110.0	14,958.3 / (110.0/100)	13598.5

Table 6.6 Converting Nominal to Real GDP (Source: Bureau of Economic Analysis, www.bea.gov)

There are a couple things to notice here. Whenever you compute a real statistic, one year (or period) plays a special role. It is called the base year (or base period). The base year is the year whose prices we use to compute the real statistic. When we calculate real GDP, for example, we take the quantities of goods and services produced in each year (for example, 1960 or 1973) and multiply them by their prices in the base year (in this case, 2005), so we get a measure of GDP that uses prices that do not change from year to year. That is why real GDP is labeled "Constant Dollars" or, in this example, "2005 Dollars," which means that real GDP is constructed using prices that existed in 2005. While the example here uses 2005 as the base year, more generally, you can use any year as the base year. The formula is:

$$\text{GDP deflator} = \frac{\text{Nominal GDP}}{\text{Real GDP}} \times 100$$

Rearranging the formula and using the data from 2005:

$$\begin{aligned}\text{Real GDP} &= \frac{\text{Nominal GDP}}{\text{Price Index} / 100} \\ &= \frac{\$13{,}095.4 \text{ billion}}{100 / 100} \\ &= \$13{,}095.4 \text{ billion}\end{aligned}$$

Comparing real GDP and nominal GDP for 2005, you see they are the same. This is no accident. It is because we have chosen 2005 as the "base year" in this example. Since the price index in the base year always has a value of 100 (by definition), nominal and real GDP are always the same in the base year.

Look at the data for 2010.

$$\begin{aligned}\text{Real GDP} &= \frac{\text{Nominal GDP}}{\text{Price Index} / 100} \\ &= \frac{\$14{,}958.3 \text{ billion}}{110 / 100} \\ &= \$13{,}598.5 \text{ billion}\end{aligned}$$

Use this data to make another observation: As long as inflation is positive, meaning prices increase on average from year to year, real GDP should be less than nominal GDP in any year after the base year. The reason for this should be clear: The value of nominal GDP is "inflated" by inflation. Similarly, as long as inflation is positive, real GDP should be greater than nominal GDP in any year before the base year.

Figure 6.9 shows the U.S. nominal and real GDP since 1960. Because 2005 is the base year, the nominal and real values are exactly the same in that year. However, over time, the rise in nominal GDP looks much larger than the rise in real GDP (that is, the nominal GDP line rises more steeply than the real GDP line), because the presence of

inflation, especially in the 1970s exaggerates the rise in nominal GDP.

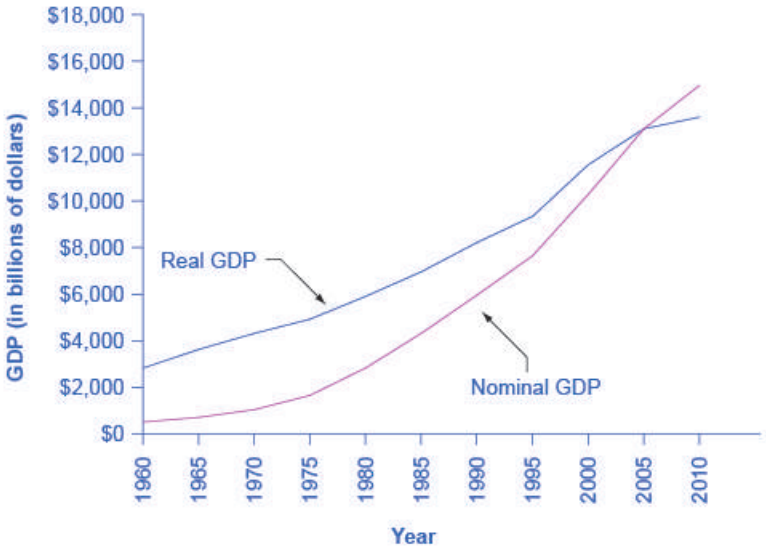


Figure 6.9 U.S. Nominal and Real GDP, 1960–2012 The red line measures U.S. GDP in nominal dollars. The black line measures U.S. GDP in real dollars, where all dollar values are converted to 2005 dollars. Since we express real GDP in 2005 dollars, the two lines cross in 2005. However, real GDP will appear higher than nominal GDP in the years before 2005, because dollars were worth less in 2005 than in previous years. Conversely, real GDP will appear lower in the years after 2005, because dollars were worth more in 2005 than in later years.

Let's return to the question that we posed originally: How much did GDP increase in real terms? What was the real GDP growth rate from 1960 to 2010? To find the real growth rate, we apply the formula for percentage change:

$$\frac{\text{2010 real GDP} - \text{1960 real GDP}}{\text{1960 real GDP}} \times 100 = \%\ \text{change}$$

$$\frac{13{,}598.5 - 2{,}859.5}{2{,}859.5} \times 100 = 376\%$$

In other words, the U.S. economy has increased real production of goods and services by nearly a factor of four since 1960. Of course, that understates the material improvement since it fails to capture improvements in the quality of products and the invention of new products.

There is a quicker way to answer this question approximately, using another math trick. Because:

$$\text{Nominal} = \text{Price} \times \text{Quantity}$$

$$\%\ \text{change in Nominal} = \%\ \text{change in Price} + \%\ \text{change in Quantity}$$

OR

$$\%\ \text{change in Quantity} = \%\ \text{change in Nominal} - \%\ \text{change in Price}$$

Therefore, real GDP growth rate (% change in quantity) equals the growth rate in nominal GDP (% change in value) minus the inflation rate (% change in price).

Note that using this equation provides an approximation for small changes in the levels. For more accurate measures, one should use the first formula.

6.3 | Tracking Real GDP over Time

By the end of this section, you will be able to:

- Explain recessions, depressions, peaks, and troughs
- Evaluate the importance of tracking real GDP over time

When news reports indicate that "the economy grew 1.2% in the first quarter," the reports are referring to the percentage change in real GDP. By convention, governments report GDP growth is at an annualized rate: Whatever the calculated growth in real GDP was for the quarter, we multiply it by four when it is reported as if the economy were growing at that rate for a full year.

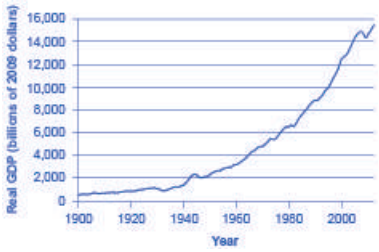


Figure 6.10 U.S. GDP, 1900–2016 Real GDP in the United States in 2016 (in 2009 dollars) was about $16.7 trillion. After adjusting to remove the effects of inflation, this represents a roughly 20-fold increase in the economy's production of goods and services since the start of the twentieth century. (Source: bea.gov)

Figure 6.10 shows the pattern of U.S. real GDP since 1900. Short term declines have regularly interrupted the generally upward long-term path of GDP. We call a significant decline in real GDP a **recession**. We call an especially lengthy and deep recession a **depression**. The severe drop in GDP that occurred during the 1930s Great Depression is clearly visible in the figure, as is the 2008–2009 Great Recession.

Real GDP is important because it is highly correlated with other measures of economic activity, like employment and unemployment. When real GDP rises, so does employment.

The most significant human problem associated with recessions (and their larger, uglier cousins, depressions) is that a slowdown in production means that firms need to lay off or fire some of their workers. Losing a job imposes painful financial and personal costs on workers, and often on their extended families as well. In addition, even those who keep their jobs are likely to find that wage raises are scanty at best—or their employers may ask them to take pay cuts.

Table 6.7 lists the pattern of recessions and expansions in the U.S. economy since 1900. We call the highest point of the economy, before the recession begins, the **peak**. Conversely, the lowest point of a recession, before a recovery begins, is the **trough**. Thus, a recession lasts from peak to trough, and an economic upswing runs from trough to peak. We call the economy's movement from peak to trough and trough to peak the **business cycle**. It is intriguing to notice that the three longest trough-to-peak expansions of the twentieth century have happened since 1960. The most recent recession started in December 2007 and ended formally in June 2009. This was the most severe recession since the 1930s Great Depression. The ongoing expansion since the June 2009 trough will also be quite long, comparatively, having already reached 90 months at the end of 2016.

Trough	Peak	Months of Contraction	Months of Expansion
December 1900	September 1902	18	21
August 1904	May 1907	23	33
June 1908	January 1910	13	19
January 1912	January 1913	24	12
December 1914	August 1918	23	44
March 1919	January 1920	7	10
July 1921	May 1923	18	22
July 1924	October 1926	14	27

Table 6.7 U.S. Business Cycles since 1900 (Source: http://www.nber.org/cycles/main.html)

Trough	Peak	Months of Contraction	Months of Expansion
November 1927	August 1929	23	21
March 1933	May 1937	43	50
June 1938	February 1945	13	80
October 1945	November 1948	8	37
October 1949	July 1953	11	45
May 1954	August 1957	10	39
April 1958	April 1960	8	24
February 1961	December 1969	10	106
November 1970	November 1973	11	36
March 1975	January 1980	16	58
July 1980	July 1981	6	12
November 1982	July 1990	16	92
March 1991	March 2001	8	120
November 2001	December 2007	8	73

Table 6.7 U.S. Business Cycles since 1900 (Source: http://www.nber.org/cycles/main.html)

A private think tank, the National Bureau of Economic Research (NBER), tracks business cycles for the U.S. economy. However, the effects of a severe recession often linger after the official ending date assigned by the NBER.

6.4 | Comparing GDP among Countries

By the end of this section, you will be able to:

- Explain how we can use GDP to compare the economic welfare of different nations
- Calculate the conversion of GDP to a common currency by using exchange rates
- Calculate GDP per capita using population data

It is common to use GDP as a measure of economic welfare or standard of living in a nation. When comparing the GDP of different nations for this purpose, two issues immediately arise. First, we measure a country's GDP in its own currency: the United States uses the U.S. dollar; Canada, the Canadian dollar; most countries of Western Europe, the euro; Japan, the yen; Mexico, the peso; and so on. Thus, comparing GDP between two countries requires converting to a common currency. A second issue is that countries have very different numbers of people. For instance, the United States has a much larger economy than Mexico or Canada, but it also has almost three times as many people as Mexico and nine times as many people as Canada. Thus, if we are trying to compare standards of living across countries, we need to divide GDP by population.

Converting Currencies with Exchange Rates

To compare the GDP of countries with different currencies, it is necessary to convert to a “common denominator” using an **exchange rate**, which is the value of one currency in terms of another currency. We express exchange rates either as the units of country A’s currency that need to be traded for a single unit of country B’s currency (for example, Japanese yen per British pound), or as the inverse (for example, British pounds per Japanese yen). We can use two types of exchange rates for this purpose, market exchange rates and purchasing power parity (PPP)

equivalent exchange rates. Market exchange rates vary on a day-to-day basis depending on supply and demand in foreign exchange markets. PPP-equivalent exchange rates provide a longer run measure of the exchange rate. For this reason, economists typically use PPP-equivalent exchange rates for GDP cross country comparisons. We will discuss exchange rates in more detail in **Exchange Rates and International Capital Flows**. The following Work It Out feature explains how to convert GDP to a common currency.

Work It Out

Converting GDP to a Common Currency

Using the exchange rate to convert GDP from one currency to another is straightforward. Say that the task is to compare Brazil's GDP in 2013 of 4.8 trillion reals with the U.S. GDP of $16.6 trillion for the same year.

Step 1. Determine the exchange rate for the specified year. In 2013, the exchange rate was 2.230 reals = $1. (These numbers are realistic, but rounded off to simplify the calculations.)

Step 2. Convert Brazil's GDP into U.S. dollars:

$$\begin{aligned}\text{Brazil's GDP in \$ U.S.} &= \frac{\text{Brazil's GDP in reals}}{\text{Exchange rate (reals/\$ U.S.)}} \\ &= \frac{\text{4.8 trillion reals}}{\text{2.230 reals per \$ U.S.}} \\ &= \text{\$2.2 trillion}\end{aligned}$$

Step 3. Compare this value to the GDP in the United States in the same year. The U.S. GDP was $16.6 trillion in 2013, which is nearly eight times that of GDP in Brazil in 2012.

Step 4. View **Table 6.8** which shows the size of and variety of GDPs of different countries in 2013, all expressed in U.S. dollars. We calculate each using the process that we explained above.

Country	GDP in Billions of Domestic Currency		Domestic Currency/U.S. Dollars (PPP Equivalent)	GDP (in billions of U.S. dollars)
Brazil	4,844.80	reals	2.157	2,246.00
Canada	1,881.20	dollars	1.030	1,826.80
China	58,667.30	yuan	6.196	9,469.10
Egypt	1,753.30	pounds	6.460	271.40
Germany	2,737.60	euros	0.753	3,636.00
India	113,550.70	rupees	60.502	1,876.80
Japan	478,075.30	yen	97.596	4,898.50
Mexico	16,104.40	pesos	12.772	1,260.90
South Korea	1,428,294.70	won	1,094.925	1,304.467
United Kingdom	1,612.80	pounds	0.639	2,523.20

Table 6.8 Comparing GDPs Across Countries, 2013 (Source: http://www.imf.org/external/pubs/ft/weo/2013/01/weodata/index.aspx)

Country	GDP in Billions of Domestic Currency		Domestic Currency/U.S. Dollars (PPP Equivalent)	GDP (in billions of U.S. dollars)
United States	16,768.10	dollars	1.000	16,768.10

Table 6.8 Comparing GDPs Across Countries, 2013 (Source: http://www.imf.org/external/pubs/ft/weo/2013/01/weodata/index.aspx)

GDP Per Capita

The U.S. economy has the largest GDP in the world, by a considerable amount. The United States is also a populous country; in fact, it is the third largest country by population in the world, although well behind China and India. Is the U.S. economy larger than other countries just because the United States has more people than most other countries, or because the U.S. economy is actually larger on a per-person basis? We can answer this question by calculating a country's **GDP per capita**; that is, the GDP divided by the population.

$$\text{GDP per capita} = \text{GDP/population}$$

The second column of **Table 6.9** lists the GDP of the same selection of countries that appeared in the previous **Tracking Real GDP over Time** and **Table 6.8**, showing their GDP as converted into U.S. dollars (which is the same as the last column of the previous table). The third column gives the population for each country. The fourth column lists the GDP per capita. We obtain GDP per capita in two steps: First, by multiplying column two (GDP, in billions of dollars) by 1000 so it has the same units as column three (Population, in millions). Then divide the result (GDP in millions of dollars) by column three (Population, in millions).

Country	GDP (in billions of U.S. dollars)	Population (in millions)	Per Capita GDP (in U.S. dollars)
Brazil	2,246.00	199.20	11,275.10
Canada	1,826.80	35.10	52,045.58
China	9,469.10	1,360.80	6,958.48
Egypt	271.40	83.70	3,242.90
Germany	3,636.00	80.80	44,999.50
India	1,876.80	1,243.30	1,509.50
Japan	4,898.50	127.3	38,479.97
Mexico	1,260.90	118.40	10,649.90
South Korea	1,304.47	50.20	25,985.46
United Kingdom	2,523.20	64.10	39,363.50
United States	16,768.10	316.30	53,013.28

Table 6.9 GDP Per Capita, 2013 (Source: http://www.imf.org/external/pubs/ft/weo/2013/01/weodata/index.aspx)

Notice that the ranking by GDP is different from the ranking by GDP per capita. India has a somewhat larger GDP than Germany, but on a per capita basis, Germany has more than 10 times India's standard of living. Will China soon have a better standard of living than the U.S.? Read the following Clear It Up feature to find out.

Clear It Up

Is China going to surpass the United States in terms of standard of living?

As **Table 6.9** shows, China has the second largest GDP of the countries: $9.5 trillion compared to the United States' $16.8 trillion. Perhaps it will surpass the United States, but probably not any time soon. China has a much larger population so that in per capita terms, its GDP is less than one fifth that of the United States ($6,958.48 compared to $53,013). The Chinese people are still quite poor relative to the United States and other developed countries. One caveat: For reasons we will discuss shortly, GDP per capita can give us only a rough idea of the differences in living standards across countries.

The world's high-income nations—including the United States, Canada, the Western European countries, and Japan—typically have GDP per capita in the range of $20,000 to $50,000. Middle-income countries, which include much of Latin America, Eastern Europe, and some countries in East Asia, have GDP per capita in the range of $6,000 to $12,000. The world's low-income countries, many of them located in Africa and Asia, often have GDP per capita of less than $2,000 per year.

6.5 | How Well GDP Measures the Well-Being of Society

By the end of this section, you will be able to:

- Discuss how productivity influences the standard of living
- Explain the limitations of GDP as a measure of the standard of living
- Analyze the relationship between GDP data and fluctuations in the standard of living

The level of GDP per capita clearly captures some of what we mean by the phrase "standard of living." Most of the migration in the world, for example, involves people who are moving from countries with relatively low GDP per capita to countries with relatively high GDP per capita.

"Standard of living" is a broader term than GDP. While GDP focuses on production that is bought and sold in markets, **standard of living** includes all elements that affect people's well-being, whether they are bought and sold in the market or not. To illuminate the difference between GDP and standard of living, it is useful to spell out some things that GDP does not cover that are clearly relevant to standard of living.

Limitations of GDP as a Measure of the Standard of Living

While GDP includes spending on recreation and travel, it does not cover leisure time. Clearly, however, there is a substantial difference between an economy that is large because people work long hours, and an economy that is just as large because people are more productive with their time so they do not have to work as many hours. The GDP per capita of the U.S. economy is larger than the GDP per capita of Germany, as **Table 6.9** showed, but does that prove that the standard of living in the United States is higher? Not necessarily, since it is also true that the average U.S. worker works several hundred hours more per year more than the average German worker. Calculating GDP does not account for the German worker's extra vacation weeks.

While GDP includes what a country spends on environmental protection, healthcare, and education, it does not include actual levels of environmental cleanliness, health, and learning. GDP includes the cost of buying pollution-control equipment, but it does not address whether the air and water are actually cleaner or dirtier. GDP includes spending on medical care, but does not address whether life expectancy or infant mortality have risen or fallen. Similarly, it counts spending on education, but does not address directly how much of the population can read, write,

or do basic mathematics.

GDP includes production that is exchanged in the market, but it does not cover production that is not exchanged in the market. For example, hiring someone to mow your lawn or clean your house is part of GDP, but doing these tasks yourself is not part of GDP. One remarkable change in the U.S. economy in recent decades is the growth in women's participation in the labor force. As of 1970, only about 42% of women participated in the paid labor force. By the second decade of the 2000s, nearly 60% of women participated in the paid labor force according to the Bureau of Labor Statistics. As women are now in the labor force, many of the services they used to produce in the non-market economy like food preparation and child care have shifted to some extent into the market economy, which makes the GDP appear larger even if people actually are not consuming more services.

GDP has nothing to say about the level of inequality in society. GDP per capita is only an average. When GDP per capita rises by 5%, it could mean that GDP for everyone in the society has risen by 5%, or that GDP of some groups has risen by more while that of others has risen by less—or even declined. GDP also has nothing in particular to say about the amount of variety available. If a family buys 100 loaves of bread in a year, GDP does not care whether they are all white bread, or whether the family can choose from wheat, rye, pumpernickel, and many others—it just looks at the total amount the family spends on bread.

Likewise, GDP has nothing much to say about what technology and products are available. The standard of living in, for example, 1950 or 1900 was not affected only by how much money people had—it was also affected by what they could buy. No matter how much money you had in 1950, you could not buy an iPhone or a personal computer.

In certain cases, it is not clear that a rise in GDP is even a good thing. If a city is wrecked by a hurricane, and then experiences a surge of rebuilding construction activity, it would be peculiar to claim that the hurricane was therefore economically beneficial. If people are led by a rising fear of crime, to pay for installing bars and burglar alarms on all their windows, it is hard to believe that this increase in GDP has made them better off. Similarly, some people would argue that sales of certain goods, like pornography or extremely violent movies, do not represent a gain to society's standard of living.

Does a Rise in GDP Overstate or Understate the Rise in the Standard of Living?

The fact that GDP per capita does not fully capture the broader idea of standard of living has led to a concern that the increases in GDP over time are illusory. It is theoretically possible that while GDP is rising, the standard of living could be falling if human health, environmental cleanliness, and other factors that are not included in GDP are worsening. Fortunately, this fear appears to be overstated.

In some ways, the rise in GDP understates the actual rise in the standard of living. For example, the typical workweek for a U.S. worker has fallen over the last century from about 60 hours per week to less than 40 hours per week. Life expectancy and health have risen dramatically, and so has the average level of education. Since 1970, the air and water in the United States have generally been getting cleaner. Companies have developed new technologies for entertainment, travel, information, and health. A much wider variety of basic products like food and clothing is available today than several decades ago. Because GDP does not capture leisure, health, a cleaner environment, the possibilities that new technology creates, or an increase in variety, the actual rise in the standard of living for Americans in recent decades has exceeded the rise in GDP.

On the other side, crime rates, traffic congestion levels, and income inequality are higher in the United States now than they were in the 1960s. Moreover, a substantial number of services that women primarily provided in the non-market economy are now part of the market economy that GDP counts. By ignoring these factors, GDP would tend to overstate the true rise in the standard of living.

Link It Up

Visit this website (http://openstaxcollege.org/l/amdreamvalue) to read about the American Dream and standards of living.

GDP is Rough, but Useful

A high level of GDP should not be the only goal of macroeconomic policy, or government policy more broadly. Even though GDP does not measure the broader standard of living with any precision, it does measure production well and it does indicate when a country is materially better or worse off in terms of jobs and incomes. In most countries, a significantly higher GDP per capita occurs hand in hand with other improvements in everyday life along many dimensions, like education, health, and environmental protection.

No single number can capture all the elements of a term as broad as "standard of living." Nonetheless, GDP per capita is a reasonable, rough-and-ready measure of the standard of living.

Bring it Home

How is the Economy Doing? How Does One Tell?

To determine the state of the economy, one needs to examine economic indicators, such as GDP. To calculate GDP is quite an undertaking. It is the broadest measure of a nation's economic activity and we owe a debt to Simon Kuznets, the creator of the measurement, for that.

The sheer size of the U.S. economy as measured by GDP is huge—as of the fourth quarter of 2016, $18.9 trillion worth of goods and services were produced annually. Real GDP informed us that the 2008–2009 recession was severe and that the recovery from that recession has been slow, but the economy is improving. GDP per capita gives a rough estimate of a nation's standard of living. This chapter is the building block for other chapters that explore more economic indicators such as unemployment, inflation, or interest rates, and perhaps more importantly, will explain how they are related and what causes them to rise or fall.

KEY TERMS

business cycle the economy's relatively short-term movement in and out of recession

depreciation the process by which capital ages over time and therefore loses its value

depression an especially lengthy and deep decline in output

double counting a potential mistake to avoid in measuring GDP, in which output is counted more than once as it travels through the stages of production

durable good long-lasting good like a car or a refrigerator

exchange rate the price of one currency in terms of another currency

final good and service output used directly for consumption, investment, government, and trade purposes; contrast with “intermediate good”

GDP per capita GDP divided by the population

gross domestic product (GDP) the value of the output of all goods and services produced within a country in a year

gross national product (GNP) includes what is produced domestically and what is produced by domestic labor and business abroad in a year

intermediate good output provided to other businesses at an intermediate stage of production, not for final users; contrast with “final good and service”

inventory good that has been produced, but not yet been sold

national income includes all income earned: wages, profits, rent, and profit income

net national product (NNP) GDP minus depreciation

nominal value the economic statistic actually announced at that time, not adjusted for inflation; contrast with real value

nondurable good short-lived good like food and clothing

peak during the business cycle, the highest point of output before a recession begins

real value an economic statistic after it has been adjusted for inflation; contrast with nominal value

recession a significant decline in national output

service product which is intangible (in contrast to goods) such as entertainment, healthcare, or education

standard of living all elements that affect people’s happiness, whether people buy or sell these elements in the market or not

structure building used as residence, factory, office building, retail store, or for other purposes

trade balance gap between exports and imports

trade deficit exists when a nation's imports exceed its exports and it calculates them as imports –exports

trade surplus exists when a nation's exports exceed its imports and it calculates them as exports – imports

trough during the business cycle, the lowest point of output in a recession, before a recovery begins

KEY CONCEPTS AND SUMMARY

6.1 Measuring the Size of the Economy: Gross Domestic Product

Economists generally express the size of a nation's economy as its gross domestic product (GDP), which measures the value of the output of all goods and services produced within the country in a year. Economists measure GDP by taking the quantities of all goods and services produced, multiplying them by their prices, and summing the total. Since GDP measures what is bought and sold in the economy, we can measure it either by the sum of what is purchased in the economy or what is produced.

We can divide demand into consumption, investment, government, exports, and imports. We can divide what is produced in the economy into durable goods, nondurable goods, services, structures, and inventories. To avoid double counting, GDP counts only final output of goods and services, not the production of intermediate goods or the value of labor in the chain of production.

6.2 Adjusting Nominal Values to Real Values

The nominal value of an economic statistic is the commonly announced value. The real value is the value after adjusting for changes in inflation. To convert nominal economic data from several different years into real, inflation-adjusted data, the starting point is to choose a base year arbitrarily and then use a price index to convert the measurements so that economists measure them in the money prevailing in the base year.

6.3 Tracking Real GDP over Time

Over the long term, U.S. real GDP have increased dramatically. At the same time, GDP has not increased the same amount each year. The speeding up and slowing down of GDP growth represents the business cycle. When GDP declines significantly, a recession occurs. A longer and deeper decline is a depression. Recessions begin at the business cycle's peak and end at the trough.

6.4 Comparing GDP among Countries

Since we measure GDP in a country's currency, in order to compare different countries' GDPs, we need to convert them to a common currency. One way to do that is with the exchange rate, which is the price of one country's currency in terms of another. Once we express GDPs in a common currency, we can compare each country's GDP per capita by dividing GDP by population. Countries with large populations often have large GDPs, but GDP alone can be a misleading indicator of a nation's wealth. A better measure is GDP per capita.

6.5 How Well GDP Measures the Well-Being of Society

GDP is an indicator of a society's standard of living, but it is only a rough indicator. GDP does not directly take account of leisure, environmental quality, levels of health and education, activities conducted outside the market, changes in inequality of income, increases in variety, increases in technology, or the (positive or negative) value that society may place on certain types of output.

SELF-CHECK QUESTIONS

1. Country A has export sales of $20 billion, government purchases of $1,000 billion, business investment is $50 billion, imports are $40 billion, and consumption spending is $2,000 billion. What is the dollar value of GDP?

2. Which of the following are included in GDP, and which are not?

a. The cost of hospital stays
b. The rise in life expectancy over time
c. Child care provided by a licensed day care center
d. Child care provided by a grandmother
e. A used car sale
f. A new car sale
g. The greater variety of cheese available in supermarkets
h. The iron that goes into the steel that goes into a refrigerator bought by a consumer.

3. Using data from **Table 6.5** how much of the nominal GDP growth from 1980 to 1990 was real GDP and how much was inflation?

4. Without looking at **Table 6.7**, return to **Figure 6.10**. If we define a recession as a significant decline in national output, can you identify any post-1960 recessions in addition to the 2008-2009 recession? (This requires a judgment call.)

5. According to **Table 6.7**, how often have recessions occurred since the end of World War II (1945)?

6. According to **Table 6.7**, how long has the average recession lasted since the end of World War II?

7. According to **Table 6.7**, how long has the average expansion lasted since the end of World War II?

8. Is it possible for GDP to rise while at the same time per capita GDP is falling? Is it possible for GDP to fall while per capita GDP is rising?

9. The Central African Republic has a GDP of 1,107,689 million CFA francs and a population of 4.862 million. The exchange rate is 284.681CFA francs per dollar. Calculate the GDP per capita of Central African Republic.

10. Explain briefly whether each of the following would cause GDP to overstate or understate the degree of change in the broad standard of living.

a. The environment becomes dirtier
b. The crime rate declines
c. A greater variety of goods become available to consumers
d. Infant mortality declines

REVIEW QUESTIONS

11. What are the main components of measuring GDP with what is demanded?

12. What are the main components of measuring GDP with what is produced?

13. Would you usually expect GDP as measured by what is demanded to be greater than GDP measured by what is supplied, or the reverse?

14. Why must you avoid double counting when measuring GDP?

15. What is the difference between a series of economic data over time measured in nominal terms versus the same data series over time measured in real terms?

16. How do you convert a series of nominal economic data over time to real terms?

17. What are typical GDP patterns for a high-income economy like the United States in the long run and the short run?

18. What are the two main difficulties that arise in comparing different countries's GDP?

19. List some of the reasons why economists should not consider GDP an effective measure of the standard of living in a country.

CRITICAL THINKING QUESTIONS

20. U.S. macroeconomic data are among the best in the world. Given what you learned in the **Clear It Up** "How do statisticians measure GDP?", does this surprise you, or does this simply reflect the complexity of a modern economy?

21. What does GDP not tell us about the economy?

22. Should people typically pay more attention to their real income or their nominal income? If you choose the latter, why would that make sense in today's world? Would your answer be the same for the 1970s?

23. Why do you suppose that U.S. GDP is so much higher today than 50 or 100 years ago?

24. Why do you think that GDP does not grow at a steady rate, but rather speeds up and slows down?

25. Cross country comparisons of GDP per capita typically use purchasing power parity equivalent exchange rates, which are a measure of the long run equilibrium value of an exchange rate. In fact, we used PPP equivalent exchange rates in this module. Why could using market exchange rates, which sometimes change dramatically in a short period of time, be misleading?

26. Why might per capita GDP be only an imperfect measure of a country's standard of living?

27. How might you measure a "green" GDP?

PROBLEMS

28. Last year, a small nation with abundant forests cut down $200 worth of trees. It then turned $100 worth of trees into $150 worth of lumber. It used $100 worth of that lumber to produce $250 worth of bookshelves. Assuming the country produces no other outputs, and there are no other inputs used in producing trees, lumber, and bookshelves, what is this nation's GDP? In other words, what is the value of the final goods the nation produced including trees, lumber and bookshelves?

29. The "prime" interest rate is the rate that banks charge their best customers. Based on the nominal interest rates and inflation rates in **Table 6.10**, in which of the years would it have been best to be a lender? Based on the nominal interest rates and inflation rates in **Table 6.10**, in which of the years given would it have been best to be a borrower?

Year	Prime Interest Rate	Inflation Rate
1970	7.9%	5.7%
1974	10.8%	11.0%
1978	9.1%	7.6%
1981	18.9%	10.3%

Table 6.10

30. A mortgage loan is a loan that a person makes to purchase a house. **Table 6.11** provides a list of the mortgage interest rate for several different years and the rate of inflation for each of those years. In which years would it have been better to be a person borrowing money from a bank to buy a home? In which years would it have been better to be a bank lending money?

Year	Mortgage Interest Rate	Inflation Rate
1984	12.4%	4.3%
1990	10%	5.4%
2001	7.0%	2.8%

Table 6.11

31. Ethiopia has a GDP of $8 billion (measured in U.S. dollars) and a population of 55 million. Costa Rica has a GDP of $9 billion (measured in U.S. dollars) and a population of 4 million. Calculate the per capita GDP for each country and identify which one is higher.

32. In 1980, Denmark had a GDP of $70 billion (measured in U.S. dollars) and a population of 5.1 million. In 2000, Denmark had a GDP of $160 billion (measured in U.S. dollars) and a population of 5.3 million. By what percentage did Denmark's GDP per capita rise between 1980 and 2000?

33. The Czech Republic has a GDP of 1,800 billion koruny. The exchange rate is 25 koruny/U.S. dollar. The Czech population is 20 million. What is the GDP per capita of the Czech Republic expressed in U.S. dollars?

7 | Economic Growth

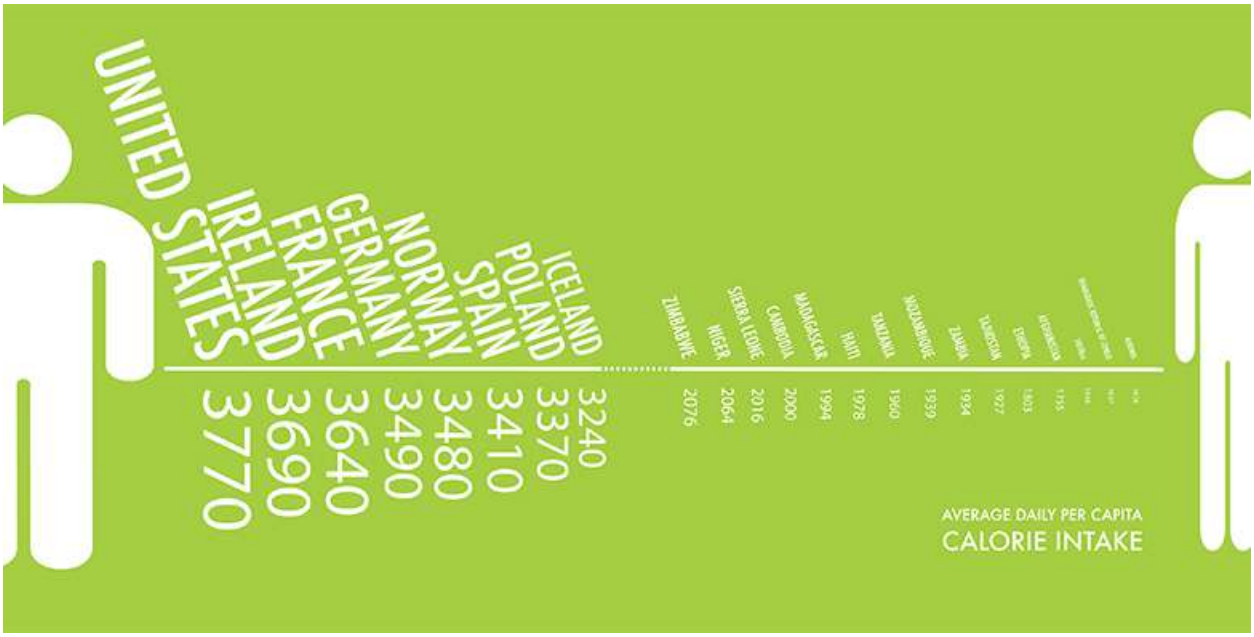


Figure 7.1 Average Daily Calorie Consumption Not only has the number of calories that people consume per day increased, so has the amount of food calories that people are able to afford based on their working wages. (Credit: modification of work by Lauren Manning/Flickr Creative Commons)

Bring it Home

Calories and Economic Growth

On average, humans need about 2,500 calories a day to survive, depending on height, weight, and gender. The economist Brad DeLong estimates that the average worker in the early 1600s earned wages that could afford him 2,500 food calories. This worker lived in Western Europe. Two hundred years later, that same worker could afford 3,000 food calories. However, between 1800 and 1875, just a time span of just 75 years, economic growth was so rapid that western European workers could purchase 5,000 food calories a day. By 2012, a low skilled worker in an affluent Western European/North American country could afford to purchase 2.4 million food calories per day.

What caused such a rapid rise in living standards between 1800 and 1875 and thereafter? Why is it that many countries, especially those in Western Europe, North America, and parts of East Asia, can feed their populations more than adequately, while others cannot? We will look at these and other questions as we examine long-run economic growth.

Introduction to Economic Growth

In this chapter, you will learn about:

- The Relatively Recent Arrival of Economic Growth
- Labor Productivity and Economic Growth
- Components of Economic Growth

- Economic Convergence

Every country worries about economic growth. In the United States and other high-income countries, the question is whether economic growth continues to provide the same remarkable gains in our standard of living as it did during the twentieth century. Meanwhile, can middle-income countries like Brazil, Egypt, or Poland catch up to the higher-income countries, or must they remain in the second tier of per capita income? Of the world's population of roughly 7.5 billion people, about 1.1 billion are scraping by on incomes that average less than $2 per day, not that different from the standard of living 2,000 years ago. Can the world's poor be lifted from their fearful poverty? As the 1995 Nobel laureate in economics, Robert E. Lucas Jr., once noted: "The consequences for human welfare involved in questions like these are simply staggering: Once one starts to think about them, it is hard to think about anything else."

Dramatic improvements in a nation's standard of living are possible. After the Korean War in the late 1950s, the Republic of Korea, often called South Korea, was one of the poorest economies in the world. Most South Koreans worked in peasant agriculture. According to the British economist Angus Maddison, who devoted life's work to measuring GDP and population in the world economy, GDP per capita in 1990 international dollars was $854 per year. From the 1960s to the early twenty-first century, a time period well within the lifetime and memory of many adults, the South Korean economy grew rapidly. Over these four decades, GDP per capita increased by more than 6% per year. According to the World Bank, GDP for South Korea now exceeds $30,000 in nominal terms, placing it firmly among high-income countries like Italy, New Zealand, and Israel. Measured by total GDP in 2015, South Korea is the eleventh-largest economy in the world. For a nation of 50 million people, this transformation is extraordinary.

South Korea is a standout example, but it is not the only case of rapid and sustained economic growth. Other East Asian nations, like Thailand and Indonesia, have seen very rapid growth as well. China has grown enormously since it enacted market-oriented economic reforms around 1980. GDP per capita in high-income economies like the United States also has grown dramatically albeit over a longer time frame. Since the Civil War, the U.S. economy has transformed from a primarily rural and agricultural economy to an economy based on services, manufacturing, and technology.

7.1 | The Relatively Recent Arrival of Economic Growth

By the end of this section, you will be able to:

- Explain the conditions that have allowed for modern economic growth in the last two centuries
- Analyze the influence of public policies on an economy's long-run economic growth

Let's begin with a brief overview of spectacular economic growth patterns around the world in the last two centuries. We commonly refer to this as the period of **modern economic growth**. (Later in the chapter we will discuss lower economic growth rates and some key ingredients for economic progress.) Rapid and sustained economic growth is a relatively recent experience for the human race. Before the last two centuries, although rulers, nobles, and conquerors could afford some extravagances and although economies rose above the subsistence level, the average person's standard of living had not changed much for centuries.

Progressive, powerful economic and institutional changes started to have a significant effect in the late eighteenth and early nineteenth centuries. According to the Dutch economic historian Jan Luiten van Zanden, slavery-based societies, favorable demographics, global trading routes, and standardized trading institutions that spread with different empires set the stage for the Industrial Revolution to succeed. The **Industrial Revolution** refers to the widespread use of power-driven machinery and the economic and social changes that resulted in the first half of the 1800s. Ingenious machines—the steam engine, the power loom, and the steam locomotive—performed tasks that otherwise would have taken vast numbers of workers to do. The Industrial Revolution began in Great Britain, and soon spread to the United States, Germany, and other countries.

The jobs for ordinary people working with these machines were often dirty and dangerous by modern standards, but the alternative jobs of that time in peasant agriculture and small-village industry were often dirty and dangerous, too. The new jobs of the Industrial Revolution typically offered higher pay and a chance for social mobility. A self-reinforcing cycle began: New inventions and investments generated profits, the profits provided funds for more new investment and inventions, and the investments and inventions provided opportunities for further profits. Slowly, a

group of national economies in Europe and North America emerged from centuries of sluggishness into a period of rapid modern growth. During the last two centuries, the average GDP growth rate per capita in the leading industrialized countries has been about 2% per year. What were times like before then? Read the following Clear It Up feature for the answer.

Clear It Up

What were economic conditions like before 1870?

Angus Maddison, a quantitative economic historian, led the most systematic inquiry into national incomes before 1870. Economists recently have refined and used his methods to compile GDP per capita estimates from year 1 C.E. to 1348. **Table 7.1** is an important counterpoint to most of the narrative in this chapter. It shows that nations can decline as well as rise. A wide array of forces, such as epidemics, natural and weather-related disasters, the inability to govern large empires, and the remarkably slow pace of technological and institutional progress explain declines in income. Institutions are the traditions and laws by which people in a community agree to behave and govern themselves. Such institutions include marriage, religion, education, and laws of governance. Institutional progress is the development and codification of these institutions to reinforce social order, and thus, economic growth.

One example of such an institution is the Magna Carta (Great Charter), which the English nobles forced King John to sign in 1215. The Magna Carta codified the principles of due process, whereby a free man could not be penalized unless his peers had made a lawful judgment against him. The United States in its own constitution later adopted this concept. This social order may have contributed to England's GDP per capita in 1348, which was second to that of northern Italy.

In studying economic growth, a country's institutional framework plays a critical role. **Table 7.1** also shows relative global equality for almost 1,300 years. After this, we begin to see significant divergence in income (not in the table).

Year	Northern Italy	Spain	England	Holland	Byzantium	Iraq	Egypt	Japan
1	$800	$600	$600	$600	$700	$700	$700	-
730	-	-	-	-	-	$920	$730	$402
1000	-	-	-	-	$600	$820	$600	-
1150	-	-	-	-	$580	$680	$660	$520
1280	-	-	-	-	-	-	$670	$527
1300	$1,588	$864	$892	-	-	-	$610	-
1348	$1,486	$907	$919	-	-	-	-	-

Table 7.1 GDP Per Capita Estimates in Current International Dollars from AD 1 to 1348 (Source: Bolt and van Zanden. "The First Update of the Maddison Project. Re-Estimating Growth Before 1820." 2013)

Another fascinating and underreported fact is the high levels of income, compared to others at that time, attained by the Islamic Empire Abbasid Caliphate—which was founded in present-day Iraq in 730 C.E. At its height, the empire spanned large regions of the Middle East, North Africa, and Spain until its gradual decline over 200 years.

The Industrial Revolution led to increasing inequality among nations. Some economies took off, whereas others, like many of those in Africa or Asia, remained close to a subsistence standard of living. General calculations show that the 17 countries of the world with the most-developed economies had, on average, 2.4 times the GDP per capita of the world's poorest economies in 1870. By 1960, the most developed economies had 4.2 times the GDP per capita of the poorest economies.

However, by the middle of the twentieth century, some countries had shown that catching up was possible. Japan's economic growth took off in the 1960s and 1970s, with a growth rate of real GDP per capita averaging 11% per year during those decades. Certain countries in Latin America experienced a boom in economic growth in the 1960s as well. In Brazil, for example, GDP per capita expanded by an average annual rate of 11.1% from 1968 to 1973. In the 1970s, some East Asian economies, including South Korea, Thailand, and Taiwan, saw rapid growth. In these countries, growth rates of 11% to 12% per year in GDP per capita were not uncommon. More recently, China, with its population of nearly 1.4 billion people, grew at a per capita rate 9% per year from 1984 into the 2000s. India, with a population of 1.3 billion, has shown promising signs of economic growth, with growth in GDP per capita of about 4% per year during the 1990s and climbing toward 7% to 8% per year in the 2000s.

Link It Up

Visit this website (http://openstaxcollege.org/l/asiadevbank) to read about the Asian Development Bank.

These waves of catch-up economic growth have not reached all shores. In certain African countries like Niger, Tanzania, and Sudan, for example, GDP per capita at the start of the 2000s was still less than $300, not much higher than it was in the nineteenth century and for centuries before that. In the context of the overall situation of low-income people around the world, the good economic news from China (population: 1.4 billion) and India (population: 1.3 billion) is, nonetheless, astounding and heartening.

Economic growth in the last two centuries has made a striking change in the human condition. Richard Easterlin, an economist at the University of Southern California, wrote in 2000:

> By many measures, a revolution in the human condition is sweeping the world. Most people today are better fed, clothed, and housed than their predecessors two centuries ago. They are healthier, live longer, and are better educated. Women's lives are less centered on reproduction and political democracy has gained a foothold. Although Western Europe and its offshoots have been the leaders of this advance, most of the less developed nations have joined in during the 20th century, with the newly emerging nations of sub-Saharan Africa the latest to participate. Although the picture is not one of universal progress, it is the greatest advance in the human condition of the world's population ever achieved in such a brief span of time.

Rule of Law and Economic Growth

Economic growth depends on many factors. Key among those factors is adherence to the **rule of law** and protection of property rights and **contractual rights** by a country's government so that markets can work effectively and efficiently. Laws must be clear, public, fair, enforced, and equally applicable to all members of society. Property rights, as you might recall from **Environmental Protection and Negative Externalities (http://cnx.org/content/m63688/latest/)** are the rights of individuals and firms to own property and use it as they see fit. If you have $100, you have the right to use that money, whether you spend it, lend it, or keep it in a jar. It is your property. The definition of property includes physical property as well as the right to your training and experience, especially

since your training is what determines your livelihood. Using this property includes the right to enter into contracts with other parties with your property. Individuals or firms must own the property to enter into a contract.

Contractual rights, then, are based on property rights and they allow individuals to enter into agreements with others regarding the use of their property providing recourse through the legal system in the event of noncompliance. One example is the employment agreement: a skilled surgeon operates on an ill person and expects payment. Failure to pay would constitute property theft by the patient. The theft is property the services that the surgeon provided. In a society with strong property rights and contractual rights, the terms of the patient–surgeon contract will be fulfilled, because the surgeon would have recourse through the court system to extract payment from that individual. Without a legal system that enforces contracts, people would not be likely to enter into contracts for current or future services because of the risk of non-payment. This would make it difficult to transact business and would slow economic growth.

The World Bank considers a country's legal system effective if it upholds property rights and contractual rights. The World Bank has developed a ranking system for countries' legal systems based on effective protection of property rights and rule-based governance using a scale from 1 to 6, with 1 being the lowest and 6 the highest rating. In 2013, the world average ranking was 2.9. The three countries with the lowest ranking of 1.5 were Afghanistan, the Central African Republic, and Zimbabwe. Their GDP per capita was $679, $333, and $1,007 respectively. The World Bank cites Afghanistan as having a low standard of living, weak government structure, and lack of adherence to the rule of law, which has stymied its economic growth. The landlocked Central African Republic has poor economic resources as well as political instability and is a source of children used in human trafficking. Zimbabwe has had declining and often negative growth for much of the period since 1998. Land redistribution and price controls have disrupted the economy, and corruption and violence have dominated the political process. Although global economic growth has increased, those countries lacking a clear system of property rights and an independent court system free from corruption have lagged far behind.

7.2 | Labor Productivity and Economic Growth

By the end of this section, you will be able to:

- Identify the role of labor productivity in promoting economic growth
- Analyze the sources of economic growth using the aggregate production function
- Measure an economy's rate of productivity growth
- Evaluate the power of sustained growth

Sustained long-term economic growth comes from increases in worker productivity, which essentially means how well we do things. In other words, how efficient is your nation with its time and workers? **Labor productivity** is the value that each employed person creates per unit of his or her input. The easiest way to comprehend labor productivity is to imagine a Canadian worker who can make 10 loaves of bread in an hour versus a U.S. worker who in the same hour can make only two loaves of bread. In this fictional example, the Canadians are more productive. More productivity essentially means you can do more in the same amount of time. This in turn frees up resources for workers to use elsewhere.

What determines how productive workers are? The answer is pretty intuitive. The first determinant of labor productivity is human capital. **Human capital** is the accumulated knowledge (from education and experience), skills, and expertise that the average worker in an economy possesses. Typically the higher the average level of education in an economy, the higher the accumulated human capital and the higher the labor productivity.

The second factor that determines labor productivity is technological change. **Technological change** is a combination of **invention**—advances in knowledge—and **innovation**, which is putting those advances to use in a new product or service. For example, the transistor was invented in 1947. It allowed us to miniaturize the footprint of electronic devices and use less power than the tube technology that came before it. Innovations since then have produced smaller and better transistors that are ubiquitous in products as varied as smart-phones, computers, and escalators. Developing the transistor has allowed workers to be anywhere with smaller devices. People can use these devices to communicate with other workers, measure product quality or do any other task in less time, improving worker productivity.

The third factor that determines labor productivity is economies of scale. Recall that economies of scale are the cost advantages that industries obtain due to size. (Read more about economies of scale in **Production, Cost and Industry Structure (http://cnx.org/content/m63680/latest/)** .) Consider again the case of the fictional

Canadian worker who could produce 10 loaves of bread in an hour. If this difference in productivity was due only to economies of scale, it could be that the Canadian worker had access to a large industrial-size oven while the U.S. worker was using a standard residential size oven.

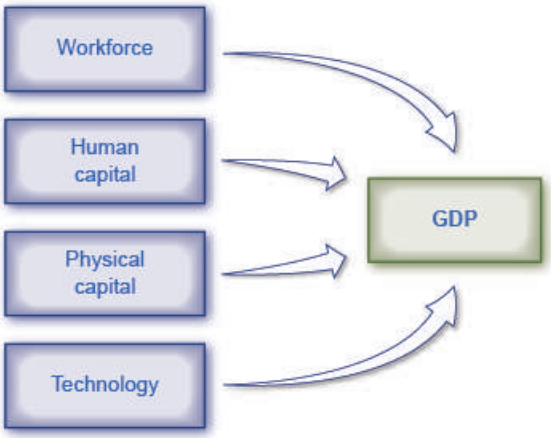
Now that we have explored the determinants of worker productivity, let's turn to how economists measure economic growth and productivity.

Sources of Economic Growth: The Aggregate Production Function

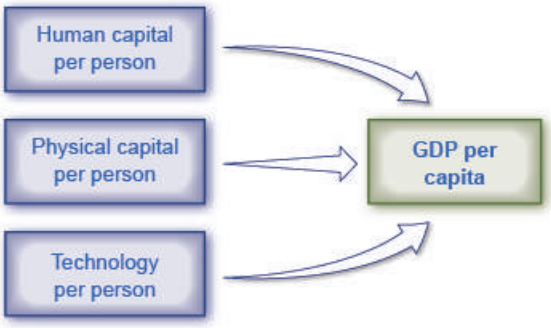
To analyze the sources of economic growth, it is useful to think about a **production function**, which is the technical relationship by which economic inputs like labor, machinery, and raw materials are turned into outputs like goods and services that consumers use. A microeconomic production function describes a firm's or perhaps an industry's inputs and outputs. In macroeconomics, we call the connection from inputs to outputs for the entire economy an **aggregate production function**.

Components of the Aggregate Production Function

Economists construct different production functions depending on the focus of their studies. **Figure 7.2** presents two examples of aggregate production functions. In the first production function in **Figure 7.2** (a), the output is GDP. The inputs in this example are workforce, human capital, physical capital, and technology. We discuss these inputs further in the module, Components of Economic Growth.



(a) Aggregate production function with GDP as its output



(b) Aggregate production function with GDP per capita as its output

Figure 7.2 Aggregate Production Functions An aggregate production function shows what goes into producing the output for an overall economy. (a) This aggregate production function has GDP as its output. (b) This aggregate production function has GDP per capita as its output. Because we calculate it on a per-person basis, we already figure the labor input into the other factors and we do not need to list it separately.

Measuring Productivity

An economy's rate of productivity growth is closely linked to the growth rate of its GDP per capita, although the two are not identical. For example, if the percentage of the population who holds jobs in an economy increases, GDP per capita will increase but the productivity of individual workers may not be affected. Over the long term, the only way that GDP per capita can grow continually is if the productivity of the average worker rises or if there are complementary increases in capital.

A common measure of U.S. productivity per worker is dollar value per hour the worker contributes to the employer's output. This measure excludes government workers, because their output is not sold in the market and so their productivity is hard to measure. It also excludes farming, which accounts for only a relatively small share of the U.S. economy. **Figure 7.3** shows an index of output per hour, with 2009 as the base year (when the index equals 100). The index equaled about 106 in 2014. In 1972, the index equaled 50, which shows that workers have more than doubled their productivity since then.

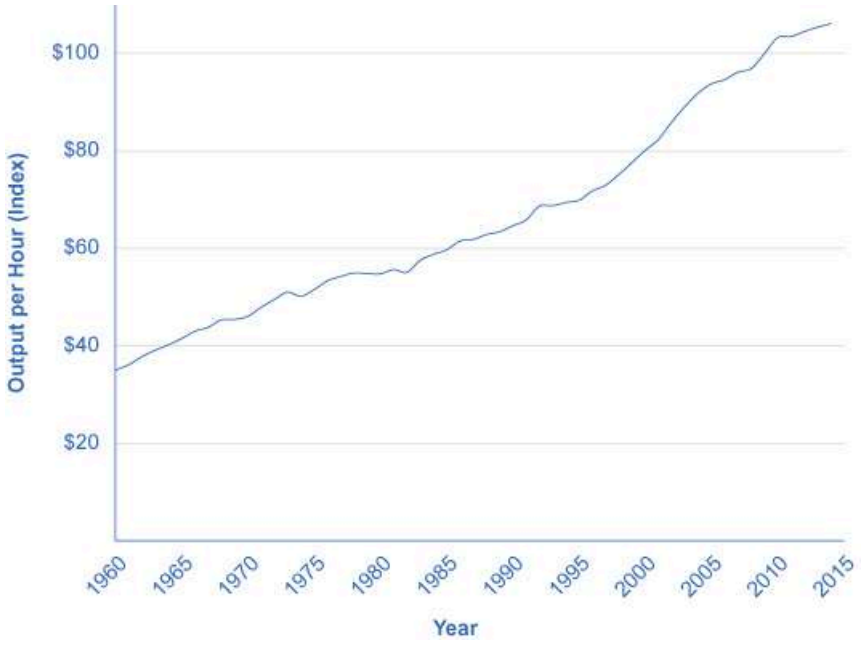


Figure 7.3 Output per Hour Worked in the U.S. Economy, 1947–2011 Output per hour worked is a measure of worker productivity. In the U.S. economy, worker productivity rose more quickly in the 1960s and the mid-1990s compared with the 1970s and 1980s. However, these growth-rate differences are only a few percentage points per year. Look carefully to see them in the changing slope of the line. The average U.S. worker produced over twice as much per hour in 2014 than he did in the early 1970s. (Source: U.S. Department of Labor, Bureau of Labor Statistics.)

According to the Department of Labor, U.S. productivity growth was fairly strong in the 1950s but then declined in the 1970s and 1980s before rising again in the second half of the 1990s and the first half of the 2000s. In fact, the rate of productivity measured by the change in output per hour worked averaged 3.2% per year from 1950 to 1970; dropped to 1.9% per year from 1970 to 1990; and then climbed back to over 2.3% from 1991 to the present, with another modest slowdown after 2001. **Figure 7.4** shows average annual rates of productivity growth averaged over time since 1950.

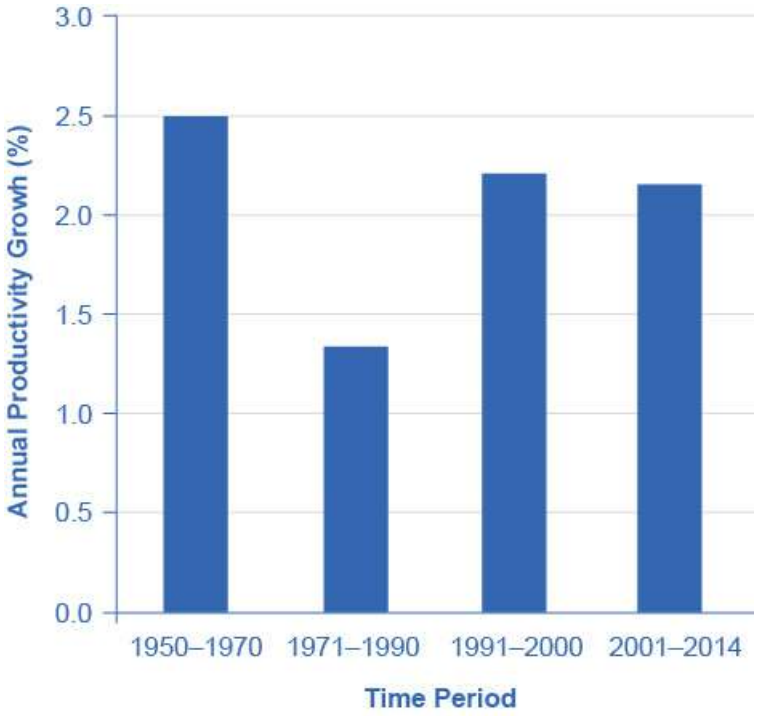


Figure 7.4 Productivity Growth Since 1950 U.S. growth in worker productivity was very high between 1950 and 1970. It then declined to lower levels in the 1970s and the 1980s. The late 1990s and early 2000s saw productivity rebound, but then productivity sagged a bit in the 2000s. Some think the productivity rebound of the late 1990s and early 2000s marks the start of a "new economy" built on higher productivity growth, but we cannot determine this until more time has passed. (Source: U.S. Department of Labor, Bureau of Labor Statistics.)

The "New Economy" Controversy

In recent years a controversy has been brewing among economists about the resurgence of U.S. productivity in the second half of the 1990s. One school of thought argues that the United States had developed a "new economy" based on the extraordinary advances in communications and information technology of the 1990s. The most optimistic proponents argue that it would generate higher average productivity growth for decades to come. The pessimists, alternatively, argue that even five or ten years of stronger productivity growth does not prove that higher productivity will last for the long term. It is hard to infer anything about long-term productivity trends during the later part of the 2000s, because the steep 2008-2009 recession, with its sharp but not completely synchronized declines in output and employment, complicates any interpretation. While productivity growth was high in 2009 and 2010 (around 3%), it has slowed down since then.

Productivity growth is also closely linked to the average level of wages. Over time, the amount that firms are willing to pay workers will depend on the value of the output those workers produce. If a few employers tried to pay their workers less than what those workers produced, then those workers would receive offers of higher wages from other profit-seeking employers. If a few employers mistakenly paid their workers more than what those workers produced, those employers would soon end up with losses. In the long run, productivity per hour is the most important determinant of the average wage level in any economy. To learn how to compare economies in this regard, follow the steps in the following Work It Out feature.

Work It Out

Comparing the Economies of Two Countries

The Organization for Economic Co-operation and Development (OECD) tracks data on the annual growth rate of real GDP per hour worked. You can find these data on the OECD data webpage "Growth in GDP per capita,

productivity and ULC" at this (http://stats.oecd.org/Index.aspx?DataSetCode=PDB_GR) website.

Step 1. Visit the OECD website given above and select two countries to compare.

Step 2. On the drop-down menu "Subject," select " GDP per capita, constant prices," and under "Measure," select "Annual growth/change." Then record the data for the countries you have chosen for the five most recent years.

Step 3. Go back to the drop-down "Subject" menu and select "GDP per hour worked, constant prices," and under "Measure" again select "Annual growth/change." Select data for the same years for which you selected GDP per capita data.

Step 4. Compare real GDP growth for both countries. Table 7.2 provides an example of a comparison between Belgium and Canada.

Australia	**2011**	**2012**	**2013**	**2014**	**2015**
Real GDP/Capita Growth (%)	2.3%	1.5%	1.3%	1.4	0.1%
Real GDP Growth/Hours Worked (%)	1.7%	−0.1%	1.4%	2.2%	−0.2%
Belgium	**2011**	**2012**	**2013**	**2014**	**2015**
Real GDP/Capita Growth (%)	0.9	−0.6	−0.5	1.2	1.0
Real GDP Growth/Hours Worked (%)	−0.5	−0.3	0.4	1.4	0.9

Table 7.2

Step 5. For both measures, growth in Canada is greater than growth in Belgium for the first four years. In addition, there are year-to-year fluctuations. Many factors can affect growth. For example, one factor that may have contributed to Canada's stronger growth may be its larger inflows of immigrants, who generally contribute to economic growth.

The Power of Sustained Economic Growth

Nothing is more important for people's standard of living than sustained economic growth. Even small changes in the rate of growth, when sustained and compounded over long periods of time, make an enormous difference in the standard of living. Consider Table 7.3, in which the rows of the table show several different rates of growth in GDP per capita and the columns show different periods of time. Assume for simplicity that an economy starts with a GDP per capita of 100. The table then applies the following formula to calculate what GDP will be at the given growth rate in the future:

$$\text{GDP at starting date} \times (1 + \text{growth rate of GDP})^{\text{years}} = \text{GDP at end date}$$

For example, an economy that starts with a GDP of 100 and grows at 3% per year will reach a GDP of 209 after 25 years; that is, $100\ (1.03)^{25} = 209$.

The slowest rate of GDP per capita growth in the table, just 1% per year, is similar to what the United States experienced during its weakest years of productivity growth. The second highest rate, 3% per year, is close to what the U.S. economy experienced during the strong economy of the late 1990s and into the 2000s. Higher rates of per capita growth, such as 5% or 8% per year, represent the experience of rapid growth in economies like Japan, Korea, and China.

Table 7.3 shows that even a few percentage points of difference in economic growth rates will have a profound effect if sustained and compounded over time. For example, an economy growing at a 1% annual rate over 50 years will see its GDP per capita rise by a total of 64%, from 100 to 164 in this example. However, a country growing at a 5% annual rate will see (almost) the same amount of growth—from 100 to 163—over just 10 years. Rapid rates of economic growth can bring profound transformation. (See the following Clear It Up feature on the relationship between compound growth rates and compound interest rates.) If the rate of growth is 8%, young adults starting at

age 20 will see the average standard of living in their country more than double by the time they reach age 30, and grow nearly sixfold by the time they reach age 45.

Growth Rate	Value of an original 100 in 10 Years	Value of an original 100 in 25 Years	Value of an original 100 in 50 Years
1%	110	128	164
3%	134	209	438
5%	163	339	1,147
8%	216	685	4,690

Table 7.3 Growth of GDP over Different Time Horizons

Clear It Up

How are compound growth rates and compound interest rates related?

The formula for GDP growth rates over different periods of time, as **Figure 7.3** shows, is exactly the same as the formula for how a given amount of financial savings grows at a certain interest rate over time, as presented in **Choice in a World of Scarcity**. Both formulas have the same ingredients:

- an original starting amount, in one case GDP and in the other case an amount of financial saving;
- a percentage increase over time, in one case the GDP growth rate and in the other case an interest rate;
- and an amount of time over which this effect happens.

Recall that compound interest is interest that is earned on past interest. It causes the total amount of financial savings to grow dramatically over time. Similarly, compound rates of economic growth, or the **compound growth rate**, means that we multiply the rate of growth by a base that includes past GDP growth, with dramatic effects over time.

For example, in 2013, the Central Intelligence Agency's World Fact Book reported that South Korea had a GDP of $1.67 trillion with a growth rate of 2.8%. We can estimate that at that growth rate, South Korea's GDP will be $1.92 trillion in five years. If we apply the growth rate to each year's ending GDP for the next five years, we will calculate that at the end of year one, GDP is $1.72 trillion. In year two, we start with the end-of-year one value of $1.72 and increase it by 2.8%. Year three starts with the end-of-year two GDP, and we increase it by 2.8% and so on, as **Table 7.4** depicts.

Year	Starting GDP	Growth Rate 2%	Year-End Amount
1	$1.67 Trillion ×	(1+0.028)	$1.72 Trillion
2	$1.72 Trillion ×	(1+0.028)	$1.76 Trillion
3	$1.76 Trillion ×	(1+0.028)	$1.81 Trillion
4	$1.81 Trillion ×	(1+0.028)	$1.87 Trillion
5	$1.87 Trillion ×	(1+0.028)	$1.92 Trillion

Table 7.4

Another way to calculate the growth rate is to apply the following formula:

$$\text{Future Value} = \text{Present Value} \times (1 + g)^{n}$$

Where "future value" is the value of GDP five years hence, "present value" is the starting GDP amount of $1.67 trillion, "g" is the growth rate of 2.8%, and "n" is the number of periods for which we are calculating growth.

$$\text{Future Value} = 1.67 \times (1+0.028)^{5} = \$1.92 \text{ trillion}$$

7.3 | Components of Economic Growth

By the end of this section, you will be able to:

- Discuss the components of economic growth, including physical capital, human capital, and technology
- Explain capital deepening and its significance
- Analyze the methods employed in economic growth accounting studies
- Identify factors that contribute to a healthy climate for economic growth

Over decades and generations, seemingly small differences of a few percentage points in the annual rate of economic growth make an enormous difference in GDP per capita. In this module, we discuss some of the components of economic growth, including physical capital, human capital, and technology.

The category of **physical capital** includes the plant and equipment that firms use as well as things like roads (also called **infrastructure**). Again, greater physical capital implies more output. Physical capital can affect productivity in two ways: (1) an increase in the *quantity* of physical capital (for example, more computers of the same quality); and (2) an increase in the *quality* of physical capital (same number of computers but the computers are faster, and so on). **Human capital** refers to the skills and knowledge that make workers productive. Human capital and physical capital accumulation are similar: In both cases, investment now pays off in higher productivity in the future.

The category of **technology** is the "joker in the deck." Earlier we described it as the combination of invention and innovation. When most people think of new technology, the invention of new products like the laser, the smartphone, or some new wonder drug come to mind. In food production, developing more drought-resistant seeds is another example of technology. Technology, as economists use the term, however, includes still more. It includes new ways of organizing work, like the invention of the assembly line, new methods for ensuring better quality of output in factories, and innovative institutions that facilitate the process of converting inputs into output. In short, technology comprises all the advances that make the existing machines and other inputs produce more, and at higher quality, as well as altogether new products.

It may not make sense to compare the GDPs of China and say, Benin, simply because of the great difference in population size. To understand economic growth, which is really concerned with the growth in living standards of an average person, it is often useful to focus on GDP per capita. Using GDP per capita also makes it easier to compare countries with smaller numbers of people, like Belgium, Uruguay, or Zimbabwe, with countries that have larger populations, like the United States, the Russian Federation, or Nigeria.

To obtain a per capita production function, divide each input in **Figure 7.2**(a) by the population. This creates a second aggregate production function where the output is GDP per capita (that is, GDP divided by population). The inputs are the average level of human capital per person, the average level of physical capital per person, and the level of technology per person—see **Figure 7.2**(b). The result of having population in the denominator is mathematically appealing. Increases in population lower per capita income. However, increasing population is important for the average person only if the rate of income growth exceeds population growth. A more important reason for constructing a per capita production function is to understand the contribution of human and physical capital.

Capital Deepening

When society increases the level of capital per person, we call the result **capital deepening**. The idea of capital deepening can apply both to additional human capital per worker and to additional physical capital per worker.

Recall that one way to measure human capital is to look at the average levels of education in an economy. **Figure 7.5** illustrates the human capital deepening for U.S. workers by showing that the proportion of the U.S. population with a high school and a college degree is rising. As recently as 1970, for example, only about half of U.S. adults had at least a high school diploma. By the start of the twenty-first century, more than 80% of adults had graduated from high school. The idea of human capital deepening also applies to the years of experience that workers have, but the average experience level of U.S. workers has not changed much in recent decades. Thus, the key dimension for deepening human capital in the U.S. economy focuses more on additional education and training than on a higher average level of work experience.

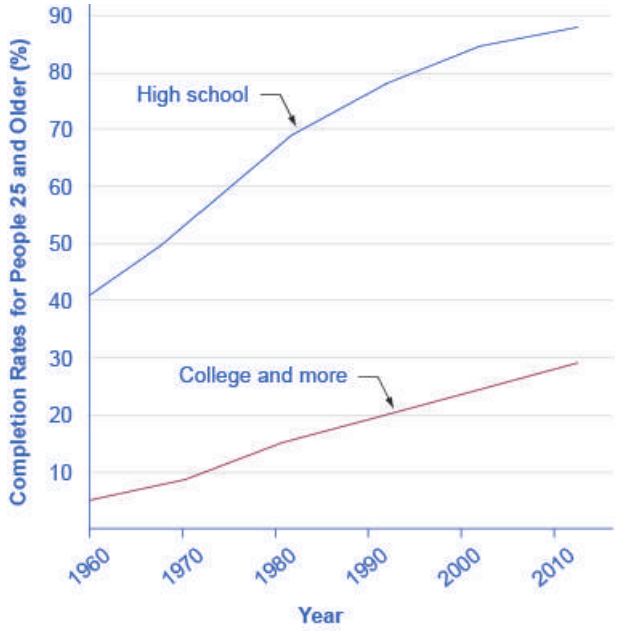


Figure 7.5 Human Capital Deepening in the U.S. Rising levels of education for persons 25 and older show the deepening of human capital in the U.S. economy. Even today, under one-third of U.S. adults have completed a four-year college degree. There is clearly room for additional deepening of human capital to occur. (Source: US Department of Education, National Center for Education Statistics)

Figure 7.6 shows physical capital deepening in the U.S. economy. The average U.S. worker in the late 2000s was working with physical capital worth almost three times as much as that of the average worker of the early 1950s.

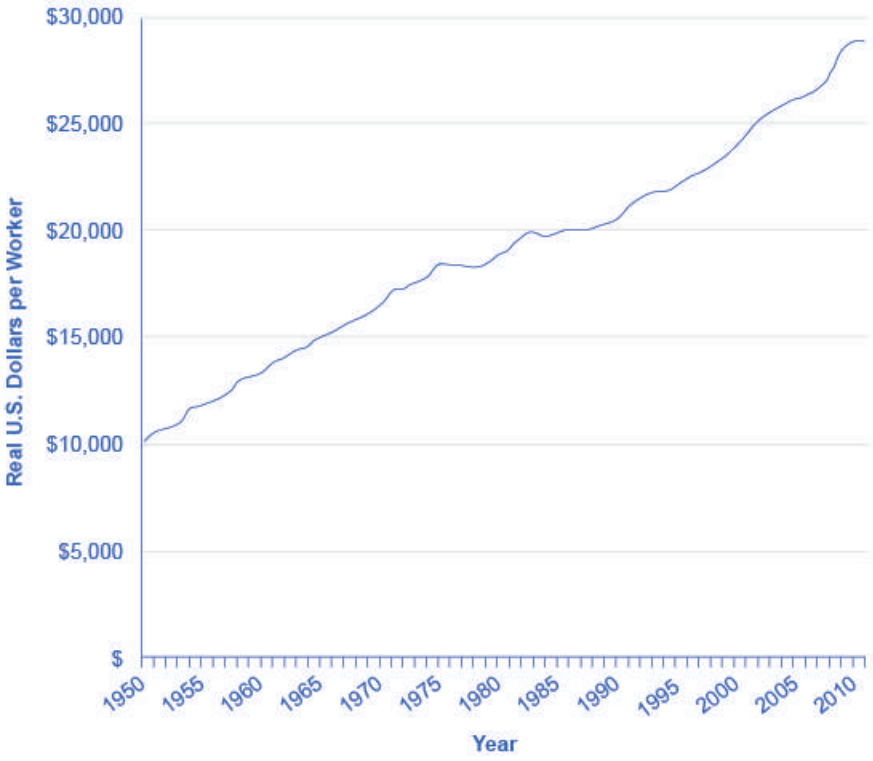


Figure 7.6 Physical Capital per Worker in the United States The value of the physical capital, measured by plant and equipment, used by the average worker in the U.S. economy has risen over the decades. The increase may have leveled off a bit in the 1970s and 1980s, which were, not coincidentally, times of slower-than-usual growth in worker productivity. We see a renewed increase in physical capital per worker in the late 1990s, followed by a flattening in the early 2000s. (Source: Center for International Comparisons of Production, Income and Prices, University of Pennsylvania)

Not only does the current U.S. economy have better-educated workers with more and improved physical capital than it did several decades ago, but these workers have access to more advanced technologies. Growth in technology is impossible to measure with a simple line on a graph, but evidence that we live in an age of technological marvels is all around us—discoveries in genetics and in the structure of particles, the wireless internet, and other inventions almost too numerous to count. The U.S. Patent and Trademark Office typically has issued more than 150,000 patents annually in recent years.

This recipe for economic growth—investing in labor productivity, with investments in human capital and technology, as well as increasing physical capital—also applies to other economies. South Korea, for example, already achieved universal enrollment in primary school (the equivalent of kindergarten through sixth grade in the United States) by 1965, when Korea's GDP per capita was still near its rock bottom low. By the late 1980s, Korea had achieved almost universal secondary school education (the equivalent of a high school education in the United States). With regard to physical capital, Korea's rates of investment had been about 15% of GDP at the start of the 1960s, but doubled to 30–35% of GDP by the late 1960s and early 1970s. With regard to technology, South Korean students went to universities and colleges around the world to obtain scientific and technical training, and South Korean firms reached out to study and form partnerships with firms that could offer them technological insights. These factors combined to foster South Korea's high rate of economic growth.

Growth Accounting Studies

Since the late 1950s, economists have conducted growth accounting studies to determine the extent to which physical and human capital deepening and technology have contributed to growth. The usual approach uses an aggregate production function to estimate how much of per capita economic growth can be attributed to growth in physical capital and human capital. We can measure these two inputs at least roughly. The part of growth that is unexplained by measured inputs, called the residual, is then attributed to growth in technology. The exact numerical estimates differ from study to study and from country to country, depending on how researchers measured these three main

factors and over what time horizons. For studies of the U.S. economy, three lessons commonly emerge from growth accounting studies.

First, technology is typically the most important contributor to U.S. economic growth. Growth in human capital and physical capital often explains only half or less than half of the economic growth that occurs. New ways of doing things are tremendously important.

Second, while investment in physical capital is essential to growth in labor productivity and GDP per capita, building human capital is at least as important. Economic growth is not just a matter of more machines and buildings. One vivid example of the power of human capital and technological knowledge occurred in Europe in the years after World War II (1939–1945). During the war, a large share of Europe's physical capital, such as factories, roads, and vehicles, was destroyed. Europe also lost an overwhelming amount of human capital in the form of millions of men, women, and children who died during the war. However, the powerful combination of skilled workers and technological knowledge, working within a market-oriented economic framework, rebuilt Europe's productive capacity to an even higher level within less than two decades.

A third lesson is that these three factors of human capital, physical capital, and technology work together. Workers with a higher level of education and skills are often better at coming up with new technological innovations. These technological innovations are often ideas that cannot increase production until they become a part of new investment in physical capital. New machines that embody technological innovations often require additional training, which builds worker skills further. If the recipe for economic growth is to succeed, an economy needs all the ingredients of the aggregate production function. See the following Clear It Up feature for an example of how human capital, physical capital, and technology can combine to significantly impact lives.

Clear It Up

How do girls' education and economic growth relate in low-income countries?

In the early 2000s, according to the World Bank, about 110 million children between the ages of 6 and 11 were not in school—and about two-thirds of them were girls. In Afghanistan, for example, the literacy rate for those aged 15-24 for the period 2005-2014 was 62% for males and only 32% for females. In Benin, in West Africa, it was 55% for males and 31% for females. In Nigeria, Africa's most populous country, it was 76% for males and 58 percent for females.

Whenever any child does not receive a basic education, it is both a human and an economic loss. In low-income countries, wages typically increase by an average of 10 to 20% with each additional year of education. There is, however, some intriguing evidence that helping girls in low-income countries to close the education gap with boys may be especially important, because of the social role that many of the girls will play as mothers and homemakers.

Girls in low-income countries who receive more education tend to grow up to have fewer, healthier, better-educated children. Their children are more likely to be better nourished and to receive basic health care like immunizations. Economic research on women in low-income economies backs up these findings. When 20 women obtain one additional year of schooling, as a group they will, on average, have one less child. When 1,000 women obtain one additional year of schooling, on average one to two fewer women from that group will die in childbirth. When a woman stays in school an additional year, that factor alone means that, on average, each of her children will spend an additional half-year in school. Education for girls is a good investment because it is an investment in economic growth with benefits beyond the current generation.

A Healthy Climate for Economic Growth

While physical and human capital deepening and better technology are important, equally important to a nation's well-being is the climate or system within which these inputs are cultivated. Both the type of market economy and a legal system that governs and sustains property rights and contractual rights are important contributors to a healthy economic climate.

A healthy economic climate usually involves some sort of market orientation at the microeconomic, individual, or firm decision-making level. Markets that allow personal and business rewards and incentives for increasing human and physical capital encourage overall macroeconomic growth. For example, when workers participate in a competitive and well-functioning labor market, they have an incentive to acquire additional human capital, because additional education and skills will pay off in higher wages. Firms have an incentive to invest in physical capital and in training workers, because they expect to earn higher profits for their shareholders. Both individuals and firms look for new technologies, because even small inventions can make work easier or lead to product improvement. Collectively, such individual and business decisions made within a market structure add up to macroeconomic growth. Much of the rapid growth since the late nineteenth century has come from harnessing the power of competitive markets to allocate resources. This market orientation typically reaches beyond national borders and includes openness to international trade.

A general orientation toward markets does not rule out important roles for government. There are times when markets fail to allocate capital or technology in a manner that provides the greatest benefit for society as a whole. The government's role is to correct these failures. In addition, government can guide or influence markets toward certain outcomes. The following examples highlight some important areas that governments around the world have chosen to invest in to facilitate capital deepening and technology:

- Education. The Danish government requires all children under 16 to attend school. They can choose to attend a public school (*Folkeskole*) or a private school. Students do not pay tuition to attend *Folkeskole*. Thirteen percent of primary/secondary (elementary/high) school is private, and the government supplies vouchers to citizens who choose private school.
- Savings and Investment. In the United States, as in other countries, the government taxes gains from private investment. Low capital gains taxes encourage investment and so also economic growth.
- Infrastructure. The Japanese government in the mid-1990s undertook significant infrastructure projects to improve roads and public works. This in turn increased the stock of physical capital and ultimately economic growth.
- Special Economic Zones. The island of Mauritius is one of the few African nations to encourage international trade in government-supported **special economic zones (SEZ)**. These are areas of the country, usually with access to a port where, among other benefits, the government does not tax trade. As a result of its SEZ, Mauritius has enjoyed above-average economic growth since the 1980s. Free trade does not have to occur in an SEZ however. Governments can encourage international trade across the board, or surrender to protectionism.
- Scientific Research. The European Union has strong programs to invest in scientific research. The researchers Abraham García and Pierre Mohnen demonstrate that firms which received support from the Austrian government actually increased their research intensity and had more sales. Governments can support scientific research and technical training that helps to create and spread new technologies. Governments can also provide a legal environment that protects the ability of inventors to profit from their inventions.

There are many more ways in which the government can play an active role in promoting economic growth. We explore them in other chapters and in particular in **Macroeconomic Policy Around the World**. A healthy climate for growth in GDP per capita and labor productivity includes human capital deepening, physical capital deepening, and technological gains, operating in a market-oriented economy with supportive government policies.

7.4 | Economic Convergence

By the end of this section, you will be able to:

- Explain economic convergence
- Analyze various arguments for and against economic convergence
- Evaluate the speed of economic convergence between high-income countries and the rest of the world

Some low-income and middle-income economies around the world have shown a pattern of **convergence**, in which their economies grow faster than those of high-income countries. GDP increased by an average rate of 2.7% per year in the 1990s and 2.3% per year from 2000 to 2008 in the high-income countries of the world, which include the

United States, Canada, the European Union countries, Japan, Australia, and New Zealand.

Table 7.5 lists 10 countries that belong to an informal "fast growth club." These countries averaged GDP growth (after adjusting for inflation) of at least 5% per year in both the time periods from 1990 to 2000 and from 2000 to 2008. Since economic growth in these countries has exceeded the average of the world's high-income economies, these countries may converge with the high-income countries. The second part of **Table 7.5** lists the "slow growth club," which consists of countries that averaged GDP growth of 2% per year or less (after adjusting for inflation) during the same time periods. The final portion of **Table 7.5** shows GDP growth rates for the countries of the world divided by income.

Country	Average Growth Rate of Real GDP 1990–2000	Average Growth Rate of Real GDP 2000–2008
Fast Growth Club (5% or more per year in both time periods)		
Cambodia	7.1%	9.1%
China	10.6%	9.9%
India	6.0%	7.1%
Ireland	7.5%	5.1%
Jordan	5.0%	6.3%
Laos	6.5%	6.8 %
Mozambique	6.4%	7.3%
Sudan	5.4%	7.3%
Uganda	7.1%	7.3%
Vietnam	7.9%	7.3%
Slow Growth Club (2% or less per year in both time periods)		
Central African Republic	2.0%	0.8%
France	2.0%	1.8%
Germany	1.8%	1.3%
Guinea-Bissau	1.2%	0.2%
Haiti	–1.5%	0.3%
Italy	1.6%	1.2%
Jamaica	0.9%	1.4%
Japan	1.3%	1.3%
Switzerland	1.0%	2.0%
United States (for reference)	3.2%	2.2%

Table 7.5 Economic Growth around the World (Source: http://databank.worldbank.org/data/views/variableSelection/selectvariables.aspx?source=world-development-indicators#c_u)

Country	Average Growth Rate of Real GDP 1990–2000	Average Growth Rate of Real GDP 2000–2008
World Overview		
High income	2.7%	2.3%
Low income	3.8%	5.6%
Middle income	4.7%	6.1%

Table 7.5 Economic Growth around the World (Source: http://databank.worldbank.org/data/views/variableSelection/selectvariables.aspx?source=world-development-indicators#c_u)

Each of the countries in **Table 7.5** has its own unique story of investments in human and physical capital, technological gains, market forces, government policies, and even lucky events, but an overall pattern of convergence is clear. The low-income countries have GDP growth that is faster than that of the middle-income countries, which in turn have GDP growth that is faster than that of the high-income countries. Two prominent members of the fast-growth club are China and India, which between them have nearly 40% of the world's population. Some prominent members of the slow-growth club are high-income countries like France, Germany, Italy, and Japan.

Will this pattern of economic convergence persist into the future? This is a controversial question among economists that we will consider by looking at some of the main arguments on both sides.

Arguments Favoring Convergence

Several arguments suggest that low-income countries might have an advantage in achieving greater worker productivity and economic growth in the future.

A first argument is based on diminishing marginal returns. Even though deepening human and physical capital will tend to increase GDP per capita, the law of diminishing returns suggests that as an economy continues to increase its human and physical capital, the marginal gains to economic growth will diminish. For example, raising the average education level of the population by two years from a tenth-grade level to a high school diploma (while holding all other inputs constant) would produce a certain increase in output. An additional two-year increase, so that the average person had a two-year college degree, would increase output further, but the marginal gain would be smaller. Yet another additional two-year increase in the level of education, so that the average person would have a four-year-college bachelor's degree, would increase output still further, but the marginal increase would again be smaller. A similar lesson holds for physical capital. If the quantity of physical capital available to the average worker increases, by, say, $5,000 to $10,000 (again, while holding all other inputs constant), it will increase the level of output. An additional increase from $10,000 to $15,000 will increase output further, but the marginal increase will be smaller.

Low-income countries like China and India tend to have lower levels of human capital and physical capital, so an investment in capital deepening should have a larger marginal effect in these countries than in high-income countries, where levels of human and physical capital are already relatively high. Diminishing returns implies that low-income economies could converge to the levels that the high-income countries achieve.

A second argument is that low-income countries may find it easier to improve their technologies than high-income countries. High-income countries must continually invent new technologies, whereas low-income countries can often find ways of applying technology that has already been invented and is well understood. The economist Alexander Gerschenkron (1904–1978) gave this phenomenon a memorable name: "the advantages of backwardness." Of course, he did not literally mean that it is an advantage to have a lower standard of living. He was pointing out that a country that is behind has some extra potential for catching up.

Finally, optimists argue that many countries have observed the experience of those that have grown more quickly and have learned from it. Moreover, once the people of a country begin to enjoy the benefits of a higher standard of living, they may be more likely to build and support the market-friendly institutions that will help provide this standard of living.

Link It Up

View this video (http://openstaxcollege.org/l/tedhansrosling) to learn about economic growth across the world.

Arguments That Convergence Is neither Inevitable nor Likely

If the economy's growth depended only on the deepening of human capital and physical capital, then we would expect that economy's growth rate to slow down over the long run because of diminishing marginal returns. However, there is another crucial factor in the aggregate production function: technology.

Developing new technology can provide a way for an economy to sidestep the diminishing marginal returns of capital deepening. **Figure 7.7** shows how. The figure's horizontal axis measures the amount of capital deepening, which on this figure is an overall measure that includes deepening of both physical and human capital. The amount of human and physical capital per worker increases as you move from left to right, from C_1 to C_2 to C_3. The diagram's vertical axis measures per capita output. Start by considering the lowest line in this diagram, labeled Technology 1. Along this aggregate production function, the level of technology is held constant, so the line shows only the relationship between capital deepening and output. As capital deepens from C_1 to C_2 to C_3 and the economy moves from R to U to W, per capita output does increase—but the way in which the line starts out steeper on the left but then flattens as it moves to the right shows the diminishing marginal returns, as additional marginal amounts of capital deepening increase output by ever-smaller amounts. The shape of the aggregate production line (Technology 1) shows that the ability of capital deepening, by itself, to generate sustained economic growth is limited, since diminishing returns will eventually set in.

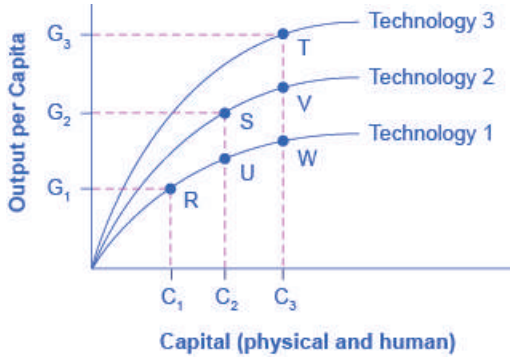


Figure 7.7 Capital Deepening and New Technology Imagine that the economy starts at point R, with the level of physical and human capital C_1 and the output per capita at G_1. If the economy relies only on capital deepening, while remaining at the technology level shown by the Technology 1 line, then it would face diminishing marginal returns as it moved from point R to point U to point W. However, now imagine that capital deepening is combined with improvements in technology. Then, as capital deepens from C_1 to C_2, technology improves from Technology 1 to Technology 2, and the economy moves from R to S. Similarly, as capital deepens from C_2 to C_3, technology increases from Technology 2 to Technology 3, and the economy moves from S to T. With improvements in technology, there is no longer any reason that economic growth must necessarily slow down.

Now, bring improvements in technology into the picture. Improved technology means that with a given set of inputs, more output is possible. The production function labeled Technology 1 in the figure is based on one level of technology, but Technology 2 is based on an improved level of technology, so for every level of capital deepening on

the horizontal axis, it produces a higher level of output on the vertical axis. In turn, production function Technology 3 represents a still higher level of technology, so that for every level of inputs on the horizontal axis, it produces a higher level of output on the vertical axis than either of the other two aggregate production functions.

Most healthy, growing economies are deepening their human and physical capital and increasing technology at the same time. As a result, the economy can move from a choice like point R on the Technology 1 aggregate production line to a point like S on Technology 2 and a point like T on the still higher aggregate production line (Technology 3). With the combination of technology and capital deepening, the rise in GDP per capita in high-income countries does not need to fade away because of diminishing returns. The gains from technology can offset the diminishing returns involved with capital deepening.

Will technological improvements themselves run into diminishing returns over time? That is, will it become continually harder and more costly to discover new technological improvements? Perhaps someday, but, at least over the last two centuries since the beginning of the Industrial Revolution, improvements in technology have not run into diminishing marginal returns. Modern inventions, like the internet or discoveries in genetics or materials science, do not seem to provide smaller gains to output than earlier inventions like the steam engine or the railroad. One reason that technological ideas do not seem to run into diminishing returns is that we often can apply widely the ideas of new technology at a marginal cost that is very low or even zero. A specific worker or group of workers must use a specific additional machine, or an additional year of education. Many workers across the economy can use a new technology or invention at very low marginal cost.

The argument that it is easier for a low-income country to copy and adapt existing technology than it is for a high-income country to invent new technology is not necessarily true, either. When it comes to adapting and using new technology, a society's performance is not necessarily guaranteed, but is the result of whether the country's economic, educational, and public policy institutions are supportive. In theory, perhaps, low-income countries have many opportunities to copy and adapt technology, but if they lack the appropriate supportive economic infrastructure and institutions, the theoretical possibility that backwardness might have certain advantages is of little practical relevance.

Link It Up

Visit this website (http://openstaxcollege.org/l/Indiapoverty) to read more about economic growth in India.

The Slowness of Convergence

Although economic convergence between the high-income countries and the rest of the world seems possible and even likely, it will proceed slowly. Consider, for example, a country that starts off with a GDP per capita of $40,000, which would roughly represent a typical high-income country today, and another country that starts out at $4,000, which is roughly the level in low-income but not impoverished countries like Indonesia, Guatemala, or Egypt. Say that the rich country chugs along at a 2% annual growth rate of GDP per capita, while the poorer country grows at the aggressive rate of 7% per year. After 30 years, GDP per capita in the rich country will be $72,450 (that is, $40,000 $(1 + 0.02)^{30}$) while in the poor country it will be $30,450 (that is, $4,000 $(1 + 0.07)^{30}$). Convergence has occurred. The rich country used to be 10 times as wealthy as the poor one, and now it is only about 2.4 times as wealthy. Even after 30 consecutive years of very rapid growth, however, people in the low-income country are still likely to feel quite poor compared to people in the rich country. Moreover, as the poor country catches up, its opportunities for catch-up growth are reduced, and its growth rate may slow down somewhat.

The slowness of convergence illustrates again that small differences in annual rates of economic growth become huge differences over time. The high-income countries have been building up their advantage in standard of living over decades—more than a century in some cases. Even in an optimistic scenario, it will take decades for the low-income countries of the world to catch up significantly.

Bring it Home

Calories and Economic Growth

We can tell the story of modern economic growth by looking at calorie consumption over time. The dramatic rise in incomes allowed the average person to eat better and consume more calories. How did these incomes increase? The neoclassical growth consensus uses the aggregate production function to suggest that the period of modern economic growth came about because of increases in inputs such as technology and physical and human capital. Also important was the way in which technological progress combined with physical and human capital deepening to create growth and convergence. The issue of distribution of income notwithstanding, it is clear that the average worker can afford more calories in 2017 than in 1875.

Aside from increases in income, there is another reason why the average person can afford more food. Modern agriculture has allowed many countries to produce more food than they need. Despite having more than enough food, however, many governments and multilateral agencies have not solved the food distribution problem. In fact, food shortages, famine, or general food insecurity are caused more often by the failure of government macroeconomic policy, according to the Nobel Prize-winning economist Amartya Sen. Sen has conducted extensive research into issues of inequality, poverty, and the role of government in improving standards of living. Macroeconomic policies that strive toward stable inflation, full employment, education of women, and preservation of property rights are more likely to eliminate starvation and provide for a more even distribution of food.

Because we have more food per capita, global food prices have decreased since 1875. The prices of some foods, however, have decreased more than the prices of others. For example, researchers from the University of Washington have shown that in the United States, calories from zucchini and lettuce are 100 times more expensive than calories from oil, butter, and sugar. Research from countries like India, China, and the United States suggests that as incomes rise, individuals want more calories from fats and protein and fewer from carbohydrates. This has very interesting implications for global food production, obesity, and environmental consequences. Affluent urban India has an obesity problem much like many parts of the United States. The forces of convergence are at work.

KEY TERMS

aggregate production function the process whereby an economy as a whole turns economic inputs such as human capital, physical capital, and technology into output measured as GDP per capita

capital deepening an increase by society in the average level of physical and/or human capital per person

compound growth rate the rate of growth when multiplied by a base that includes past GDP growth

contractual rights the rights of individuals to enter into agreements with others regarding the use of their property providing recourse through the legal system in the event of noncompliance

convergence pattern in which economies with low per capita incomes grow faster than economies with high per capita incomes

human capital the accumulated skills and education of workers

Industrial Revolution the widespread use of power-driven machinery and the economic and social changes that occurred in the first half of the 1800s

infrastructure a component of physical capital such as roads and rail systems

innovation putting advances in knowledge to use in a new product or service

invention advances in knowledge

labor productivity the value of what is produced per worker, or per hour worked (sometimes called worker productivity)

modern economic growth the period of rapid economic growth from 1870 onward

physical capital the plant and equipment that firms use in production; this includes infrastructure

production function the process whereby a firm turns economic inputs like labor, machinery, and raw materials into outputs like goods and services that consumers use

rule of law the process of enacting laws that protect individual and entity rights to use their property as they see fit. Laws must be clear, public, fair, and enforced, and applicable to all members of society

special economic zone (SEZ) area of a country, usually with access to a port where, among other benefits, the government does not tax trade

technological change a combination of invention—advances in knowledge—and innovation

technology all the ways in which existing inputs produce more or higher quality, as well as different and altogether new products

KEY CONCEPTS AND SUMMARY

7.1 The Relatively Recent Arrival of Economic Growth

Since the early nineteenth century, there has been a spectacular process of long-run economic growth during which the world's leading economies—mostly those in Western Europe and North America—expanded GDP per capita at an average rate of about 2% per year. In the last half-century, countries like Japan, South Korea, and China have shown the potential to catch up. The Industrial Revolution facilitated the extensive process of economic growth, that economists often refer to as modern economic growth. This increased worker productivity and trade, as well as the

development of governance and market institutions.

7.2 Labor Productivity and Economic Growth

We can measure productivity, the value of what is produced per worker, or per hour worked, as the level of GDP per worker or GDP per hour. The United States experienced a productivity slowdown between 1973 and 1989. Since then, U.S. productivity has rebounded for the most part, but annual growth in productivity in the nonfarm business sector has been less than one percent each year between 2011 and 2016. It is not clear what productivity growth will be in the coming years. The rate of productivity growth is the primary determinant of an economy's rate of long-term economic growth and higher wages. Over decades and generations, seemingly small differences of a few percentage points in the annual rate of economic growth make an enormous difference in GDP per capita. An aggregate production function specifies how certain inputs in the economy, like human capital, physical capital, and technology, lead to the output measured as GDP per capita.

Compound interest and compound growth rates behave in the same way as productivity rates. Seemingly small changes in percentage points can have big impacts on income over time.

7.3 Components of Economic Growth

Over decades and generations, seemingly small differences of a few percentage points in the annual rate of economic growth make an enormous difference in GDP per capita. Capital deepening refers to an increase in the amount of capital per worker, either human capital per worker, in the form of higher education or skills, or physical capital per worker. Technology, in its economic meaning, refers broadly to all new methods of production, which includes major scientific inventions but also small inventions and even better forms of management or other types of institutions. A healthy climate for growth in GDP per capita consists of improvements in human capital, physical capital, and technology, in a market-oriented environment with supportive public policies and institutions.

7.4 Economic Convergence

When countries with lower GDP levels per capita catch up to countries with higher GDP levels per capita, we call the process convergence. Convergence can occur even when both high- and low-income countries increase investment in physical and human capital with the objective of growing GDP. This is because the impact of new investment in physical and human capital on a low-income country may result in huge gains as new skills or equipment combine with the labor force. In higher-income countries, however, a level of investment equal to that of the low income country is not likely to have as big an impact, because the more developed country most likely already has high levels of capital investment. Therefore, the marginal gain from this additional investment tends to be successively less and less. Higher income countries are more likely to have diminishing returns to their investments and must continually invent new technologies. This allows lower-income economies to have a chance for convergent growth. However, many high-income economies have developed economic and political institutions that provide a healthy economic climate for an ongoing stream of technological innovations. Continuous technological innovation can counterbalance diminishing returns to investments in human and physical capital.

SELF-CHECK QUESTIONS

1. Explain what the Industrial Revolution was and where it began.

2. Explain the difference between property rights and contractual rights. Why do they matter to economic growth?

3. Are there other ways in which we can measure productivity besides the amount produced per hour of work?

4. Assume there are two countries: South Korea and the United States. South Korea grows at 4% and the United States grows at 1%. For the sake of simplicity, assume they both start from the same fictional income level, $10,000. What will the incomes of the United States and South Korea be in 20 years? By how many multiples will each country's income grow in 20 years?

5. What do the growth accounting studies conclude are the determinants of growth? Which is more important, the determinants or how they are combined?

6. What policies can the government of a free-market economy implement to stimulate economic growth?

7. List the areas where government policy can help economic growth.

8. Use an example to explain why, after periods of rapid growth, a low-income country that has not caught up to a high-income country may feel poor.

9. Would the following events usually lead to capital deepening? Why or why not?

a. A weak economy in which businesses become reluctant to make long-term investments in physical capital.
b. A rise in international trade.
c. A trend in which many more adults participate in continuing education courses through their employers and at colleges and universities.

10. What are the "advantages of backwardness" for economic growth?

11. Would you expect capital deepening to result in diminished returns? Why or why not? Would you expect improvements in technology to result in diminished returns? Why or why not?

12. Why does productivity growth in high-income economies not slow down as it runs into diminishing returns from additional investments in physical capital and human capital? Does this show one area where the theory of diminishing returns fails to apply? Why or why not?

REVIEW QUESTIONS

13. How did the Industrial Revolution increase the economic growth rate and income levels in the United States?

14. How much should a nation be concerned if its rate of economic growth is just 2% slower than other nations?

15. How is GDP per capita calculated differently from labor productivity?

16. How do gains in labor productivity lead to gains in GDP per capita?

17. What is an aggregate production function?

18. What is capital deepening?

19. What do economists mean when they refer to improvements in technology?

20. For a high-income economy like the United States, what aggregate production function elements are most important in bringing about growth in GDP per capita? What about a middle-income country such as Brazil? A low-income country such as Niger?

21. List some arguments for and against the likelihood of convergence.

CRITICAL THINKING QUESTIONS

22. Over the past 50 years, many countries have experienced an annual growth rate in real GDP per capita greater than that of the United States. Some examples are China, Japan, South Korea, and Taiwan. Does that mean the United States is regressing relative to other countries? Does that mean these countries will eventually overtake the United States in terms of the growth rate of real GDP per capita? Explain.

23. **Labor Productivity and Economic Growth** outlined the logic of how increased productivity is associated with increased wages. Detail a situation where this is not the case and explain why it is not.

24. Change in labor productivity is one of the most watched international statistics of growth. Visit the St. Louis Federal Reserve website and find the data section (http://research.stlouisfed.org). Find international comparisons of labor productivity, listed under the FRED Economic database (Growth Rate of Total Labor Productivity), and compare two countries in the recent past. State what you think the reasons for differences in labor productivity could be.

25. Refer back to the **Work It Out** about Comparing the Economies of Two Countries and examine the data for the two countries you chose. How are they similar? How are they different?

26. Education seems to be important for human capital deepening. As people become better educated and more knowledgeable, are there limits to how much additional benefit more education can provide? Why or why not?

27. Describe some of the political and social tradeoffs that might occur when a less developed country adopts a strategy to promote labor force participation and economic growth via investment in girls' education.

28. Why is investing in girls' education beneficial for growth?

29. How is the concept of technology, as defined with the aggregate production function, different from our everyday use of the word?

30. What sorts of policies can governments implement to encourage convergence?

31. As technological change makes us more sedentary and food costs increase, obesity is likely. What factors do you think may limit obesity?

PROBLEMS

32. An economy starts off with a GDP per capita of $5,000. How large will the GDP per capita be if it grows at an annual rate of 2% for 20 years? 2% for 40 years? 4% for 40 years? 6% for 40 years?

33. An economy starts off with a GDP per capita of 12,000 euros. How large will the GDP per capita be if it grows at an annual rate of 3% for 10 years? 3% for 30 years? 6% for 30 years?

34. Say that the average worker in Canada has a productivity level of $30 per hour while the average worker in the United Kingdom has a productivity level of $25 per hour (both measured in U.S. dollars). Over the next five years, say that worker productivity in Canada grows at 1% per year while worker productivity in the UK grows 3% per year. After five years, who will have the higher productivity level, and by how much?

35. Say that the average worker in the U.S. economy is eight times as productive as an average worker in Mexico. If the productivity of U.S. workers grows at 2% for 25 years and the productivity of Mexico's workers grows at 6% for 25 years, which country will have higher worker productivity at that point?

8 | Unemployment

Figure 8.1 Out of Business Borders was one of the many companies unable to recover from the 2008-2009 economic recession. (Credit: modification of work by Luis Villa del Campo/Flickr Creative Commons)

Bring it Home

Unemployment and the Great Recession

Nearly eight million U.S. jobs were lost as a consequence of the Great Recession, which lasted from December 2007 to June 2009. At the outset of the recession, the unemployment rate was 5.0%. The rate began rising several months after the recession began, and it peaked at 10.0% in October 2009, several months after the recession ended, according to the Bureau of Labor Statistics (BLS). The job loss represented a huge number of positions gone. Subsequently, the recovery was tepid. Companies added some positions, but as of summer 2013, four years after the end of the recession, unemployment was about 7.5%, well above the pre-recession rate. Employment began increasing at the outset of 2010, and reached its pre-recession level in mid-2014. However, because of population and labor force growth, the unemployment rate at that point was still slightly above 6%. The economy only returned to an unemployment rate of 5.0% in September 2015, and it has remained at or slightly below that level since then, up through January 2017.

This brief overview of unemployment during and after the Great Recession highlights a few important points. First, unemployment is a lagging indicator of business activity. It didn't begin to increase until a few months after the onset of the recession, and it didn't begin to decline until several months after the recovery. Second, the decline in the unemployment rate was quite slow, with the pre-recession unemployment rate only reaching a higher level than six years after the recession ended. This reflects a combination of slow increase in the number of jobs and ongoing increases in the size of the population and the labor force.

It turns out that recent recessions, going back to the early 1990s, have been characterized by longer periods of recovery than their predecessors. We will return to this point at the end of the chapter. However, first we

need to examine unemployment. What constitutes it, and how do we measure it?

Introduction to Unemployment

In this chapter, you will learn about:

- How Economists Define and Compute Unemployment Rate
- Patterns of Unemployment
- What Causes Changes in Unemployment over the Short Run
- What Causes Changes in Unemployment over the Long Run

Unemployment can be a terrible and wrenching life experience—like a serious automobile accident or a messy divorce—whose consequences only someone who has gone through it can fully understand. For unemployed individuals and their families, there is the day-to-day financial stress of not knowing from where the next paycheck is coming. There are painful adjustments, like watching your savings account dwindle, selling a car and buying a cheaper one, or moving to a less expensive place to live. Even when the unemployed person finds a new job, it may pay less than the previous one. For many people, their job is an important part of their self worth. When unemployment separates people from the workforce, it can affect family relationships as well as mental and physical health.

The human costs of unemployment alone would justify making a low level of unemployment an important public policy priority. However, unemployment also includes economic costs to the broader society. When millions of unemployed but willing workers cannot find jobs, economic resource are unused. An economy with high unemployment is like a company operating with a functional but unused factory. The opportunity cost of unemployment is the output that the unemployed workers could have produced.

This chapter will discuss how economists define and compute the unemployment rate. It will examine the patterns of unemployment over time, for the U.S. economy as a whole, for different demographic groups in the U.S. economy, and for other countries. It will then consider an economic explanation for unemployment, and how it explains the patterns of unemployment and suggests public policies for reducing it.

8.1 | How Economists Define and Compute Unemployment Rate

By the end of this section, you will be able to:

- Calculate the labor force participation rate and the unemployment rate
- Explain hidden unemployment and what it means to be in or out of the labor force
- Evaluate the collection and interpretation of unemployment data

Newspaper or television reports typically describe unemployment as a percentage or a rate. A recent report might have said, for example, *from August 2009 to November 2009, the U.S. unemployment rate rose from 9.7% to 10.0%, but by June 2010, it had fallen to 9.5%*. At a glance, the changes between the percentages may seem small. However, remember that the U.S. economy has about 160 million adults (as of the beginning of 2017) who either have jobs or are looking for them. A rise or fall of just 0.1% in the unemployment rate of 160 million potential workers translates into 160,000 people, which is roughly the total population of a city like Syracuse, New York, Brownsville, Texas, or Pasadena, California. Large rises in the unemployment rate mean large numbers of job losses. In November 2009, at the peak of the recession, about 15 million people were out of work. Even with the unemployment rate now at 4.8% as of January 2017, about 7.6 million people who would like to have jobs are out of work.

Link It Up

The Bureau of Labor Statistics (http://openstaxcollege.org/l/BLS1) tracks and reports all data related to unemployment.

Who's In or Out of the Labor Force?

Should we count everyone without a job as unemployed? Of course not. For example, we should not count children as unemployed. Surely, we should not count the retired as unemployed. Many full-time college students have only a part-time job, or no job at all, but it seems inappropriate to count them as suffering the pains of unemployment. Some people are not working because they are rearing children, ill, on vacation, or on parental leave.

The point is that we do not just divide the adult population into employed and unemployed. A third group exists: people who do not have a job, and for some reason—retirement, looking after children, taking a voluntary break before a new job—are not interested in having a job, either. It also includes those who do want a job but have quit looking, often due to discouragement due to their inability to find suitable employment. Economists refer to this third group of those who are not working and not looking for work as **out of the labor force** or not in the labor force.

The U.S. unemployment rate, which is based on a monthly survey carried out by the U.S. Bureau of the Census, asks a series of questions to divide the adult population into employed, unemployed, or not in the labor force. To be classified as unemployed, a person must be without a job, currently available to work, and actively looking for work in the previous four weeks. Thus, a person who does not have a job but who is not currently available to work or has not actively looked for work in the last four weeks is counted as out of the labor force.

Employed: currently working for pay

Unemployed: Out of work and actively looking for a job

Out of the labor force: Out of paid work and not actively looking for a job

Labor force: the number of employed plus the unemployed

Calculating the Unemployment Rate

Figure 8.2 shows the three-way division of the 16-and-over population. In January 2017, about 62.9% of the adult population was "in the labor force"; that is, people are either employed or without a job but looking for work. We can divide those in the labor force into the employed and the unemployed. **Table 8.1** shows those values. The **unemployment rate** is not the percentage of the total adult population without jobs, but rather the percentage of adults who are in the labor force but who do not have jobs:

$$\text{Unemployment rate} = \frac{\text{Unemployed people}}{\text{Total labor force}} \times 100$$

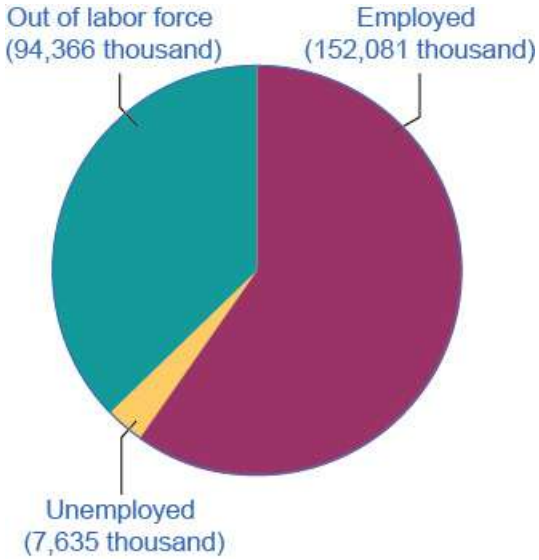


Figure 8.2 Employed, Unemployed, and Out of the Labor Force Distribution of Adult Population (age 16 and older), January 2017 The total adult, working-age population in January 2017 was 254.1 million. Out of this total population, 152.1 were classified as employed, and 7.6 million were classified as unemployed. The remaining 94.4 were classified as out of the labor force. As you will learn, however, this seemingly simple chart does not tell the whole story.

Total adult population over the age of 16	254.082 million
In the labor force	159.716 million (62.9%)
Employed	152.081 million
Unemployed	7.635 million
Out of the labor force	94.366 million (37.1%)

Table 8.1 U.S. Employment and Unemployment, January 2017 (Source: https://data.bls.gov)

In this example, we can calculate the unemployment rate as 7.635 million unemployed people divided by 159.716 million people in the labor force, which works out to a 4.8% rate of unemployment. The following Work It Out feature will walk you through the steps of this calculation.

Work It Out

Calculating Labor Force Percentages

How do economists arrive at the percentages in and out of the labor force and the unemployment rate? We will use the values in **Table 8.1** to illustrate the steps.

To determine the percentage in the labor force:

Step 1. Divide the number of people in the labor force (159.716 million) by the total adult (working-age) population (254.082 million).

Step 2. Multiply by 100 to obtain the percentage.

$$\begin{aligned}\text{Percentage in the labor force} &= \frac{159.716}{254.082}\\ &= 0.6286\\ &= 62.9\%\end{aligned}$$

To determine the percentage out of the labor force:

Step 1. Divide the number of people out the labor force (94.366 million) by the total adult (working-age) population (254.082 million).

Step 2. Multiply by 100 to obtain the percentage.

$$\begin{aligned}\text{Percentage in the labor force} &= \frac{94.366}{254.082}\\ &= 0.3714\\ &= 37.1\%\end{aligned}$$

To determine the unemployment rate:

Step 1. Divide the number of unemployed people (7.635 million) by the total labor force (157 million).

Step 2. Multiply by 100 to obtain the rate.

$$\begin{aligned}\text{Unemployment rate} &= \frac{7.635}{159.716}\\ &= 0.0478\\ &= 4.8\%\end{aligned}$$

Hidden Unemployment

Even with the "out of the labor force" category, there are still some people who are mislabeled in the categorization of employed, unemployed, or out of the labor force. There are some people who have only part time or temporary jobs, and they are looking for full time and permanent employment that are counted as employed, although they are not employed in the way they would like or need to be. Additionally, there are individuals who are **underemployed**. This includes those who are trained or skilled for one type or level of work but are working in a lower paying job or one that does not utilize their skills. For example, we would consider an individual with a college degree in finance who is working as a sales clerk underemployed. They are, however, also counted in the employed group. All of these individuals fall under the umbrella of the term "hidden unemployment." **Discouraged workers**, those who have stopped looking for employment and, hence, are no longer counted in the unemployed also fall into this group

Labor Force Participation Rate

Another important statistic is the **labor force participation rate**. This is the percentage of adults in an economy who are either employed or who are unemployed and looking for a job. Using the data in **Figure 8.2** and **Table 8.1**, those included in this calculation would be the 159.716 million individuals in the labor force. We calculate the rate by taking the number of people in the labor force, that is, the number employed and the number unemployed, divided by the total adult population and multiplying by 100 to get the percentage. For the data from January 2017, the labor force participation rate is 62.9%. Historically, the civilian labor force participation rate in the United States climbed beginning in the 1960s as women increasingly entered the workforce, and it peaked at just over 67% in late 1999 to early 2000. Since then, the labor force participation rate has steadily declined, slowly to about 66% in 2008, early in the Great Recession, and then more rapidly during and after that recession, reaching its present level, where it has remained stable, near the end of 2013.

The Establishment Payroll Survey

When the unemployment report comes out each month, the Bureau of Labor Statistics (BLS) also reports on the number of jobs created—which comes from the establishment payroll survey. The payroll survey is based on a survey of about 147,000 businesses and government agencies throughout the United States. It generates payroll employment estimates by the following criteria: all employees, average weekly hours worked, and average hourly, weekly, and overtime earnings. One of the criticisms of this survey is that it does not count the self-employed. It also does not make a distinction between new, minimum wage, part time or temporary jobs and full time jobs with "decent" pay.

How Does the U.S. Bureau of Labor Statistics Collect the U.S.

Unemployment Data?

The unemployment rate announced by the U.S. Bureau of Labor Statistics on the first Friday of each month for the previous month is based on the Current Population Survey (CPS), which the Bureau has carried out every month since 1940. The Bureau takes great care to make this survey representative of the country as a whole. The country is first divided into 3,137 areas. The U.S. Bureau of the Census then selects 729 of these areas to survey. It divides the 729 areas into districts of about 300 households each, and divides each district into clusters of about four dwelling units. Every month, Census Bureau employees call about 15,000 of the four-household clusters, for a total of 60,000 households. Employees interview households for four consecutive months, then rotate them out of the survey for eight months, and then interview them again for the same four months the following year, before leaving the sample permanently.

Based on this survey, state, industry, urban and rural areas, gender, age, race or ethnicity, and level of education statistics comprise components that contribute to unemployment rates. A wide variety of other information is available, too. For example, how long have people been unemployed? Did they become unemployed because they quit, or were laid off, or their employer went out of business? Is the unemployed person the only wage earner in the family? The Current Population Survey is a treasure trove of information about employment and unemployment. If you are wondering what the difference is between the CPS and EPS, read the following Clear it Up feature.

Clear It Up

What is the difference between CPS and EPS?

The United States Census Bureau conducts the Current Population Survey (CPS), which measures the percentage of the labor force that is unemployed. The Bureau of Labor Statistics' establishment payroll survey (EPS) is a payroll survey that measures the net change in jobs created for the month.

Criticisms of Measuring Unemployment

There are always complications in measuring the number of unemployed. For example, what about people who do not have jobs and would be available to work, but are discouraged by the lack of available jobs in their area and stopped looking? Such people, and their families, may be suffering the pains of unemployment. However, the survey counts them as out of the labor force because they are not actively looking for work. Other people may tell the Census Bureau that they are ready to work and looking for a job but, truly, they are not that eager to work and are not looking very hard at all. They are counted as unemployed, although they might more accurately be classified as out of the labor force. Still other people may have a job, perhaps doing something like yard work, child care, or cleaning houses, but are not reporting the income earned to the tax authorities. They may report being unemployed, when they actually are working.

Although the unemployment rate gets most of the public and media attention, economic researchers at the Bureau of Labor Statistics publish a wide array of surveys and reports that try to measure these kinds of issues and to develop a more nuanced and complete view of the labor market. It is not exactly a hot news flash that economic statistics are imperfect. Even imperfect measures like the unemployment rate, however, can still be quite informative, when interpreted knowledgeably and sensibly.

Link It Up

Click here (http://openstaxcollege.org/l/BLS_CPS) to learn more about the CPS and to read frequently asked questions about employment and labor.

8.2 | Patterns of Unemployment

By the end of this section, you will be able to:

- Explain historical patterns of unemployment in the U.S.
- Identify trends of unemployment based on demographics
- Evaluate global unemployment rates

Let's look at how unemployment rates have changed over time and how various groups of people are affected by unemployment differently.

The Historical U.S. Unemployment Rate

Figure 8.3 shows the historical pattern of U.S. unemployment since 1955.

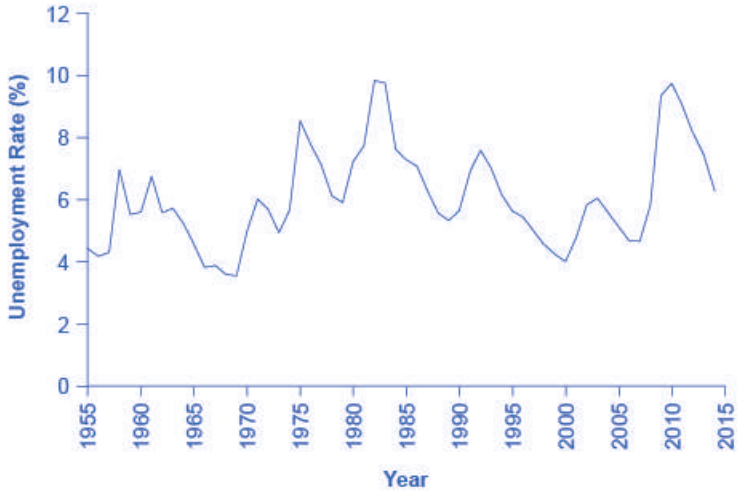


Figure 8.3 The U.S. Unemployment Rate, 1955–2015 The U.S. unemployment rate moves up and down as the economy moves in and out of recessions. However, over time, the unemployment rate seems to return to a range of 4% to 6%. There does not seem to be a long-term trend toward the rate moving generally higher or generally lower. (Source: Federal Reserve Economic Data (FRED) https://research.stlouisfed.org/fred2/series/LRUN64TTUSA156S0)

As we look at this data, several patterns stand out:

1. Unemployment rates do fluctuate over time. During the deep recessions of the early 1980s and of 2007–2009, unemployment reached roughly 10%. For comparison, during the 1930s Great Depression, the unemployment rate reached almost 25% of the labor force.
2. Unemployment rates in the late 1990s and into the mid-2000s were rather low by historical standards. The unemployment rate was below 5% from 1997 to 2000, and near 5% during almost all of 2006–2007, and 5% or slightly less from September 2015 through January 2017 (the latest date for which data are available as of this writing). The previous time unemployment had been less than 5% for three consecutive years was three

decades earlier, from 1968 to 1970.

3. The unemployment rate never falls all the way to zero. It almost never seems to get below 3%—and it stays that low only for very short periods. (We discuss reasons why this is the case later in this chapter.)
4. The timing of rises and falls in unemployment matches fairly well with the timing of upswings and downswings in the overall economy, except that unemployment tends to lag changes in economic activity, and especially so during upswings of the economy following a recession. During periods of recession and depression, unemployment is high. During periods of economic growth, unemployment tends to be lower.
5. No significant upward or downward trend in unemployment rates is apparent. This point is especially worth noting because the U.S. population more than quadrupled from 76 million in 1900 to over 324 million by 2017. Moreover, a higher proportion of U.S. adults are now in the paid workforce, because women have entered the paid labor force in significant numbers in recent decades. Women comprised 18% of the paid workforce in 1900 and nearly half of the paid workforce in 2017. However, despite the increased number of workers, as well as other economic events like globalization and the continuous invention of new technologies, the economy has provided jobs without causing any long-term upward or downward trend in unemployment rates.

Unemployment Rates by Group

Unemployment is not distributed evenly across the U.S. population. **Figure 8.4** shows unemployment rates broken down in various ways: by gender, age, and race/ethnicity.

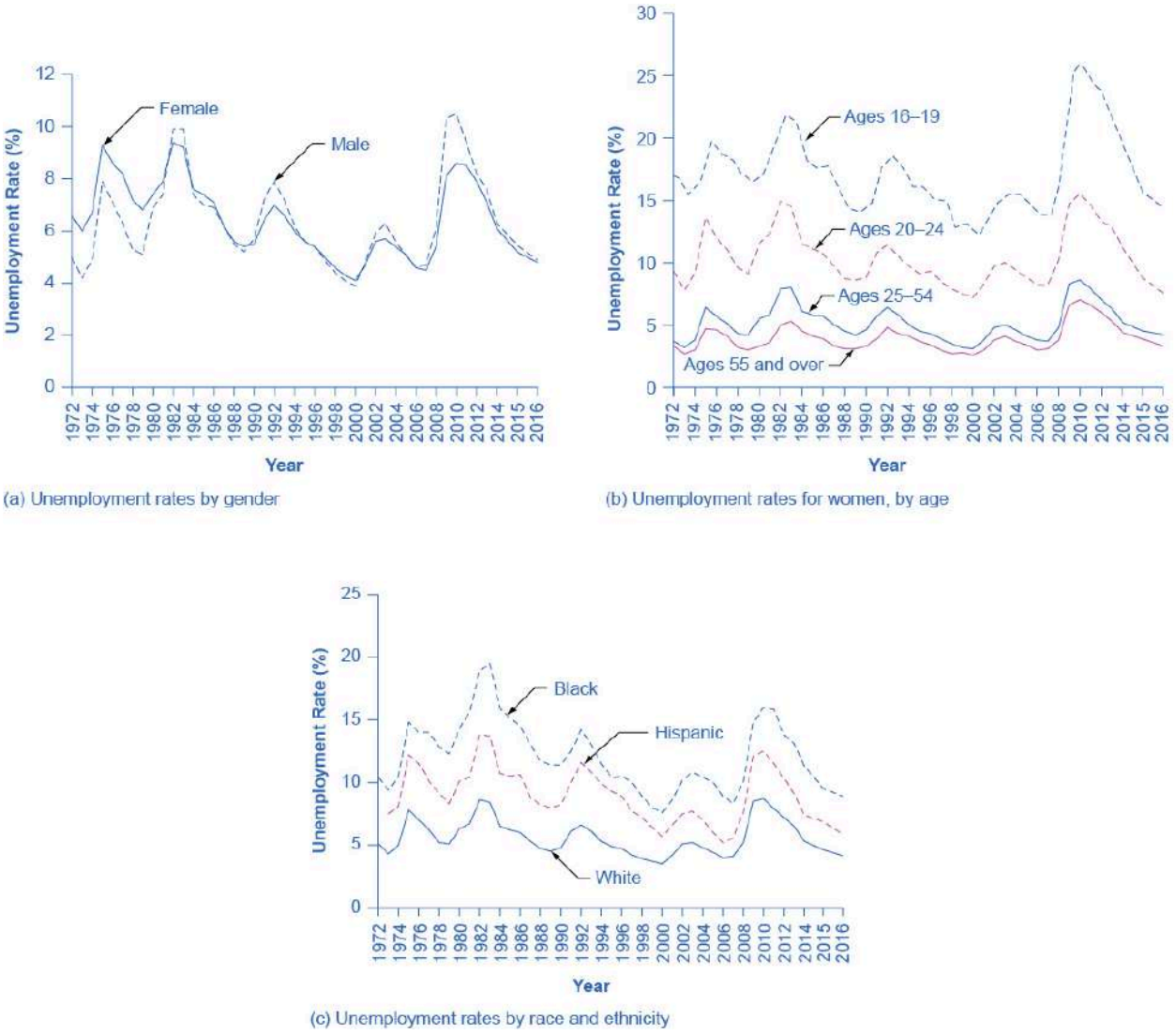


Figure 8.4 Unemployment Rate by Demographic Group (a) By gender, 1972–2016. Unemployment rates for men used to be lower than unemployment rates for women, but in recent decades, the two rates have been very close, often– and especially during and soon after the Great Recession – with the unemployment rate for men somewhat higher. (b) By age, 1972–2016. Unemployment rates are highest for the very young and become lower with age. (c) By race and ethnicity, 1972–2016. Although unemployment rates for all groups tend to rise and fall together, the unemployment rate for blacks is typically about twice as high as that for whites, while the unemployment rate for Hispanics is in between. (Source: www.bls.gov)

The unemployment rate for women had historically tended to be higher than the unemployment rate for men, perhaps reflecting the historical pattern that women were seen as "secondary" earners. By about 1980, however, the unemployment rate for women was essentially the same as that for men, as **Figure 8.4** (a) shows. During the 2008-2009 recession and in the immediate aftermath, the unemployment rate for men exceeded the unemployment rate for women. Subsequently, however, the gap has narrowed.

Link It Up

Read this report (http://openstaxcollege.org/l/BLS_recession) for detailed information on the 2008–2009 recession. It also provides some very useful information on the statistics of unemployment.

Younger workers tend to have higher unemployment, while middle-aged workers tend to have lower unemployment, probably because the middle-aged workers feel the responsibility of needing to have a job more heavily. Younger workers move in and out of jobs more than middle-aged workers, as part of the process of matching of workers and jobs, and this contributes to their higher unemployment rates. In addition, middle-aged workers are more likely to feel the responsibility of needing to have a job more heavily. Elderly workers have extremely low rates of unemployment, because those who do not have jobs often exit the labor force by retiring, and thus are not counted in the unemployment statistics. **Figure 8.4** (b) shows unemployment rates for women divided by age. The pattern for men is similar.

The unemployment rate for African-Americans is substantially higher than the rate for other racial or ethnic groups, a fact that surely reflects, to some extent, a pattern of discrimination that has constrained blacks' labor market opportunities. However, the gaps between unemployment rates for whites and for blacks and Hispanics diminished in the 1990s, as **Figure 8.4** (c) shows. In fact, unemployment rates for blacks and Hispanics were at the lowest levels for several decades in the mid-2000s before rising during the recent Great Recession.

Finally, those with less education typically suffer higher unemployment. In January 2017, for example, the unemployment rate for those with a college degree was 2.5%; for those with some college but not a four year degree, the unemployment rate was 3.8%; for high school graduates with no additional degree, the unemployment rate was 5.3%; and for those without a high school diploma, the unemployment rate was 7.7%. This pattern arises because additional education typically offers better connections to the labor market and higher demand. With less attractive labor market opportunities for low-skilled workers compared to the opportunities for the more highly-skilled, including lower pay, low-skilled workers may be less motivated to find jobs.

Breaking Down Unemployment in Other Ways

The Bureau of Labor Statistics also gives information about the reasons for unemployment, as well as the length of time individuals have been unemployed. **Table 8.2**, for example, shows the four reasons for unemployment and the percentages of the currently unemployed that fall into each category. **Table 8.3** shows the length of unemployment. For both of these, the data is from January 2017.(bls.gov)

Reason	Percentage
New Entrants	10.8%
Re-entrants	28.7%
Job Leavers	11.4%
Job Losers: Temporary	14.0%
Job Losers: Non Temporary	35.1%

Table 8.2 Reasons for Unemployment, January 2017

Length of Time	Percentage
Under 5 weeks	32.5%
5 to 14 weeks	27.5%
15 to 26 weeks	15.7%
Over 27 weeks	27.4%

Table 8.3 Length of Unemployment, January 2017

Link It Up

Watch this speech (http://openstaxcollege.org/l/droids) on the impact of droids on the labor market.

International Unemployment Comparisons

From an international perspective, the U.S. unemployment rate typically has looked a little better than average. **Table 8.4** compares unemployment rates for 1991, 1996, 2001, 2006 (just before the recession), and 2012 (somewhat after the recession) from several other high-income countries.

Country	1991	1996	2001	2006	2012
United States	6.8%	5.4%	4.8%	4.4%	8.1%
Canada	9.8%	8.8%	6.4%	6.2%	6.3%
Japan	2.1%	3.4%	5.1%	4.5%	3.9%
France	9.5%	12.5%	8.7%	10.1%	10.0%
Germany	5.6%	9.0%	8.9%	9.8%	5.5%
Italy	6.9%	11.7%	9.6%	7.8%	10.8%
Sweden	3.1%	9.9%	5.0%	5.2%	7.9%
United Kingdom	8.8%	8.1%	5.1%	5.5%	8.0%

Table 8.4 International Comparisons of Unemployment Rates

However, we need to treat cross-country comparisons of unemployment rates with care, because each country has slightly different definitions of unemployment, survey tools for measuring unemployment, and also different labor markets. For example, Japan's unemployment rates appear quite low, but Japan's economy has been mired in slow growth and recession since the late 1980s, and Japan's unemployment rate probably paints too rosy a picture of its

labor market. In Japan, workers who lose their jobs are often quick to exit the labor force and not look for a new job, in which case they are not counted as unemployed. In addition, Japanese firms are often quite reluctant to fire workers, and so firms have substantial numbers of workers who are on reduced hours or officially employed, but doing very little. We can view this Japanese pattern as an unusual method for society to provide support for the unemployed, rather than a sign of a healthy economy.

Link It Up

We hear about the Chinese economy in the news all the time. The value of the Chinese yuan in comparison to the U.S. dollar is likely to be part of the nightly business report, so why is the Chinese economy not included in this discussion of international unemployment? The lack of reliable statistics is the reason. This article (http://openstaxcollege.org/l/ChinaEmployment) explains why.

Comparing unemployment rates in the United States and other high-income economies with unemployment rates in Latin America, Africa, Eastern Europe, and Asia is very difficult. One reason is that the statistical agencies in many poorer countries lack the resources and technical capabilities of the U.S. Bureau of the Census. However, a more difficult problem with international comparisons is that in many low-income countries, most workers are not involved in the labor market through an employer who pays them regularly. Instead, workers in these countries are engaged in short-term work, subsistence activities, and barter. Moreover, the effect of unemployment is very different in high-income and low-income countries. Unemployed workers in the developed economies have access to various government programs like unemployment insurance, welfare, and food stamps. Such programs may barely exist in poorer countries. Although unemployment is a serious problem in many low-income countries, it manifests itself in a different way than in high-income countries.

8.3 | What Causes Changes in Unemployment over the Short Run

By the end of this section, you will be able to:

- Analyze cyclical unemployment
- Explain the relationship between sticky wages and employment using various economic arguments
- Apply supply and demand models to unemployment and wages

We have seen that unemployment varies across times and places. What causes changes in unemployment? There are different answers in the short run and in the long run. Let's look at the short run first.

Cyclical Unemployment

Let’s make the plausible assumption that in the short run, from a few months to a few years, the quantity of hours that the average person is willing to work for a given wage does not change much, so the labor supply curve does not shift much. In addition, make the standard *ceteris paribus* assumption that there is no substantial short-term change in the age structure of the labor force, institutions and laws affecting the labor market, or other possibly relevant factors.

One primary determinant of the demand for labor from firms is how they perceive the state of the macro economy. If firms believe that business is expanding, then at any given wage they will desire to hire a greater quantity of labor, and

the labor demand curve shifts to the right. Conversely, if firms perceive that the economy is slowing down or entering a recession, then they will wish to hire a lower quantity of labor at any given wage, and the labor demand curve will shift to the left. Economists call the variation in unemployment that the economy causes moving from expansion to recession or from recession to expansion (i.e. the business cycle) **cyclical unemployment**.

From the standpoint of the supply-and-demand model of competitive and flexible labor markets, unemployment represents something of a puzzle. In a supply-and-demand model of a labor market, as **Figure 8.5** illustrates, the labor market should move toward an equilibrium wage and quantity. At the equilibrium wage (We), the equilibrium quantity (Qe) of labor supplied by workers should be equal to the quantity of labor demanded by employers.

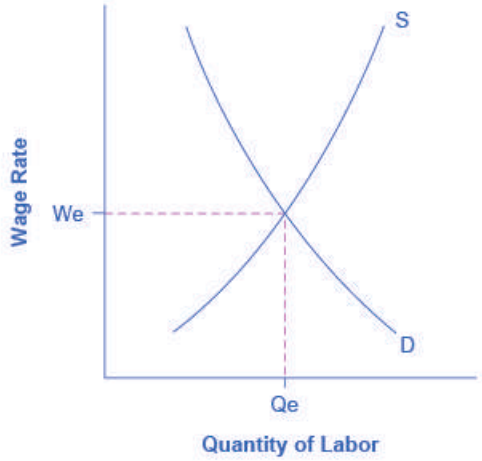


Figure 8.5 The Unemployment and Equilibrium in the Labor Market In a labor market with flexible wages, the equilibrium will occur at wage We and quantity Qe, where the number of people who want jobs (shown by S) equals the number of jobs available (shown by D).

One possibility for unemployment is that people who are unemployed are those who are not willing to work at the current equilibrium wage, say $10 an hour, but would be willing to work at a higher wage, like $20 per hour. The monthly Current Population Survey would count these people as unemployed, because they say they are ready and looking for work (at $20 per hour). However, from an economist's perspective, these people are choosing to be unemployed.

Probably a few people are unemployed because of unrealistic expectations about wages, but they do not represent the majority of the unemployed. Instead, unemployed people often have friends or acquaintances of similar skill levels who are employed, and the unemployed would be willing to work at the jobs and wages similar to what those people are receiving. However, the employers of their friends and acquaintances do not seem to be hiring. In other words, these people are involuntarily unemployed. What causes involuntary unemployment?

Why Wages Might Be Sticky Downward

If a labor market model with flexible wages does not describe unemployment very well—because it predicts that anyone willing to work at the going wage can always find a job—then it may prove useful to consider economic models in which wages are not flexible or adjust only very slowly. In particular, even though wage increases may occur with relative ease, wage decreases are few and far between.

One set of reasons why wages may be "sticky downward," as economists put it, involves economic laws and institutions. For low-skilled workers receiving minimum wage, it is illegal to reduce their wages. For union workers operating under a multiyear contract with a company, wage cuts might violate the contract and create a labor dispute or a strike. However, minimum wages and union contracts are not a sufficient reason why wages would be sticky downward for the U.S. economy as a whole. After all, out of the 150 million or so employed workers in the U.S. economy, only about 2.6 million—less than 2% of the total—do not receive compensation above the minimum wage. Similarly, labor unions represent only about 11% of American wage and salary workers. In other high-income countries, more workers may have their wages determined by unions or the minimum wage may be set at a level that applies to a larger share of workers. However, for the United States, these two factors combined affect only about 15% or less of the labor force.

Economists looking for reasons why wages might be sticky downwards have focused on factors that may characterize most labor relationships in the economy, not just a few. Many have proposed a number of different theories, but they share a common tone.

One argument is that even employees who are not union members often work under an **implicit contract**, which is that the employer will try to keep wages from falling when the economy is weak or the business is having trouble, and the employee will not expect huge salary increases when the economy or the business is strong. This wage-setting behavior acts like a form of insurance: the employee has some protection against wage declines in bad times, but pays for that protection with lower wages in good times. Clearly, this sort of implicit contract means that firms will be hesitant to cut wages, lest workers feel betrayed and work less hard or even leave the firm.

Efficiency wage theory argues that workers' productivity depends on their pay, and so employers will often find it worthwhile to pay their employees somewhat more than market conditions might dictate. One reason is that employees who receive better pay than others will be more productive because they recognize that if they were to lose their current jobs, they would suffer a decline in salary. As a result, they are motivated to work harder and to stay with the current employer. In addition, employers know that it is costly and time-consuming to hire and train new employees, so they would prefer to pay workers a little extra now rather than to lose them and have to hire and train new workers. Thus, by avoiding wage cuts, the employer minimizes costs of training and hiring new workers, and reaps the benefits of well-motivated employees.

The **adverse selection of wage cuts argument** points out that if an employer reacts to poor business conditions by reducing wages for all workers, then the best workers, those with the best employment alternatives at other firms, are the most likely to leave. The least attractive workers, with fewer employment alternatives, are more likely to stay. Consequently, firms are more likely to choose which workers should depart, through layoffs and firings, rather than trimming wages across the board. Sometimes companies that are experiencing difficult times can persuade workers to take a pay cut for the short term, and still retain most of the firm’s workers. However, it is far more typical for companies to lay off some workers, rather than to cut wages for everyone.

The **insider-outsider model** of the labor force, in simple terms, argues that those already working for firms are “insiders,” while new employees, at least for a time, are “outsiders.” A firm depends on its insiders to keep the organization running smoothly, to be familiar with routine procedures, and to train new employees. However, cutting wages will alienate the insiders and damage the firm’s productivity and prospects.

Finally, the **relative wage coordination argument** points out that even if most workers were hypothetically willing to see a decline in their own wages in bad economic times as long as everyone else also experiences such a decline, there is no obvious way for a decentralized economy to implement such a plan. Instead, workers confronted with the possibility of a wage cut will worry that other workers will not have such a wage cut, and so a wage cut means being worse off both in absolute terms and relative to others. As a result, workers fight hard against wage cuts.

These theories of why wages tend not to move downward differ in their logic and their implications, and figuring out the strengths and weaknesses of each theory is an ongoing subject of research and controversy among economists. All tend to imply that wages will decline only very slowly, if at all, even when the economy or a business is having tough times. When wages are inflexible and unlikely to fall, then either short-run or long-run unemployment can result. **Figure 8.6** illustrates this.

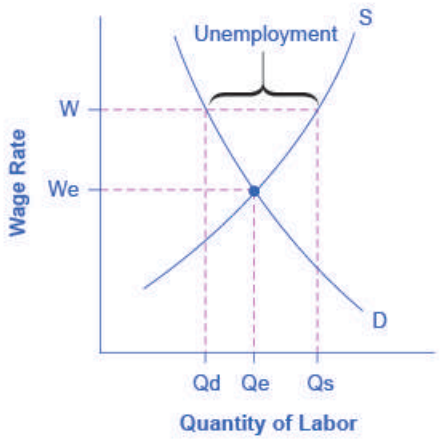


Figure 8.6 Sticky Wages in the Labor Market Because the wage rate is stuck at W, above the equilibrium, the number of those who want jobs (Qs) is greater than the number of job openings (Qd). The result is unemployment, shown by the bracket in the figure.

Figure 8.7 shows the interaction between shifts in labor demand and wages that are sticky downward. **Figure 8.7** (a) illustrates the situation in which the demand for labor shifts to the right from D_0 to D_1. In this case, the equilibrium wage rises from W_0 to W_1 and the equilibrium quantity of labor hired increases from Q_0 to Q_1. It does not hurt employee morale at all for wages to rise.

Figure 8.7 (b) shows the situation in which the demand for labor shifts to the left, from D_0 to D_1, as it would tend to do in a recession. Because wages are sticky downward, they do not adjust toward what would have been the new equilibrium wage (W_1), at least not in the short run. Instead, after the shift in the labor demand curve, the same quantity of workers is willing to work at that wage as before; however, the quantity of workers demanded at that wage has declined from the original equilibrium (Q_0) to Q_2. The gap between the original equilibrium quantity (Q_0) and the new quantity demanded of labor (Q_2) represents workers who would be willing to work at the going wage but cannot find jobs. The gap represents the economic meaning of unemployment.

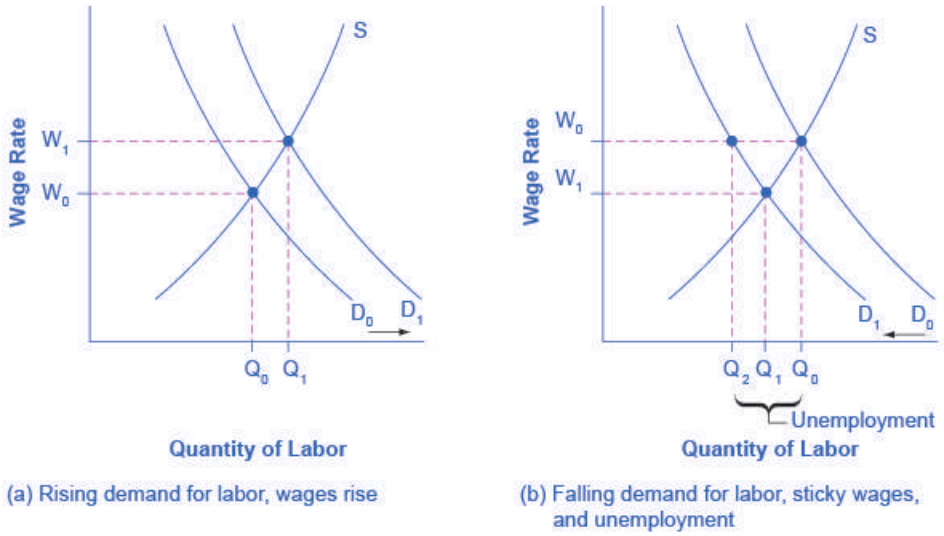


Figure 8.7 Rising Wage and Low Unemployment: Where Is the Unemployment in Supply and Demand? (a) In a labor market where wages are able to rise, an increase in the demand for labor from D_0 to D_1 leads to an increase in equilibrium quantity of labor hired from Q_0 to Q_1 and a rise in the equilibrium wage from W_0 to W_1. (b) In a labor market where wages do not decline, a fall in the demand for labor from D_0 to D_1 leads to a decline in the quantity of labor demanded at the original wage (W_0) from Q_0 to Q_2. These workers will want to work at the prevailing wage (W_0), but will not be able to find jobs.

This analysis helps to explain the connection that we noted earlier: that unemployment tends to rise in recessions and to decline during expansions. The overall state of the economy shifts the labor demand curve and, combined with wages that are sticky downwards, unemployment changes. The rise in unemployment that occurs because of a recession is cyclical unemployment.

Link It Up

The St. Louis Federal Reserve Bank is the best resource for macroeconomic time series data, known as the Federal Reserve Economic Data (FRED). FRED (http://openstaxcollege.org/l/FRED_employment) provides complete data sets on various measures of the unemployment rate as well as the monthly Bureau of Labor Statistics report on the results of the household and employment surveys.

8.4 | What Causes Changes in Unemployment over the Long Run

By the end of this section, you will be able to:

- Explain frictional and structural unemployment
- Assess relationships between the natural rate of employment and potential real GDP, productivity, and public policy
- Identify recent patterns in the natural rate of employment
- Propose ways to combat unemployment

Cyclical unemployment explains why unemployment rises during a recession and falls during an economic expansion, but what explains the remaining level of unemployment even in good economic times? Why is the unemployment rate never zero? Even when the U.S. economy is growing strongly, the unemployment rate only rarely dips as low as 4%. Moreover, the discussion earlier in this chapter pointed out that unemployment rates in many European countries like Italy, France, and Germany have often been remarkably high at various times in the last few decades. Why does some level of unemployment persist even when economies are growing strongly? Why are unemployment rates continually higher in certain economies, through good economic years and bad? Economists have a term to describe the remaining level of unemployment that occurs even when the economy is healthy: they call it the **natural rate of unemployment**.

The Long Run: The Natural Rate of Unemployment

The natural rate of unemployment is not “natural” in the sense that water freezes at 32 degrees Fahrenheit or boils at 212 degrees Fahrenheit. It is not a physical and unchanging law of nature. Instead, it is only the “natural” rate because it is the unemployment rate that would result from the combination of economic, social, and political factors that exist at a time—assuming the economy was neither booming nor in recession. These forces include the usual pattern of companies expanding and contracting their workforces in a dynamic economy, social and economic forces that affect the labor market, or public policies that affect either the eagerness of people to work or the willingness of businesses to hire. Let’s discuss these factors in more detail.

Frictional Unemployment

In a market economy, some companies are always going broke for a variety of reasons: old technology; poor management; good management that happened to make bad decisions; shifts in tastes of consumers so that less of the firm's product is desired; a large customer who went broke; or tough domestic or foreign competitors. Conversely, other companies will be doing very well for just the opposite reasons and looking to hire more employees. In a perfect world, all of those who lost jobs would immediately find new ones. However, in the real world, even if the number of job seekers is equal to the number of job vacancies, it takes time to find out about new jobs, to interview and figure out if the new job is a good match, or perhaps to sell a house and buy another in proximity to a new job. Economists call the unemployment that occurs in the meantime, as workers move between jobs, **frictional unemployment**. Frictional unemployment is not inherently a bad thing. It takes time on part of both the employer and the individual to match those looking for employment with the correct job openings. For individuals and companies to be successful and productive, you want people to find the job for which they are best suited, not just the first job offered.

In the mid-2000s, before the 2008–2009 recession, it was true that about 7% of U.S. workers saw their jobs disappear in any three-month period. However, in periods of economic growth, these destroyed jobs are counterbalanced for the economy as a whole by a larger number of jobs created. In 2005, for example, there were typically about 7.5 million unemployed people at any given time in the U.S. economy. Even though about two-thirds of those unemployed people found a job in 14 weeks or fewer, the unemployment rate did not change much during the year, because those who found new jobs were largely offset by others who lost jobs.

Of course, it would be preferable if people who were losing jobs could immediately and easily move into newly created jobs, but in the real world, that is not possible. Someone who is laid off by a textile mill in South Carolina cannot turn around and immediately start working for a textile mill in California. Instead, the adjustment process happens in ripples. Some people find new jobs near their old ones, while others find that they must move to new locations. Some people can do a very similar job with a different company, while others must start new career paths. Some people may be near retirement and decide to look only for part-time work, while others want an employer that offers a long-term career path. The frictional unemployment that results from people moving between jobs in a dynamic economy may account for one to two percentage points of total unemployment.

The level of frictional unemployment will depend on how easy it is for workers to learn about alternative jobs, which may reflect the ease of communications about job prospects in the economy. The extent of frictional unemployment will also depend to some extent on how willing people are to move to new areas to find jobs—which in turn may depend on history and culture.

Frictional unemployment and the natural rate of unemployment also seem to depend on the age distribution of the population. **Figure 8.4** (b) showed that unemployment rates are typically lower for people between 25–54 years of age or aged 55 and over than they are for those who are younger. "Prime-age workers," as those in the 25–54 age bracket are sometimes called, are typically at a place in their lives when they want to have a job and income arriving at all times. In addition, older workers who lose jobs may prefer to opt for retirement. By contrast, it is likely that a relatively high proportion of those who are under 25 will be trying out jobs and life options, and this leads to greater job mobility and hence higher frictional unemployment. Thus, a society with a relatively high proportion of young workers, like the U.S. beginning in the mid-1960s when Baby Boomers began entering the labor market, will tend to have a higher unemployment rate than a society with a higher proportion of its workers in older ages.

Structural Unemployment

Another factor that influences the natural rate of unemployment is the amount of **structural unemployment**. The structurally unemployed are individuals who have no jobs because they lack skills valued by the labor market, either because demand has shifted away from the skills they do have, or because they never learned any skills. An example of the former would be the unemployment among aerospace engineers after the U.S. space program downsized in the 1970s. An example of the latter would be high school dropouts.

Some people worry that technology causes structural unemployment. In the past, new technologies have put lower skilled employees out of work, but at the same time they create demand for higher skilled workers to use the new technologies. Education seems to be the key in minimizing the amount of structural unemployment. Individuals who have degrees can be retrained if they become structurally unemployed. For people with no skills and little education, that option is more limited.

Natural Unemployment and Potential Real GDP

The natural unemployment rate is related to two other important concepts: full employment and potential real GDP. Economists consider the economy to be at full employment when the actual unemployment rate is equal to the natural unemployment rate. When the economy is at full employment, real GDP is equal to potential real GDP. By contrast, when the economy is below full employment, the unemployment rate is greater than the natural unemployment rate and real GDP is less than potential. Finally, when the economy is above full employment, then the unemployment rate is less than the natural unemployment rate and real GDP is greater than potential. Operating above potential is only possible for a short while, since it is analogous to all workers working overtime.

Productivity Shifts and the Natural Rate of Unemployment

Unexpected shifts in productivity can have a powerful effect on the natural rate of unemployment. Over time, workers' productivity determines the level of wages in an economy. After all, if a business paid workers more than could be justified by their productivity, the business will ultimately lose money and go bankrupt. Conversely, if a business tries to pay workers less than their productivity then, in a competitive labor market, other businesses will find it worthwhile to hire away those workers and pay them more.

However, adjustments of wages to productivity levels will not happen quickly or smoothly. Employers typically review wages only once or twice a year. In many modern jobs, it is difficult to measure productivity at the individual level. For example, how precisely would one measure the quantity produced by an accountant who is one of many people working in the tax department of a large corporation? Because productivity is difficult to observe, employers often determine wage increases based on recent experience with productivity. If productivity has been rising at, say, 2% per year, then wages rise at that level as well. However, when productivity changes unexpectedly, it can affect the natural rate of unemployment for a time.

The U.S. economy in the 1970s and 1990s provides two vivid examples of this process. In the 1970s, productivity growth slowed down unexpectedly (as we discussed in **Economic Growth**). For example, output per hour of U.S. workers in the business sector increased at an annual rate of 3.3% per year from 1960 to 1973, but only 0.8% from 1973 to 1982. **Figure 8.8** (a) illustrates the situation where the demand for labor—that is, the quantity of labor that business is willing to hire at any given wage—has been shifting out a little each year because of rising productivity, from D_0 to D_1 to D_2. As a result, equilibrium wages have been rising each year from W_0 to W_1 to W_2. However, when productivity unexpectedly slows down, the pattern of wage increases does not adjust right away. Wages keep rising each year from W_2 to W_3 to W_4, but the demand for labor is no longer shifting up. A gap opens where the quantity of labor supplied at wage level W_4 is greater than the quantity demanded. The natural rate of unemployment rises. In the aftermath of this unexpectedly low productivity in the 1970s, the national unemployment rate did not fall below 7% from May, 1980 until 1986. Over time, the rise in wages will adjust to match the slower gains in productivity, and the unemployment rate will ease back down, but this process may take years.

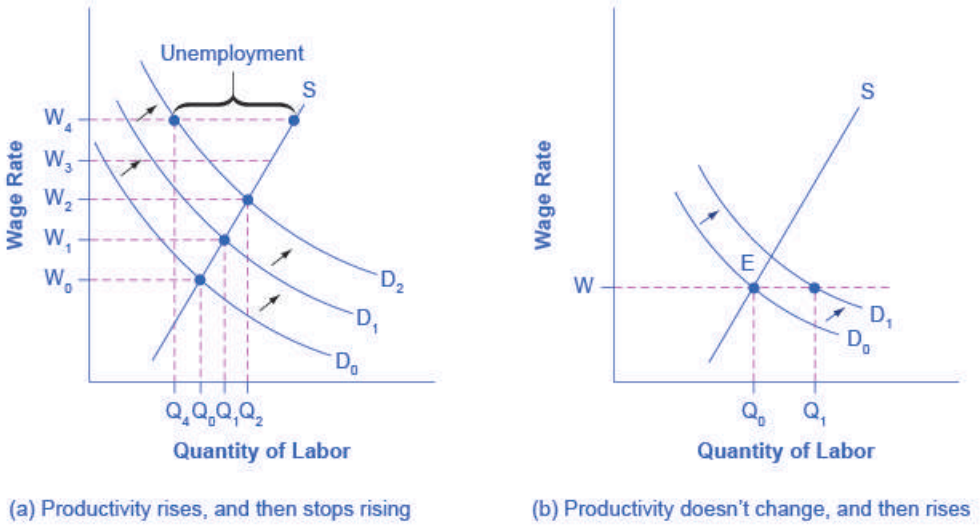


Figure 8.8 Unexpected Productivity Changes and Unemployment (a) Productivity is rising, increasing the demand for labor. Employers and workers become used to the pattern of wage increases. Then productivity suddenly stops increasing. However, the expectations of employers and workers for wage increases do not shift immediately, so wages keep rising as before. However, the demand for labor has not increased, so at wage W_4, unemployment exists where the quantity supplied of labor exceeds the quantity demanded. (b) The rate of productivity increase has been zero for a time, so employers and workers have come to accept the equilibrium wage level (W). Then productivity increases unexpectedly, shifting demand for labor from D_0 to D_1. At the wage (W), this means that the quantity demanded of labor exceeds the quantity supplied, and with job offers plentiful, the unemployment rate will be low.

The late 1990s provide an opposite example: instead of the surprise decline in productivity that occurred in the 1970s, productivity unexpectedly rose in the mid-1990s. The annual growth rate of real output per hour of labor increased from 1.7% from 1980–1995, to an annual rate of 2.6% from 1995–2001. Let's simplify the situation a bit, so that the economic lesson of the story is easier to see graphically, and say that productivity had not been increasing at all in earlier years, so the intersection of the labor market was at point E in **Figure 8.8** (b), where the demand curve for labor (D_0) intersects the supply curve for labor. As a result, real wages were not increasing. Now, productivity jumps upward, which shifts the demand for labor out to the right, from D_0 to D_1. At least for a time, however, wages are still set according to the earlier expectations of no productivity growth, so wages do not rise. The result is that at the prevailing wage level (W), the quantity of labor demanded (Qd) will for a time exceed the quantity of labor supplied (Qs), and unemployment will be very low—actually below the natural level of unemployment for a time. This pattern of unexpectedly high productivity helps to explain why the unemployment rate stayed below 4.5%—quite a low level by historical standards—from 1998 until after the U.S. economy had entered a recession in 2001.

Levels of unemployment will tend to be somewhat higher on average when productivity is unexpectedly low, and conversely, will tend to be somewhat lower on average when productivity is unexpectedly high. However, over time, wages do eventually adjust to reflect productivity levels.

Public Policy and the Natural Rate of Unemployment

Public policy can also have a powerful effect on the natural rate of unemployment. On the supply side of the labor market, public policies to assist the unemployed can affect how eager people are to find work. For example, if a worker who loses a job is guaranteed a generous package of unemployment insurance, welfare benefits, food stamps, and government medical benefits, then the opportunity cost of unemployment is lower and that worker will be less eager to seek a new job.

What seems to matter most is not just the amount of these benefits, but how long they last. A society that provides generous help for the unemployed that cuts off after, say, six months, may provide less of an incentive for unemployment than a society that provides less generous help that lasts for several years. Conversely, government assistance for job search or retraining can in some cases encourage people back to work sooner. See the Clear it Up to learn how the U.S. handles unemployment insurance.

Clear It Up

How does U.S. unemployment insurance work?

Unemployment insurance is a joint federal–state program that the federal government enacted in 1935. While the federal government sets minimum standards for the program, state governments conduct most of the administration.

The funding for the program is a federal tax collected from employers. The federal government requires tax collection on the first $7,000 in wages paid to each worker; however, states can choose to collect the tax on a higher amount if they wish, and 41 states have set a higher limit. States can choose the length of time that they pay benefits, although most states limit unemployment benefits to 26 weeks—with extensions possible in times of especially high unemployment. The states then use the fund to pay benefits to those who become unemployed. Average unemployment benefits are equal to about one-third of the wage that the person earned in his or her previous job, but the level of unemployment benefits varies considerably across states.

Bottom 10 States That Pay the Lowest Benefit per Week		Top 10 States That Pay the Highest Benefit per Week	
Delaware	$330	Massachusetts	$672
Georgia	$330	Minnesota	$683
South Carolina	$326	Washington	$681
Missouri	$320	New Jersey	$657
Florida	$275	North Dakota	$633
Tennessee	$275	Connecticut	$598
Alabama	$265	Oregon	$590
Louisiana	$247	Pennsylvania	$573
Arizona	$240	Colorado	$568
Mississippi	$235	Rhode Island	$566

Table 8.5 Maximum Weekly Unemployment Benefits by State in 2017 (Source: http://www.savingtoinvest.com/maximum-weekly-unemployment-benefits-by-state/)

One other interesting thing to note about the classifications of unemployment—an individual does not have to collect unemployment benefits to be classified as unemployed. While there are statistics kept and studied relating to how many people are collecting unemployment insurance, this is not the source of unemployment rate information.

Link It Up

View this article (http://openstaxcollege.org/l/NYT_Benefits) for an explanation of exactly who is eligible for unemployment benefits.

On the demand side of the labor market, government rules, social institutions, and the presence of unions can affect the willingness of firms to hire. For example, if a government makes it hard for businesses to start up or to expand, by wrapping new businesses in bureaucratic red tape, then businesses will become more discouraged about hiring. Government regulations can make it harder to start a business by requiring that a new business obtain many permits and pay many fees, or by restricting the types and quality of products that a company can sell. Other government regulations, like zoning laws, may limit where companies can conduct business, or whether businesses are allowed to be open during evenings or on Sunday.

Whatever defenses may be offered for such laws in terms of social value—like the value some Christians place on not working on Sunday, or Orthodox Jews or highly observant Muslims on Saturday—these kinds of restrictions impose a barrier between some willing workers and other willing employers, and thus contribute to a higher natural rate of unemployment. Similarly, if government makes it difficult to fire or lay off workers, businesses may react by trying not to hire more workers than strictly necessary—since laying these workers off would be costly and difficult. High minimum wages may discourage businesses from hiring low-skill workers. Government rules may encourage and support powerful unions, which can then push up wages for union workers, but at a cost of discouraging businesses from hiring those workers.

The Natural Rate of Unemployment in Recent Years

The underlying economic, social, and political factors that determine the natural rate of unemployment can change over time, which means that the natural rate of unemployment can change over time, too.

Estimates by economists of the natural rate of unemployment in the U.S. economy in the early 2000s run at about 4.5 to 5.5%. This is a lower estimate than earlier. We outline three of the common reasons that economists propose for this change below.

1. The internet has provided a remarkable new tool through which job seekers can find out about jobs at different companies and can make contact with relative ease. An internet search is far easier than trying to find a list of local employers and then hunting up phone numbers for all of their human resources departments, and requesting a list of jobs and application forms. Social networking sites such as LinkedIn have changed how people find work as well.
2. The growth of the temporary worker industry has probably helped to reduce the natural rate of unemployment. In the early 1980s, only about 0.5% of all workers held jobs through temp agencies. By the early 2000s, the figure had risen above 2%. Temp agencies can provide jobs for workers while they are looking for permanent work. They can also serve as a clearinghouse, helping workers find out about jobs with certain employers and getting a tryout with the employer. For many workers, a temp job is a stepping-stone to a permanent job that they might not have heard about or obtained any other way, so the growth of temp jobs will also tend to reduce frictional unemployment.
3. The aging of the “baby boom generation”—the especially large generation of Americans born between 1946 and 1964—meant that the proportion of young workers in the economy was relatively high in the 1970s, as the boomers entered the labor market, but is relatively low today. As we noted earlier, middle-aged and older workers are far more likely to experience low unemployment than younger workers, a factor that tends to reduce the natural rate of unemployment as the baby boomers age.

The combined result of these factors is that the natural rate of unemployment was on average lower in the 1990s and the early 2000s than in the 1980s. The 2008–2009 Great Recession pushed monthly unemployment rates up to 10% in late 2009. However, even at that time, the Congressional Budget Office was forecasting that by 2015, unemployment

rates would fall back to about 5%. During the last four months of 2015 the unemployment rate held steady at 5.0%. Throughout 2016 and up through January 2017, the unemployment rate has remained at or slightly below 5%. As of the first quarter of 2017, the Congressional Budget Office estimates the natural rate to be 4.74%, and the measured unemployment rate for January 2017 is 4.8%.

The Natural Rate of Unemployment in Europe

By the standards of other high-income economies, the natural rate of unemployment in the U.S. economy appears relatively low. Through good economic years and bad, many European economies have had unemployment rates hovering near 10%, or even higher, since the 1970s. European rates of unemployment have been higher not because recessions in Europe have been deeper, but rather because the conditions underlying supply and demand for labor have been different in Europe, in a way that has created a much higher natural rate of unemployment.

Many European countries have a combination of generous welfare and unemployment benefits, together with nests of rules that impose additional costs on businesses when they hire. In addition, many countries have laws that require firms to give workers months of notice before laying them off and to provide substantial severance or retraining packages after laying them off. The legally required notice before laying off a worker can be more than three months in Spain, Germany, Denmark, and Belgium, and the legally required severance package can be as high as a year’s salary or more in Austria, Spain, Portugal, Italy, and Greece. Such laws will surely discourage laying off or firing current workers. However, when companies know that it will be difficult to fire or lay off workers, they also become hesitant about hiring in the first place.

We can attribute the typically higher levels of unemployment in many European countries in recent years, which have prevailed even when economies are growing at a solid pace, to the fact that the sorts of laws and regulations that lead to a high natural rate of unemployment are much more prevalent in Europe than in the United States.

A Preview of Policies to Fight Unemployment

The **Government Budgets and Fiscal Policy** and **Macroeconomic Policy Around the World** chapters provide a detailed discussion of how to fight unemployment, when we can discuss these policies in the context of the full array of macroeconomic goals and frameworks for analysis. However, even at this preliminary stage, it is useful to preview the main issues concerning policies to fight unemployment.

The remedy for unemployment will depend on the diagnosis. Cyclical unemployment is a short-term problem, caused because the economy is in a recession. Thus, the preferred solution will be to avoid or minimize recessions. As **Government Budgets and Fiscal Policy** discusses, governments can enact this policy by stimulating the overall buying power in the economy, so that firms perceive that sales and profits are possible, which makes them eager to hire.

Dealing with the natural rate of unemployment is trickier. In a market-oriented economy, firms will hire and fire workers. Governments cannot control this. Furthermore, the evolving age structure of the economy's population, or unexpected shifts in productivity are beyond a government's control and, will affect the natural rate of unemployment for a time. However, as the example of high ongoing unemployment rates for many European countries illustrates, government policy clearly can affect the natural rate of unemployment that will persist even when GDP is growing.

When a government enacts policies that will affect workers or employers, it must examine how these policies will affect the information and incentives employees and employers have to find one another. For example, the government may have a role to play in helping some of the unemployed with job searches. Governments may need to rethink the design of their programs that offer assistance to unemployed workers and protections to employed workers so that they will not unduly discourage the supply of labor. Similarly, governments may need to reassess rules that make it difficult for businesses to begin or to expand so that they will not unduly discourage the demand for labor. The message is not that governments should repeal all laws affecting labor markets, but only that when they enact such laws, a society that cares about unemployment will need to consider the tradeoffs involved.

Bring it Home

Unemployment and the Great Recession

In the review of unemployment during and after the Great Recession at the outset of this chapter, we noted that unemployment tends to be a lagging indicator of business activity. This has historically been the case, and it is evident for all recessions that have taken place since the end of World War II. In brief, this results from the costs to employers of recruitment, hiring, and training workers. Those costs represent investments by firms in their work forces.

At the outset of a recession, when a firm realizes that demand for its product or service is not as strong as anticipated, it has an incentive to lay off workers. However, doing so runs the risk of losing those workers, and if the weak demand proves to be only temporary, the firm will be obliged to recruit, hire, and train new workers. Thus, firms tend to retain workers initially in a downturn. Similarly, as business begins to pick up when a recession is over, firms are not sure if the improvement will last. Rather than incur the costs of hiring and training new workers, they will wait, and perhaps resort to overtime work for existing workers, until they are confident that the recession is over.

Another point that we noted at the outset is that the duration of recoveries in employment following recessions has been longer following the last three recessions (going back to the early 1990s) than previously. Nir Jaimovich and Henry Siu have argued that these “jobless recoveries” are a consequence of job polarization – the disappearance of employment opportunities focused on “routine” tasks. Job polarization refers to the increasing concentration of employment in the highest- and lowest-wage occupations, as jobs in middle-skill occupations disappear. Job polarization is an outcome of technological progress in robotics, computing, and information and communication technology. The result of this progress is a decline in demand for labor in occupations that perform “routine” tasks – tasks that are limited in scope and can be performed by following a well-defined set of procedures – and hence a decline in the share of total employment that is composed of routine occupations. Jaimovich and Siu have shown that job polarization characterizes the aftermath of the last three recessions, and this appears to be responsible for the jobless recoveries.

KEY TERMS

adverse selection of wage cuts argument if employers reduce wages for all workers, the best will leave

cyclical unemployment unemployment closely tied to the business cycle, like higher unemployment during a recession

discouraged workers those who have stopped looking for employment due to the lack of suitable positions available

efficiency wage theory the theory that the productivity of workers, either individually or as a group, will increase if the employer pays them more

frictional unemployment unemployment that occurs as workers move between jobs

implicit contract an unwritten agreement in the labor market that the employer will try to keep wages from falling when the economy is weak or the business is having trouble, and the employee will not expect huge salary increases when the economy or the business is strong

insider-outsider model those already working for the firm are "insiders" who know the procedures; the other workers are "outsiders" who are recent or prospective hires

labor force participation rate this is the percentage of adults in an economy who are either employed or who are unemployed and looking for a job

natural rate of unemployment the unemployment rate that would exist in a growing and healthy economy from the combination of economic, social, and political factors that exist at a given time

out of the labor force those who are not working and not looking for work—whether they want employment or not; also termed "not in the labor force"

relative wage coordination argument across-the-board wage cuts are hard for an economy to implement, and workers fight against them

structural unemployment unemployment that occurs because individuals lack skills valued by employers

underemployed individuals who are employed in a job that is below their skills

unemployment rate the percentage of adults who are in the labor force and thus seeking jobs, but who do not have jobs

KEY CONCEPTS AND SUMMARY

8.1 How Economists Define and Compute Unemployment Rate

Unemployment imposes high costs. Unemployed individuals suffer from loss of income and from stress. An economy with high unemployment suffers an opportunity cost of unused resources. We can divide the adult population into those in the labor force and those out of the labor force. In turn, we divide those in the labor force into employed and unemployed. A person without a job must be willing and able to work and actively looking for work to be counted as unemployed; otherwise, a person without a job is counted as out of the labor force. Economists define the unemployment rate as the number of unemployed persons divided by the number of persons in the labor force (not the overall adult population). The Current Population Survey (CPS) conducted by the United States Census Bureau measures the percentage of the labor force that is unemployed. The establishment payroll survey by the Bureau of Labor Statistics measures the net change in jobs created for the month.

8.2 Patterns of Unemployment

The U.S. unemployment rate rises during periods of recession and depression, but falls back to the range of 4% to

6% when the economy is strong. The unemployment rate never falls to zero. Despite enormous growth in the size of the U.S. population and labor force in the twentieth century, along with other major trends like globalization and new technology, the unemployment rate shows no long-term rising trend.

Unemployment rates differ by group: higher for African-Americans and Hispanics than for whites; higher for less educated than more educated; higher for the young than the middle-aged. Women's unemployment rates used to be higher than men's, but in recent years men's and women's unemployment rates have been very similar. In recent years, unemployment rates in the United States have compared favorably with unemployment rates in most other high-income economies.

8.3 What Causes Changes in Unemployment over the Short Run

Cyclical unemployment rises and falls with the business cycle. In a labor market with flexible wages, wages will adjust in such a market so that quantity demanded of labor always equals the quantity supplied of labor at the equilibrium wage. Economists have proposed many theories for why wages might not be flexible, but instead may adjust only in a "sticky" way, especially when it comes to downward adjustments: implicit contracts, efficiency wage theory, adverse selection of wage cuts, insider-outsider model, and relative wage coordination.

8.4 What Causes Changes in Unemployment over the Long Run

The natural rate of unemployment is the rate of unemployment that the economic, social, and political forces in the economy would cause even when the economy is not in a recession. These factors include the frictional unemployment that occurs when people either choose to change jobs or are put out of work for a time by the shifts of a dynamic and changing economy. They also include any laws concerning conditions of hiring and firing that have the undesired side effect of discouraging job formation. They also include structural unemployment, which occurs when demand shifts permanently away from a certain type of job skill.

SELF-CHECK QUESTIONS

1. Suppose the adult population over the age of 16 is 237.8 million and the labor force is 153.9 million (of whom 139.1 million are employed). How many people are "not in the labor force?" What are the proportions of employed, unemployed and not in the labor force in the population? *Hint*: Proportions are percentages.

2. Using the above data, what is the unemployment rate? These data are U.S. statistics from 2010. How does it compare to the February 2015 unemployment rate computed earlier?

3. Over the long term, has the U.S. unemployment rate generally trended up, trended down, or remained at basically the same level?

4. Whose unemployment rates are commonly higher in the U.S. economy:

a. Whites or nonwhites?
b. The young or the middle-aged?
c. College graduates or high school graduates?

5. Beginning in the 1970s and continuing for three decades, women entered the U.S. labor force in a big way. If we assume that wages are sticky in a downward direction, but that around 1970 the demand for labor equaled the supply of labor at the current wage rate, what do you imagine happened to the wage rate, employment, and unemployment as a result of increased labor force participation?

6. Is the increase in labor force participation rates among women better thought of as causing an increase in cyclical unemployment or an increase in the natural rate of unemployment? Why?

7. Many college students graduate from college before they have found a job. When graduates begin to look for a job, they are counted as what category of unemployed?

REVIEW QUESTIONS

8. What is the difference between being unemployed and being out of the labor force?

9. How do you calculate the unemployment rate? How do you calculate the labor force participation rate?

10. Are all adults who do not hold jobs counted as unemployed?

11. If you are out of school but working part time, are you considered employed or unemployed in U.S. labor statistics? If you are a full time student and working 12 hours a week at the college cafeteria are you considered employed or not in the labor force? If you are a senior citizen who is collecting social security and a pension and working as a greeter at Wal-Mart are you considered employed or not in the labor force?

12. What happens to the unemployment rate when unemployed workers are reclassified as discouraged workers?

13. What happens to the labor force participation rate when employed individuals are reclassified as unemployed? What happens when they are reclassified as discouraged workers?

14. What are some of the problems with using the unemployment rate as an accurate measure of overall joblessness?

15. What criteria do the BLS use to count someone as employed? As unemployed?

16. Assess whether the following would be counted as "unemployed" in the Current Employment Statistics survey.

a. A husband willingly stays home with children while his wife works.
b. A manufacturing worker whose factory just closed down.
c. A college student doing an unpaid summer internship.
d. A retiree.
e. Someone who has been out of work for two years but keeps looking for a job.
f. Someone who has been out of work for two months but isn't looking for a job.
g. Someone who hates her present job and is actively looking for another one.
h. Someone who decides to take a part time job because she could not find a full time position.

17. Are U.S. unemployment rates typically higher, lower, or about the same as unemployment rates in other high-income countries?

18. Are U.S. unemployment rates distributed evenly across the population?

19. When would you expect cyclical unemployment to be rising? Falling?

20. Why is there unemployment in a labor market with flexible wages?

21. Name and explain some of the reasons why wages are likely to be sticky, especially in downward adjustments.

22. What term describes the remaining level of unemployment that occurs even when the economy is healthy?

23. What forces create the natural rate of unemployment for an economy?

24. Would you expect the natural rate of unemployment to be roughly the same in different countries?

25. Would you expect the natural rate of unemployment to remain the same within one country over the long run of several decades?

26. What is frictional unemployment? Give examples of frictional unemployment.

27. What is structural unemployment? Give examples of structural unemployment.

28. After several years of economic growth, would you expect the unemployment in an economy to be mainly cyclical or mainly due to the natural rate of unemployment? Why?

29. What type of unemployment (cyclical, frictional, or structural) applies to each of the following:

a. landscapers laid off in response to a drop in new housing construction during a recession.
b. coal miners laid off due to EPA regulations that shut down coal fired power
c. a financial analyst who quits his/her job in Chicago and is pursing similar work in Arizona
d. printers laid off due to drop in demand for printed catalogues and flyers as firms go the internet to promote an advertise their products.
e. factory workers in the U.S. laid off as the plants shut down and move to Mexico and Ireland.

CRITICAL THINKING QUESTIONS

30. Using the definition of the unemployment rate, is an increase in the unemployment rate necessarily a bad thing for a nation?

31. Is a decrease in the unemployment rate necessarily a good thing for a nation? Explain.

32. If many workers become discouraged from looking for jobs, explain how the number of jobs could decline but the unemployment rate could fall at the same time.

33. Would you expect hidden unemployment to be higher, lower, or about the same when the unemployment rate is high, say 10%, versus low, say 4%? Explain.

34. Is the higher unemployment rates for minority workers necessarily an indication of discrimination? What could be some other reasons for the higher unemployment rate?

35. While unemployment is highly negatively correlated with the level of economic activity, in the real world it responds with a lag. In other words, firms do not immediately lay off workers in response to a sales decline. They wait a while before responding. Similarly, firms do not immediately hire workers when sales pick up. What do you think accounts for the lag in response time?

36. Why do you think that unemployment rates are lower for individuals with more education?

37. Do you think it is rational for workers to prefer sticky wages to wage cuts, when the consequence of sticky wages is unemployment for some workers? Why or why not? How do the reasons for sticky wages explained in this section apply to your argument?

38. Under what condition would a decrease in unemployment be bad for the economy?

39. Under what condition would an increase in the unemployment rate be a positive sign?

40. As the baby boom generation retires, the ratio of retirees to workers will increase noticeably. How will this affect the Social Security program? How will this affect the standard of living of the average American?

41. Unemployment rates have been higher in many European countries in recent decades than in the United States. Is the main reason for this long-term difference in unemployment rates more likely to be cyclical unemployment or the natural rate of unemployment? Explain briefly.

42. Is it desirable to pursue a goal of zero unemployment? Why or why not?

43. Is it desirable to eliminate natural unemployment? Why or why not? *Hint*: Think about what our economy would look like today and what assumptions would have to be met to have a zero rate of natural unemployment.

44. The U.S. unemployment rate increased from 4.6% in July 2001 to 5.9% by June 2002. Without studying the subject in any detail, would you expect that a change of this kind is more likely to be due to cyclical unemployment or a change in the natural rate of unemployment? Why?

PROBLEMS

45. A country with a population of eight million adults has five million employed, 500,000 unemployed, and the rest of the adult population is out of the labor force. What's the unemployment rate? What share of population is in the labor force? Sketch a pie chart that divides the adult population into these three groups.

46. A government passes a family-friendly law that no companies can have evening, nighttime, or weekend hours, so that everyone can be home with their families during these times. Analyze the effect of this law using a demand and supply diagram for the labor market: first assuming that wages are flexible, and then assuming that wages are sticky downward.

47. As the baby boomer generation retires, what should happen to wages and employment? Can you show this graphically?

9 | Inflation

Figure 9.1 Big Bucks in Zimbabwe This bill was worth 100 billion Zimbabwean dollars when issued in 2008. There were even bills issued with a face value of 100 trillion Zimbabwean dollars. The bills had $100,000,000,000,000 written on them. Unfortunately, they were almost worthless. At one point, 621,984,228 Zimbabwean dollars were equal to one U.S. dollar. Eventually, the country abandoned its own currency and allowed people to use foreign currency for purchases. (Credit: modification of work by Samantha Marx/Flickr Creative Commons)

Bring it Home

A $550 Million Loaf of Bread?

If you were born within the last three decades in the United States, Canada, or many other countries in the developed world, you probably have no real experience with a high rate of inflation. Inflation is when most prices in an entire economy are rising. However, there is an extreme form of inflation called hyperinflation. This occurred in Germany between 1921 and 1928, and more recently in Zimbabwe between 2008 and 2009. In November 2008, Zimbabwe had an inflation rate of 79.6 billion percent. In contrast, in 2014, the United States had an average annual rate of inflation of 1.6%.

Zimbabwe's inflation rate was so high it is difficult to comprehend, so let's put it into context. It is equivalent to price increases of 98% per day. This means that, from one day to the next, prices essentially double. What is life like in an economy afflicted with hyperinflation? Most of you reading this will have never experienced this phenomenon. The government adjusted prices for commodities in Zimbabwean dollars several times *each day*. There was no desire to hold on to currency since it lost value by the minute. The people there spent a great deal of time getting rid of any cash they acquired by purchasing whatever food or other commodities they could find. At one point, a loaf of bread cost 550 million Zimbabwean dollars. Teachers' salaries were in the trillions a month; however, this was equivalent to only one U.S. dollar a day. At its height, it took 621,984,228 Zimbabwean dollars to purchase one U.S. dollar.

Government agencies had no money to pay their workers so they started printing money to pay their bills rather than raising taxes. Rising prices caused the government to enact price controls on private businesses, which led to shortages and the emergence of black markets. In 2009, the country abandoned its currency and allowed people to use foreign currencies for purchases.

How does this happen? How can both government and the economy fail to function at the most basic level? Before we consider these extreme cases of hyperinflation, let's first look at inflation itself.

Introduction to Inflation

In this chapter, you will learn about:

- Tracking Inflation
- How to Measure Changes in the Cost of Living
- How the U.S. and Other Countries Experience Inflation
- The Confusion Over Inflation
- Indexing and Its Limitations

Inflation is a general and ongoing rise in the level of prices in an entire economy. Inflation does not refer to a change in relative prices. A relative price change occurs when you see that the price of tuition has risen, but the price of laptops has fallen. Inflation, on the other hand, means that there is pressure for prices to rise in most markets in the economy. In addition, price increases in the supply-and-demand model were one-time events, representing a shift from a previous equilibrium to a new one. Inflation implies an ongoing rise in prices. If inflation happened for one year and then stopped, then it would not be inflation any more.

This chapter begins by showing how to combine prices of individual goods and services to create a measure of overall inflation. It discusses the historical and recent experience of inflation, both in the United States and in other countries around the world. Other chapters have sometimes included a note under an exhibit or a parenthetical reminder in the text saying that the numbers have been adjusted for inflation. In this chapter, it is time to show how to use inflation statistics to adjust other economic variables, so that you can tell how much of, for example, we can attribute the rise in GDP over different periods of time to an actual increase in the production of goods and services and how much we should attribute to the fact that prices for most items have risen.

Inflation has consequences for people and firms throughout the economy, in their roles as lenders and borrowers, wage-earners, taxpayers, and consumers. The chapter concludes with a discussion of some imperfections and biases in the inflation statistics, and a preview of policies for fighting inflation that we will discuss in other chapters.

9.1 | Tracking Inflation

By the end of this section, you will be able to:

- Calculate the annual rate of inflation
- Explain and use index numbers and base years when simplifying the total quantity spent over a year for products
- Calculate inflation rates using index numbers

Dinner table conversations where you might have heard about inflation usually entail reminiscing about when "everything seemed to cost so much less. You used to be able to buy three gallons of gasoline for a dollar and then go see an afternoon movie for another dollar." **Table 9.1** compares some prices of common goods in 1970 and 2017. Of course, the average prices in this table may not reflect the prices where you live. The cost of living in New York City is much higher than in Houston, Texas, for example. In addition, certain products have evolved over recent decades. A new car in 2017, loaded with antipollution equipment, safety gear, computerized engine controls, and many other technological advances, is a more advanced machine (and more fuel efficient) than your typical 1970s car. However, put details like these to one side for the moment, and look at the overall pattern. The primary reason behind the price rises in **Table 9.1**—and all the price increases for the other products in the economy—is not specific to the market for housing or cars or gasoline or movie tickets. Instead, it is part of a general rise in the level of all prices. At the beginning of 2017, $1 had about the same purchasing power in overall terms of goods and services as 18 cents did in

1972, because of the amount of inflation that has occurred over that time period.

Items	1970	2017
Pound of ground beef	$0.66	$3.62
Pound of butter	$0.87	$2.03
Movie ticket	$1.55	$8.65
Sales price of new home (median)	$22,000	$312,900
New car	$3,000	$4,077
Gallon of gasoline	$0.36	$2.35
Average hourly wage for a manufacturing worker	$3.23	$20.65
Per capita GDP	$5,069	$57,294

Table 9.1 Price Comparisons, 1970 and 2017 (Sources: See chapter References at end of book.)

Moreover, the power of inflation does not affect just goods and services, but wages and income levels, too. The second-to-last row of **Table 9.1** shows that the average hourly wage for a manufacturing worker increased nearly six-fold from 1970 to 2017. The average worker in 2017 is better educated and more productive than the average worker in 1970—but not six times more productive. Per capita GDP increased substantially from 1970 to 2017, but is the average person in the U.S. economy really more than eleven times better off in just 47 years? Not likely.

A modern economy has millions of goods and services whose prices are continually quivering in the breezes of supply and demand. How can all of these shifts in price attribute to a single inflation rate? As with many problems in economic measurement, the conceptual answer is reasonably straightforward: Economists combine prices of a variety of goods and services into a single price level. The inflation rate is simply the percentage change in the price level. Applying the concept, however, involves some practical difficulties.

The Price of a Basket of Goods

To calculate the price level, economists begin with the concept of a **basket of goods and services**, consisting of the different items individuals, businesses, or organizations typically buy. The next step is to look at how the prices of those items change over time. In thinking about how to combine individual prices into an overall price level, many people find that their first impulse is to calculate the average of the prices. Such a calculation, however, could easily be misleading because some products matter more than others.

Changes in the prices of goods for which people spend a larger share of their incomes will matter more than changes in the prices of goods for which people spend a smaller share of their incomes. For example, an increase of 10% in the rental rate on housing matters more to most people than whether the price of carrots rises by 10%. To construct an overall measure of the price level, economists compute a weighted average of the prices of the items in the basket, where the weights are based on the actual quantities of goods and services people buy. The following Work It Out feature walks you through the steps of calculating the annual rate of inflation based on a few products.

Work It Out

Calculating an Annual Rate of Inflation

Consider the simple basket of goods with only three items, represented in **Table 9.2**. Say that in any given month, a college student spends money on 20 hamburgers, one bottle of aspirin, and five movies. The table provides prices for these items over four years through each time period (Pd). Prices of some goods in the basket may rise while others fall. In this example, the price of aspirin does not change over the four years, while movies increase in price and hamburgers bounce up and down. The table shows the cost of buying the

given basket of goods at the prices prevailing at that time.

Items	Hamburger	Aspirin	Movies	Total	Inflation Rate
Qty	20	1 bottle	5	-	-
(Pd 1) Price	$3.00	$10.00	$6.00	-	-
(Pd 1) Amount Spent	$60.00	$10.00	$30.00	$100.00	-
(Pd 2) Price	$3.20	$10.00	$6.50	-	-
(Pd 2) Amount Spent	$64.00	$10.00	$32.50	$106.50	6.5%
(Pd 3) Price	$3.10	$10.00	$7.00	-	-
(Pd 3) Amount Spent	$62.00	$10.00	$35.00	$107.00	0.5%
(Pd 4) Price	$3.50	$10.00	$7.50	-	-
(Pd 4) Amount Spent	$70.00	$10.00	$37.50	$117.50	9.8%

Table 9.2 A College Student's Basket of Goods

To calculate the annual rate of inflation in this example:

Step 1. Find the percentage change in the cost of purchasing the overall basket of goods between the time periods. The general equation for percentage changes between two years, whether in the context of inflation or in any other calculation, is:

$$\frac{(\text{Level in new year} - \text{Level in previous year})}{\text{Level in previous year}} \times 100 = \text{Percentage change}$$

Step 2. From period 1 to period 2, the total cost of purchasing the basket of goods in Table 9.2 rises from $100 to $106.50. Therefore, the percentage change over this time—the inflation rate—is:

$$\frac{(106.50 - 100)}{100.0} = 0.065 = 6.5\%$$

Step 3. From period 2 to period 3, the overall change in the cost of purchasing the basket rises from $106.50 to $107. Thus, the inflation rate over this time, again calculated by the percentage change, is approximately:

$$\frac{(107 - 106.50)}{106.50} = 0.0047 = 0.47\%$$

Step 4. From period 3 to period 4, the overall cost rises from $107 to $117.50. The inflation rate is thus:

$$\frac{(117.50 - 107)}{107} = 0.098 = 9.8\%$$

This calculation of the change in the total cost of purchasing a basket of goods accounts for how much a student spends on each good. Hamburgers are the lowest-priced good in this example, and aspirin is the highest-priced. If an individual buys a greater quantity of a low-price good, then it makes sense that changes in the price of that good should have a larger impact on the buying power of that person's money. The larger impact of hamburgers shows up in the "amount spent" row, where, in all time periods, hamburgers are the largest item within the amount spent row.

Index Numbers

The numerical results of a calculation based on a basket of goods can get a little messy. The simplified example in Table 9.2 has only three goods and the prices are in even dollars, not numbers like 79 cents or $124.99. If the list of products were much longer, and we used more realistic prices, the total quantity spent over a year might be some

messy-looking number like $17,147.51 or $27,654.92.

To simplify the task of interpreting the price levels for more realistic and complex baskets of goods, economists typically report the price level in each period as an **index number**, rather than as the dollar amount for buying the basket of goods. Economists create price indices to calculate an overall average change in relative prices over time. To convert the money spent on the basket to an index number, economists arbitrarily choose one year to be the **base year**, or starting point from which we measure changes in prices. The base year, by definition, has an index number equal to 100. This sounds complicated, but it is really a simple math trick. In the example above, say that we choose time period 3 as the base year. Since the total amount of spending in that year is $107, we divide that amount by itself ($107) and multiply by 100. Again, this is because the index number in the base year *always* has to have a value of 100. Then, to figure out the values of the index number for the other years, we divide the dollar amounts for the other years by 1.07 as well. Note also that the dollar signs cancel out so that index numbers have no units.

Table 9.3 shows calculations for the other values of the index number, based on the example in **Table 9.2**. Because we calculate the index numbers so that they are in exactly the same proportion as the total dollar cost of purchasing the basket of goods, we can calculate the inflation rate based on the index numbers, using the percentage change formula. Thus, the inflation rate from period 1 to period 2 would be

$$\frac{(99.5 - 93.4)}{93.4} = 0.065 = 6.5\%$$

This is the same answer that we derived when measuring inflation based on the dollar cost of the basket of goods for the same time period.

	Total Spending	Index Number	Inflation Rate Since Previous Period
Period 1	$100	$\frac{100}{1.07} = 93.4$	
Period 2	$106.50	$\frac{106.50}{1.07} = 99.5$	$\frac{(99.5 - 93.4)}{93.4} = 0.065 = 6.5\%$
Period 3	$107	$\frac{107}{1.07} = 100.0$	$\frac{100 - 99.5}{99.5} = 0.005 = 0.5\%$
Period 4	$117.50	$\frac{117.50}{1.07} = 109.8$	$\frac{109.8 - 100}{100} = 0.098 = 9.8\%$

Table 9.3 Calculating Index Numbers When Period 3 is the Base Year

If the inflation rate is the same whether it is based on dollar values or index numbers, then why bother with the index numbers? The advantage is that indexing allows easier eyeballing of the inflation numbers. If you glance at two index numbers like 107 and 110, you know automatically that the rate of inflation between the two years is about, but not quite exactly equal to, 3%. By contrast, imagine that we express the price levels in absolute dollars of a large basket of goods, so that when you looked at the data, the numbers were $19,493.62 and $20,040.17. Most people find it difficult to eyeball those kinds of numbers and say that it is a change of about 3%. However, the two numbers expressed in absolute dollars are exactly in the same proportion of 107 to 110 as the previous example. If you're wondering why simple subtraction of the index numbers wouldn't work, read the following Clear It Up feature.

Clear It Up

Why do you not just subtract index numbers?

A word of warning: When a price index moves from, say, 107 to 110, the rate of inflation is not *exactly* 3%. Remember, the inflation rate is not derived by subtracting the index numbers, but rather through the percentage-change calculation. We calculate the precise inflation rate as the price index moves from 107 to

110 as 100 x (110 – 107) / 107 = 100 x 0.028 = 2.8%. When the base year is fairly close to 100, a quick subtraction is not a terrible shortcut to calculating the inflation rate—but when precision matters down to tenths of a percent, subtracting will not give the right answer.

Two final points about index numbers are worth remembering. First, index numbers have no dollar signs or other units attached to them. Although we can use index numbers to calculate a percentage inflation rate, the index numbers themselves do not have percentage signs. Index numbers just mirror the proportions that we find in other data. They transform the other data so that it is easier to work with the data.

Second, the choice of a base year for the index number—that is, the year that is automatically set equal to 100—is arbitrary. We choose it as a starting point from which we can track changes in prices. In the official inflation statistics, it is common to use one base year for a few years, and then to update it, so that the base year of 100 is relatively close to the present. However, any base year that we choose for the index numbers will result in exactly the same inflation rate. To see this in the previous example, imagine that period 1 is the base year when total spending was $100, and we assign it an index number of 100. At a glance, you can see that the index numbers would now exactly match the dollar figures, and the inflation rate in the first period would be 6.5%.

Now that we see how indexes work to track inflation, the next module will show us how economists measure the cost of living.

Link It Up

Watch this video (http://openstaxcollege.org/l/Duck_Tales) from the cartoon *Duck Tales* to view a mini-lesson on inflation.

9.2 | How to Measure Changes in the Cost of Living

By the end of this section, you will be able to:

- Use the Consumer Price Index (CPI) to calculate U.S. inflation rates
- Identify several ways the Bureau of Labor Statistics avoids biases in the Consumer Price Index (CPI)
- Differentiate among the Consumer Price Index (CPI), the Producer Price Index (PPI), the International Price Index, the Employment Cost Index, and the GDP deflator.

The most commonly cited measure of inflation in the United States is the **Consumer Price Index (CPI)**. Government statisticians at the U.S. Bureau of Labor Statistics calculate the CPI based on the prices in a fixed basket of goods and services that represents the purchases of the average family of four. In recent years, the statisticians have paid considerable attention to a subtle problem: that the change in the total cost of buying a fixed basket of goods and services over time is conceptually not quite the same as the change in the cost of living, because the cost of living represents how much it costs for a person to feel that his or her consumption provides an equal level of satisfaction or utility.

To understand the distinction, imagine that over the past 10 years, the cost of purchasing a fixed basket of goods increased by 25% and your salary also increased by 25%. Has your personal standard of living held constant? If you

do not necessarily purchase an identical fixed basket of goods every year, then an inflation calculation based on the cost of a fixed basket of goods may be a misleading measure of how your cost of living has changed. Two problems arise here: substitution bias and quality/new goods bias.

When the price of a good rises, consumers tend to purchase less of it and to seek out substitutes instead. Conversely, as the price of a good falls, people will tend to purchase more of it. This pattern implies that goods with generally rising prices should tend over time to become less important in the overall basket of goods used to calculate inflation, while goods with falling prices should tend to become more important. Consider, as an example, a rise in the price of peaches by $100 per pound. If consumers were utterly inflexible in their demand for peaches, this would lead to a big rise in the price of food for consumers. Alternatively, imagine that people are utterly indifferent to whether they have peaches or other types of fruit. Now, if peach prices rise, people completely switch to other fruit choices and the average price of food does not change at all. A fixed and unchanging basket of goods assumes that consumers are locked into buying exactly the same goods, regardless of price changes—not a very likely assumption. Thus, **substitution bias**—the rise in the price of a fixed basket of goods over time—tends to overstate the rise in a consumer's true cost of living, because it does not take into account that the person can substitute away from goods whose relative prices have risen.

The other major problem in using a fixed basket of goods as the basis for calculating inflation is how to deal with the arrival of improved versions of older goods or altogether new goods. Consider the problem that arises if a cereal is improved by adding 12 essential vitamins and minerals—and also if a box of the cereal costs 5% more. It would clearly be misleading to count the entire resulting higher price as inflation, because the new price reflects a higher quality (or at least different) product. Ideally, one would like to know how much of the higher price is due to the quality change, and how much of it is just a higher price. The Bureau of Labor Statistics, which is responsible for computing the Consumer Price Index, must deal with these difficulties in adjusting for quality changes.

Link It Up

Visit this website (http://openstax.org/l/Fords) to view a list of Ford car prices between 1909 and 1927. Consider how these prices compare to today's models. Is the product today of a different quality?

We can think of a new product as an extreme improvement in quality—from something that did not exist to something that does. However, the basket of goods that was fixed in the past obviously does not include new goods created since then. The basket of goods and services in the Consumer Price Index (CPI) is revised and updated over time, and so new products are gradually included. However, the process takes some time. For example, room air conditioners were widely sold in the early 1950s, but were not introduced into the basket of goods behind the Consumer Price Index until 1964. The VCR and personal computer were available in the late 1970s and widely sold by the early 1980s, but did not enter the CPI basket of goods until 1987. By 1996, there were more than 40 million cellular phone subscribers in the United States—but cell phones were not yet part of the CPI basket of goods. The parade of inventions has continued, with the CPI inevitably lagging a few years behind.

The arrival of new goods creates problems with respect to the accuracy of measuring inflation. The reason people buy new goods, presumably, is that the new goods offer better value for money than existing goods. Thus, if the price index leaves out new goods, it overlooks one of the ways in which the cost of living is improving. In addition, the price of a new good is often higher when it is first introduced and then declines over time. If the new good is not included in the CPI for some years, until its price is already lower, the CPI may miss counting this price decline altogether. Taking these arguments together, the **quality/new goods bias** means that the rise in the price of a fixed basket of goods over time tends to overstate the rise in a consumer's true cost of living, because it does not account

for how improvements in the quality of existing goods or the invention of new goods improves the standard of living. The following Clear It Up feature is a must-read on how statisticians comprise and calculate the CPI.

Clear It Up

How do U.S. government statisticians measure the Consumer Price Index?

When the U.S. Bureau of Labor Statistics (BLS) calculates the Consumer Price Index, the first task is to decide on a basket of goods that is representative of the purchases of the average household. We do this by using the Consumer Expenditure Survey, a national survey of about 7,000 households, which provides detailed information on spending habits. Statisticians divide consumer expenditures into eight major groups (seen below), which in turn they divide into more than 200 individual item categories. The BLS currently uses 1982–1984 as the base period.

For each of the 200 individual expenditure items, the BLS chooses several hundred very specific examples of that item and looks at the prices of those examples. In figuring out the "breakfast cereal" item under the overall category of "foods and beverages," the BLS picks several hundred examples of breakfast cereal. One example might be the price of a 24-oz. box of a particular brand of cereal sold at a particular store. The BLS statistically selects specific products and sizes and stores to reflect what people buy and where they shop. The basket of goods in the Consumer Price Index thus consists of about 80,000 products; that is, several hundred specific products in over 200 broad-item categories. Statisticians rotate about one-quarter of these 80,000 specific products of the sample each year, and replace them with a different set of products.

The next step is to collect data on prices. Data collectors visit or call about 23,000 stores in 87 urban areas all over the United States every month to collect prices on these 80,000 specific products. The BLS also conducts a survey of 50,000 landlords or tenants to collect information about rents.

Statisticians then calculate the Consumer Price Index by taking the 80,000 prices of individual products and combining them, using weights (see **Figure 9.2**) determined by the quantities of these products that people buy and allowing for factors like substitution between goods and quality improvements, into price indices for the 200 or so overall items. Then, the statisticians combine the price indices for the 200 items into an overall Consumer Price Index. According the Consumer Price Index website, there are eight categories that data collectors use:

The Eight Major Categories in the Consumer Price Index

1. Food and beverages (breakfast cereal, milk, coffee, chicken, wine, full-service meals, and snacks)
2. Housing (renter's cost of housing, homeowner's cost of housing, fuel oil, bedroom furniture)
3. Apparel (men's shirts and sweaters, women's dresses, jewelry)
4. Transportation (new vehicles, airline fares, gasoline, motor vehicle insurance)
5. Medical care (prescription drugs and medical supplies, physicians' services, eyeglasses and eye care, hospital services)
6. Recreation (televisions, cable television, pets and pet products, sports equipment, admissions)
7. Education and communication (college tuition, postage, telephone services, computer software and accessories)
8. Other goods and services (tobacco and smoking products, haircuts and other personal services, funeral expenses)

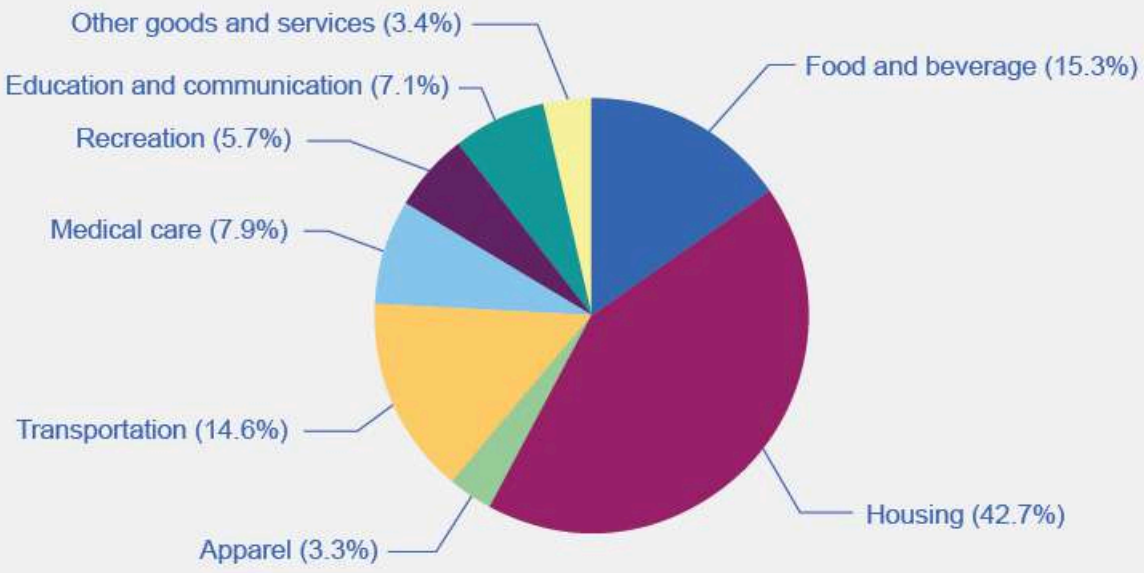


Figure 9.2 The Weighting of CPI Components Of the eight categories used to generate the Consumer Price Index, housing is the highest at 42.7%. The next highest category, food and beverage at 15.3%, is less than half the size of housing. Other goods and services, and apparel, are the lowest at 3.4% and 3.3%, respectively. (Source: www.bls.gov/cpi)

The CPI and Core Inflation Index

Imagine if you were driving a company truck across the country- you probably would care about things like the prices of available roadside food and motel rooms as well as the truck's operating condition. However, the manager of the firm might have different priorities. He would care mostly about the truck's on-time performance and much less so about the food you were eating and the places you were staying. In other words, the company manager would be paying attention to the firm's production, while ignoring transitory elements that impacted you, but did not affect the company's bottom line.

In a sense, a similar situation occurs with regard to measures of inflation. As we've learned, CPI measures prices as they affect everyday household spending. Economists typically calculate a **core inflation index** by taking the CPI and excluding volatile economic variables. In this way, economists have a better sense of the underlying trends in prices that affect the cost of living.

Examples of excluded variables include energy and food prices, which can jump around from month to month because of the weather. According to an article by Kent Bernhard, during Hurricane Katrina in 2005, a key supply point for the nation's gasoline was nearly knocked out. Gas prices quickly shot up across the nation, in some places by up to 40 cents a gallon in one day. This was not the cause of an economic policy but rather a short-lived event until the pumps were restored in the region. In this case, the CPI that month would register the change as a cost of living event to households, but the core inflation index would remain unchanged. As a result, the Federal Reserve's decisions on interest rates would not be influenced. Similarly, droughts can cause world-wide spikes in food prices that, if temporary, do not affect the nation's economic capability.

As former Chairman of the Federal Reserve Ben Bernanke noted in 1999 about the core inflation index, "It provide(s) a better guide to monetary policy than the other indices, since it measures the more persistent underlying inflation rather than transitory influences on the price level." Bernanke also noted that it helps communicate that the Federal Reserve does not need to respond to every inflationary shock since some price changes are transitory and not part of a structural change in the economy.

In sum, both the CPI and the core inflation index are important, but serve different audiences. The CPI helps households understand their overall cost of living from month to month, while the core inflation index is a preferred gauge from which to make important government policy changes.

Practical Solutions for the Substitution and the Quality/New Goods Biases

By the early 2000s, the Bureau of Labor Statistics was using alternative mathematical methods for calculating the Consumer Price Index, more complicated than just adding up the cost of a fixed basket of goods, to allow for some substitution between goods. It was also updating the basket of goods behind the CPI more frequently, so that it could include new and improved goods more rapidly. For certain products, the BLS was carrying out studies to try to measure the quality improvement. For example, with computers, an economic study can try to adjust for changes in speed, memory, screen size, and other product characteristics, and then calculate the change in price after accounting for these product changes. However, these adjustments are inevitably imperfect, and exactly how to make these adjustments is often a source of controversy among professional economists.

By the early 2000s, the substitution bias and quality/new goods bias had been somewhat reduced, so that since then the rise in the CPI probably overstates the true rise in inflation by only about 0.5% per year. Over one or a few years, this is not much. Over a period of a decade or two, even half of a percent per year compounds to a more significant amount. In addition, the CPI tracks prices from physical locations, and not at online sites like Amazon, where prices can be lower.

When measuring inflation (and other economic statistics, too), a tradeoff arises between simplicity and interpretation. If we calculate the inflation rate with a basket of goods that is fixed and unchanging, then the calculation of an inflation rate is straightforward, but the problems of substitution bias and quality/new goods bias will arise. However, when the basket of goods is allowed to shift and evolve to reflect substitution toward lower relative prices, quality improvements, and new goods, the technical details of calculating the inflation rate grow more complex.

Additional Price Indices: PPI, GDP Deflator, and More

The basket of goods behind the Consumer Price Index represents an average hypothetical U.S. household's consumption, which is to say that it does not exactly capture anyone’s personal experience. When the task is to calculate an average level of inflation, this approach works fine. What if, however, you are concerned about inflation experienced by a certain group, like the elderly, or the poor, or single-parent families with children, or Hispanic-Americans? In specific situations, a price index based on the buying power of the average consumer may not feel quite right.

This problem has a straightforward solution. If the Consumer Price Index does not serve the desired purpose, then invent another index, based on a basket of goods appropriate for the group of interest. The Bureau of Labor Statistics publishes a number of experimental price indices: some for particular groups like the elderly or the poor, some for different geographic areas, and some for certain broad categories of goods like food or housing.

The BLS also calculates several price indices that are not based on baskets of consumer goods. For example, the **Producer Price Index (PPI)** is based on prices paid for supplies and inputs by producers of goods and services. We can break it down into price indices for different industries, commodities, and stages of processing (like finished goods, intermediate goods, or crude materials for further processing). There is an **International Price Index** based on the prices of merchandise that is exported or imported. An **Employment Cost Index** measures wage inflation in the labor market. The **GDP deflator**, which the Bureau of Economic Analysis measures, is a price index that includes all the GDP components (that is, consumption plus investment plus government plus exports minus imports). Unlike the CPI, its baskets are not fixed but re-calculate what that year’s GDP would have been worth using the base-year’s prices. MIT's Billion Prices Project is a more recent alternative attempt to measure prices: economists collect data online from retailers and then put them into an index that they compare to the CPI (Source: http://bpp.mit.edu/usa/).

What’s the best measure of inflation? If one is concerned with the most accurate measure of inflation, one should use the GDP deflator as it picks up the prices of goods and services produced. However, it is not a good measure of the cost of living as it includes prices of many products not purchased by households (for example, aircraft, fire engines, factory buildings, office complexes, and bulldozers). If one wants the most accurate measure of inflation as it impacts households, one should use the CPI, as it only picks up prices of products purchased by households. That is why economists sometimes refer to the CPI as the cost-of-living index. As the Bureau of Labor Statistics states on its website: “The ‘best’ measure of inflation for a given application depends on the intended use of the data.”

9.3 | How the U.S. and Other Countries Experience Inflation

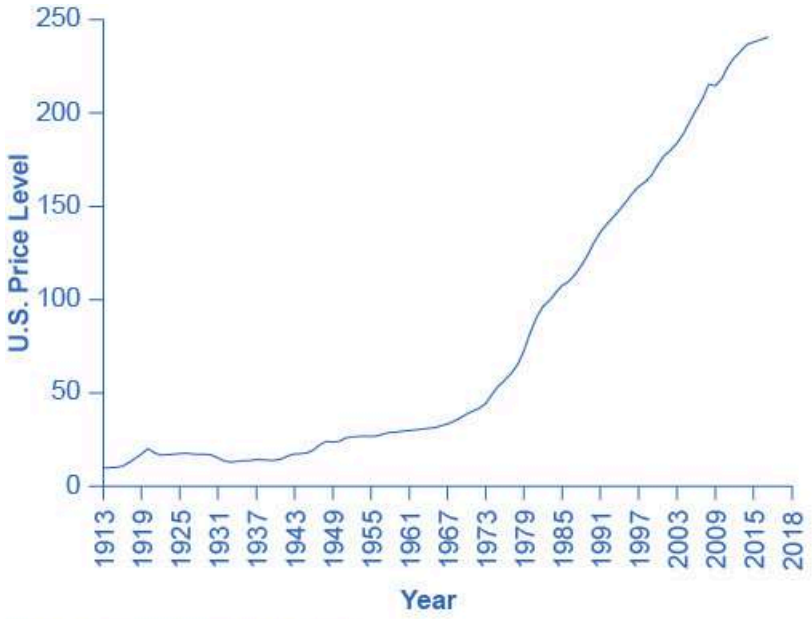
By the end of this section, you will be able to:

- Identify patterns of inflation for the United States using data from the Consumer Price Index
- Identify patterns of inflation on an international level

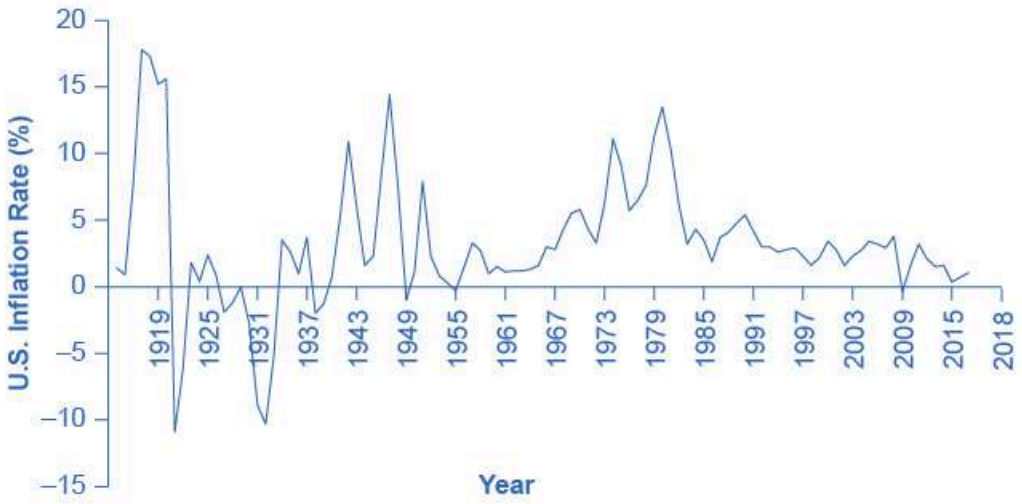
In the last three decades, inflation has been relatively low in the U.S. economy, with the Consumer Price Index typically rising 2% to 4% per year. Looking back over the twentieth century, there have been several periods where inflation caused the price level to rise at double-digit rates, but nothing has come close to hyperinflation.

Historical Inflation in the U.S. Economy

Figure 9.3 (a) shows the level of prices in the Consumer Price Index stretching back to 1913. In this case, the base years (when the CPI is defined as 100) are set for the average level of prices that existed from 1982 to 1984. **Figure 9.3** (b) shows the annual percentage changes in the CPI over time, which is the inflation rate.



(a) U.S. price level 1913-2016



(b) U.S. inflation rate 1913-2016

Figure 9.3 U.S. Price Level and Inflation Rates since 1913 Graph a shows the trends in the U.S. price level from the year 1913 to 2016. In 1913, the graph starts out close to 10, rises to around 20 in 1920, stays around 16 or 17 until 1931, then falls to 13 or 14 until 1940. It gradually increases until about 1973, then increases more rapidly through the remainder of the 1970s and beyond, with periodic dips, until 2016, when it reached around 240. Graph b shows the trends in U.S. inflation rates from the year 1914 to 2016. In 1916, the graph starts out with inflation at almost 8%, jumps to about 17% in 1917, drops drastically to close to –11% in 1921, goes up and down periodically, with peaks in the 1940s and the 1970s, until settling to around 1.3% in 2016.

The first two waves of inflation are easy to characterize in historical terms: they are right after World War I and World War II. However, there are also two periods of severe negative inflation—called **deflation**—in the early decades of the twentieth century: one following the deep 1920-21 recession of and the other during the 1930s Great Depression of the 1930s. (Since inflation is a time when the buying power of money in terms of goods and services is reduced, deflation will be a time when the buying power of money in terms of goods and services increases.) For the period from 1900 to about 1960, the major inflations and deflations nearly balanced each other out, so the average annual rate of inflation over these years was only about 1% per year. A third wave of more severe inflation arrived in the

1970s and departed in the early 1980s.

Link It Up

Visit this website (http://openstax.org/l/CPI_calculator) to use an inflation calculator and discover how prices have changed in the last 100 years.

Times of recession or depression often seem to be times when the inflation rate is lower, as in the recession of 1920–1921, the Great Depression, the recession of 1980–1982, and the Great Recession in 2008–2009. There were a few months in 2009 that were deflationary, but not at an annual rate. High levels of unemployment typically accompany recessions, and the total demand for goods falls, pulling the price level down. Conversely, the rate of inflation often, but not always, seems to start moving up when the economy is growing very strongly, like right after wartime or during the 1960s. The frameworks for macroeconomic analysis, that we developed in other chapters, will explain why recession often accompanies higher unemployment and lower inflation, while rapid economic growth often brings lower unemployment but higher inflation.

Inflation around the World

Around the rest of the world, the pattern of inflation has been very mixed; **Figure 9.4** shows inflation rates over the last several decades. Many industrialized countries, not just the United States, had relatively high inflation rates in the 1970s. For example, in 1975, Japan's inflation rate was over 8% and the inflation rate for the United Kingdom was almost 25%. In the 1980s, inflation rates came down in the United States and in Europe and have largely stayed down.

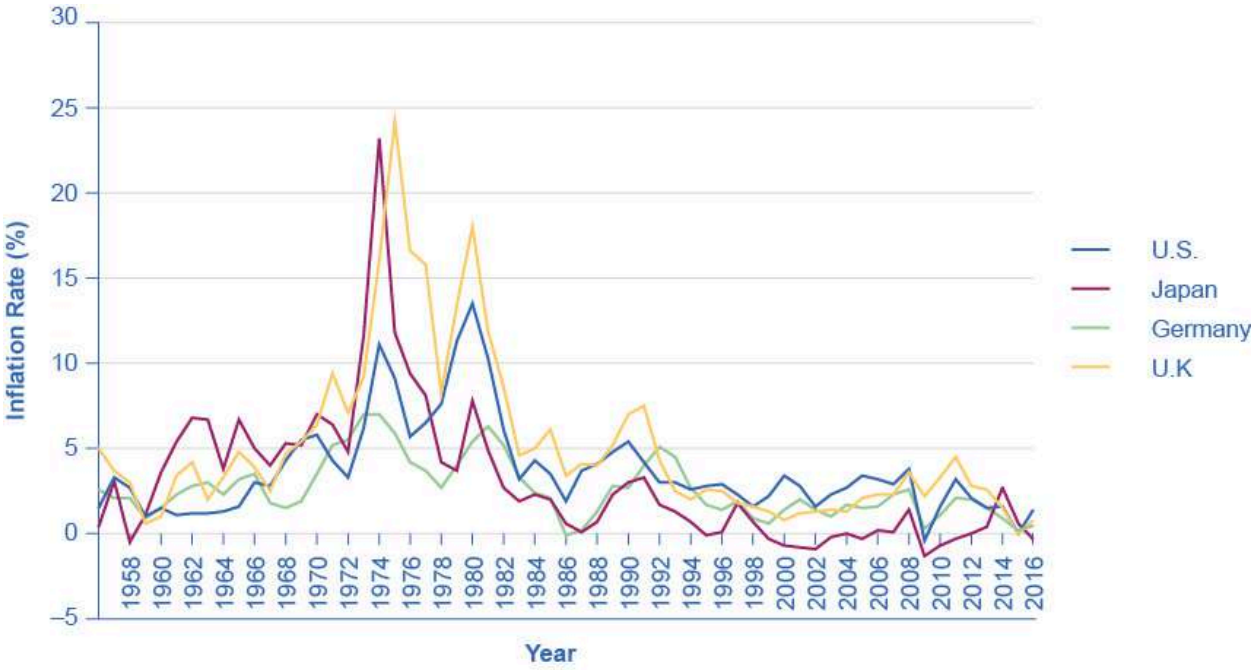
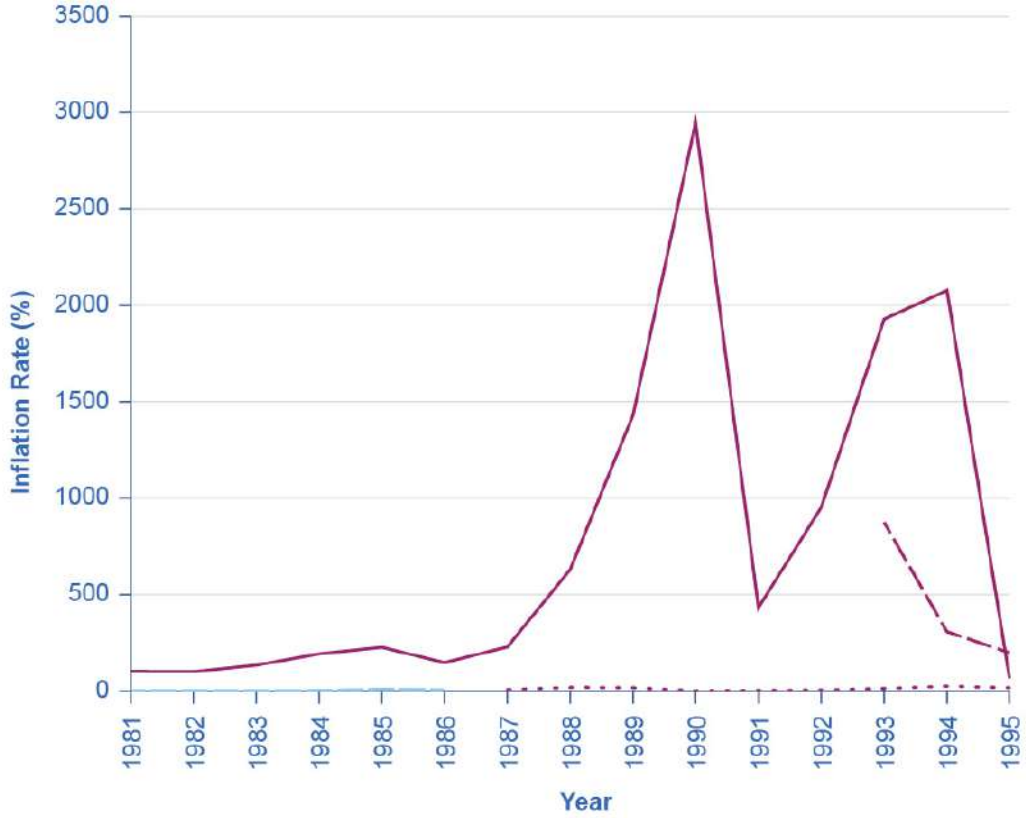
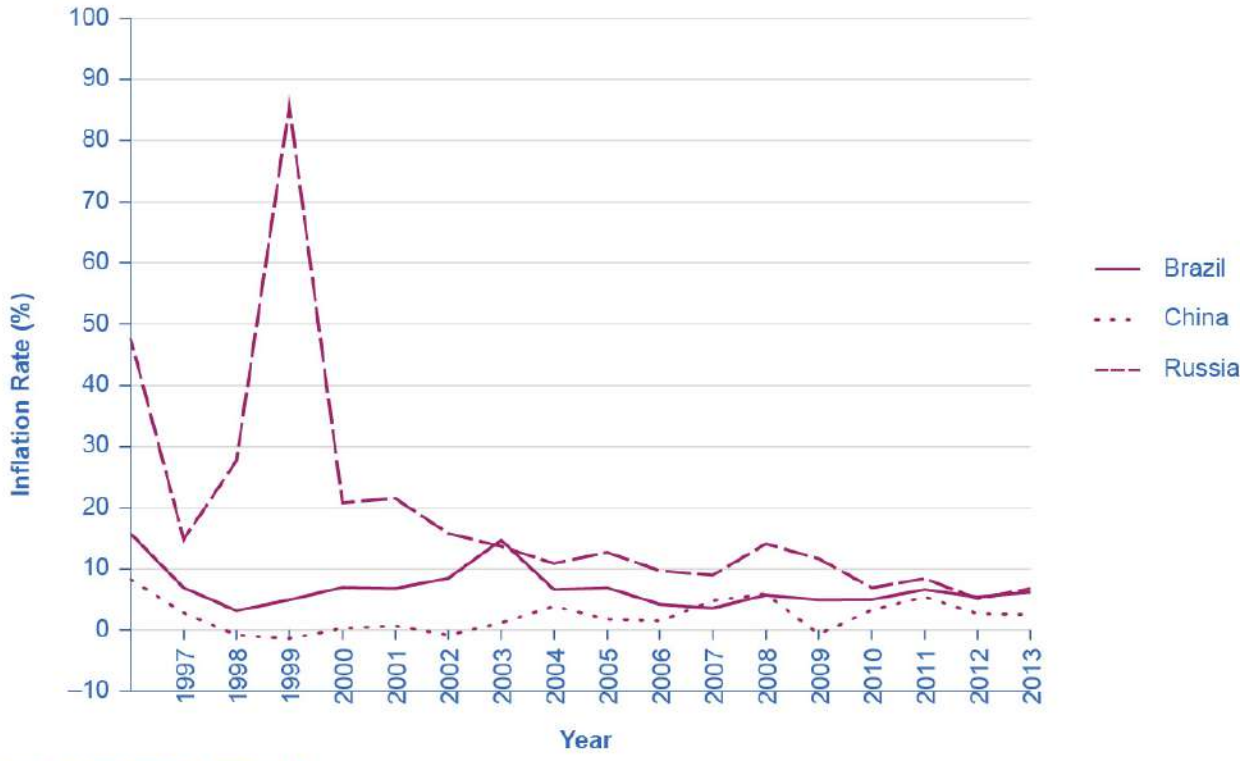


Figure 9.4 Countries with Relatively Low Inflation Rates, 1960–2016 This chart shows the annual percentage change in consumer prices compared with the previous year's consumer prices in the United States, the United Kingdom, Japan, and Germany.

Countries with controlled economies in the 1970s, like the Soviet Union and China, historically had very low rates of measured inflation—because prices were forbidden to rise by law, except for the cases where the government deemed a price increase to be due to quality improvements. However, these countries also had perpetual shortages of goods, since forbidding prices to rise acts like a price ceiling and creates a situation where quantity demanded often exceeds quantity supplied. As Russia and China made a transition toward more market-oriented economies, they also experienced outbursts of inflation, although we should regard the statistics for these economies as somewhat shakier. Inflation in China averaged about 10% per year for much of the 1980s and early 1990s, although it has dropped off since then. Russia experienced **hyperinflation**—an outburst of high inflation—of 2,500% per year in the early 1990s, although by 2006 Russia's consumer price inflation had dipped below 10% per year, as **Figure 9.5** shows. The closest the United States has ever reached hyperinflation was during the 1860–1865 Civil War, in the Confederate states.



(a) Inflation rates 1980-1995



(b) Inflation rates 1996-2013

Figure 9.5 Countries with Relatively High Inflation Rates, 1980–2016 These charts show the percentage change in consumer prices compared with the previous year's consumer prices in Brazil, China, and Russia. (a) Of these, Brazil and Russia experienced very high inflation at some point between the late-1980s and late-1990s. (b) Though not as high, China also had high inflation rates in the mid-1990s. Even though their inflation rates have come down over the last two decades, several of these countries continue to see significant inflation rates. (Sources: http://www.inflation.eu/inflation-rates; http://research.stlouisfed.org/fred2/series/FPCPITOTLZGBRA; http://research.stlouisfed.org/fred2/series/CHNCPIALLMINMEI; http://research.stlouisfed.org/fred2/series/FPCPITOTLZGRUS)

Many countries in Latin America experienced raging inflation during the 1980s and early 1990s, with inflation rates often well above 100% per year. In 1990, for example, both Brazil and Argentina saw inflation climb above 2000%. Certain countries in Africa experienced extremely high rates of inflation, sometimes bordering on hyperinflation, in the 1990s. Nigeria, the most populous country in Africa, had an inflation rate of 75% in 1995.

In the early 2000s, the problem of inflation appears to have diminished for most countries, at least in comparison to the worst times of recent decades. As we noted in this earlier Bring it Home feature, in recent years, the world's worst example of hyperinflation was in Zimbabwe, where at one point the government was issuing bills with a face value of $100 trillion (in Zimbabwean dollars)—that is, the bills had $100,000,000,000,000 written on the front, but were almost worthless. In many countries, the memory of double-digit, triple-digit, and even quadruple-digit inflation is not very far in the past.

9.4 | The Confusion Over Inflation

By the end of this section, you will be able to:

- Explain how inflation can cause redistributions of purchasing power
- Identify ways inflation can blur the perception of supply and demand
- Explain the economic benefits and challenges of inflation

Economists usually oppose high inflation, but they oppose it in a milder way than many non-economists. Robert Shiller, one of 2013's Nobel Prize winners in economics, carried out several surveys during the 1990s about attitudes toward inflation. One of his questions asked, "Do you agree that preventing high inflation is an important national priority, as important as preventing drug abuse or preventing deterioration in the quality of our schools?" Answers were on a scale of 1–5, where 1 meant "Fully agree" and 5 meant "Completely disagree." For the U.S. population as a whole, 52% answered "Fully agree" that preventing high inflation was a highly important national priority and just 4% said "Completely disagree." However, among professional economists, only 18% answered "Fully agree," while the same percentage of 18% answered "Completely disagree."

The Land of Funny Money

What are the economic problems caused by inflation, and why do economists often regard them with less concern than the general public? Consider a very short story: "The Land of Funny Money."

One morning, everyone in the Land of Funny Money awakened to find that everything denominated in money had increased by 20%. The change was completely unexpected. Every price in every store was 20% higher. Paychecks were 20% higher. Interest rates were 20 % higher. The amount of money, everywhere from wallets to savings accounts, was 20% larger. This overnight inflation of prices made newspaper headlines everywhere in the Land of Funny Money. However, the headlines quickly disappeared, as people realized that in terms of what they could actually buy with their incomes, this inflation had no economic impact. Everyone's pay could still buy exactly the same set of goods as it did before. Everyone's savings were still sufficient to buy exactly the same car, vacation, or retirement that they could have bought before. Equal levels of inflation in all wages and prices ended up not mattering much at all.

When the people in Robert Shiller's surveys explained their concern about inflation, one typical reason was that they feared that as prices rose, they would not be able to afford to buy as much. In other words, people were worried because they did not live in a place like the Land of Funny Money, where all prices and wages rose simultaneously. Instead, people live here on Planet Earth, where prices might rise while wages do not rise at all, or where wages rise more slowly than prices.

Economists note that over most periods, the inflation level in prices is roughly similar to the inflation level in wages, and so they reason that, on average, over time, people's economic status is not greatly changed by inflation. If all prices, wages, and interest rates adjusted automatically and immediately with inflation, as in the Land of Funny Money, then no one's purchasing power, profits, or real loan payments would change. However, if other economic variables do not move exactly in sync with inflation, or if they adjust for inflation only after a time lag, then inflation can cause three types of problems: unintended redistributions of purchasing power, blurred price signals, and difficulties in long-term planning.

Unintended Redistributions of Purchasing Power

Inflation can cause redistributions of purchasing power that hurt some and help others. People who are hurt by inflation include those who are holding considerable cash, whether it is in a safe deposit box or in a cardboard box under the bed. When inflation happens, the buying power of cash diminishes. However, cash is only an example of a more general problem: anyone who has financial assets invested in a way that the nominal return does not keep up with inflation will tend to suffer from inflation. For example, if a person has money in a bank account that pays 4% interest, but inflation rises to 5%, then the real rate of return for the money invested in that bank account is negative 1%.

The problem of a good-looking nominal interest rate transforming into an ugly-looking real interest rate can be worsened by taxes. The U.S. income tax is charged on the nominal interest received in dollar terms, without an adjustment for inflation. Thus, the government taxes a person who invests $10,000 and receives a 5% nominal rate of interest on the $500 received—no matter whether the inflation rate is 0%, 5%, or 10%. If inflation is 0%, then the real interest rate is 5% and all $500 is a gain in buying power. However, if inflation is 5%, then the real interest rate is zero and the person had no real gain—but owes income tax on the nominal gain anyway. If inflation is 10%, then the real interest rate is *negative* 5% and the person is actually falling behind in buying power, but would still owe taxes on the $500 in nominal gains.

Inflation can cause unintended redistributions for wage earners, too. Wages do typically creep up with inflation over time, eventually. The last row of **Table 9.1** at the start of this chapter showed that the average hourly wage in manufacturing in the U.S. economy increased from $3.23 in 1970 to $20.65 in 2017, which is an increase by a factor of more than six. Over that time period, the Consumer Price Index increased by an almost identical amount. However, increases in wages may lag behind inflation for a year or two, since wage adjustments are often somewhat sticky and occur only once or twice a year. Moreover, the extent to which wages keep up with inflation creates insecurity for workers and may involve painful, prolonged conflicts between employers and employees. If the government adjusts minimum wage for inflation only infrequently, minimum wage workers are losing purchasing power from their nominal wages, as **Figure 9.6** shows.

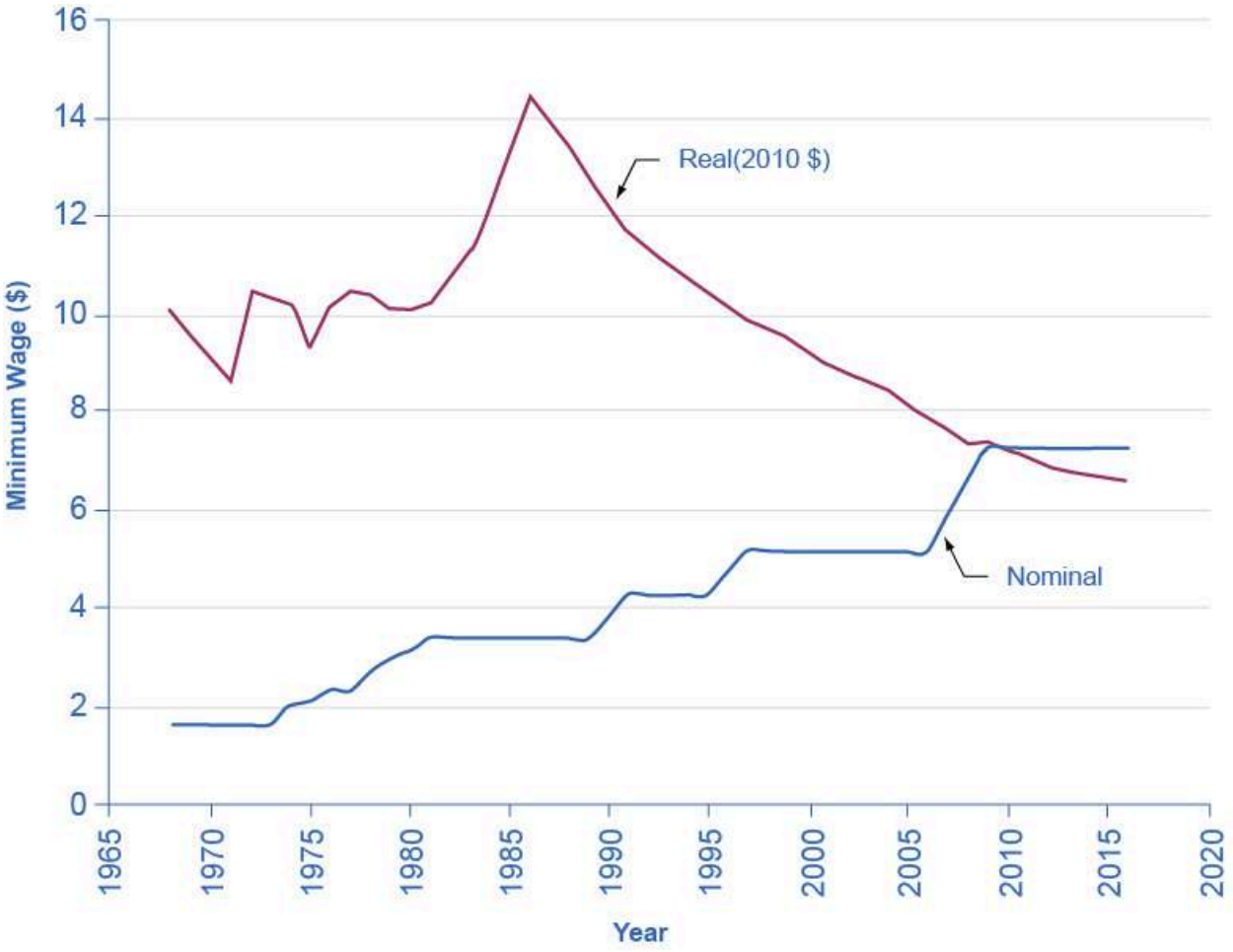


Figure 9.6 U.S. Minimum Wage and Inflation After adjusting for inflation, the federal minimum wage dropped more than 30 percent from 1967 to 2010, even though the nominal figure climbed from $1.40 to $7.25 per hour. Increases in the minimum wage in between 2008 and 2010 kept the decline from being worse—as it would have been if the wage had remained the same as it did from 1997 through 2007. Since 2010, the real minimum wage has continued to decline. (Sources: http://www.dol.gov/whd/minwage/chart.htm; http://data.bls.gov/cgi-bin/surveymost?cu)

One sizable group of people has often received a large share of their income in a form that does not increase over time: retirees who receive a private company pension. Most pensions have traditionally been set as a fixed nominal dollar amount per year at retirement. For this reason, economists call pensions "defined benefits" plans. Even if inflation is low, the combination of inflation and a fixed income can create a substantial problem over time. A person who retires on a fixed income at age 65 will find that losing just 1% to 2% of buying power per year to inflation compounds to a considerable loss of buying power after a decade or two.

Fortunately, pensions and other defined benefits retirement plans are increasingly rare, replaced instead by "defined contribution" plans, such as 401(k)s and 403(b)s. In these plans, the employer contributes a fixed amount to the worker's retirement account on a regular basis (usually every pay check). The employee often contributes as well. The worker invests these funds in a wide range of investment vehicles. These plans are tax deferred, and they are portable so that if the individual takes a job with a different employer, their 401(k) comes with them. To the extent that the investments made generate real rates of return, retirees do not suffer from the inflation costs of traditional pensioners.

However, ordinary people can sometimes benefit from the unintended redistributions of inflation. Consider someone who borrows $10,000 to buy a car at a fixed interest rate of 9%. If inflation is 3% at the time the loan is made, then he or she must repay the loan at a real interest rate of 6%. However, if inflation rises to 9%, then the real interest rate on the loan is zero. In this case, the borrower's benefit from inflation is the lender's loss. A borrower paying a fixed interest rate, who benefits from inflation, is just the flip side of an investor receiving a fixed interest rate, who suffers

from inflation. The lesson is that when interest rates are fixed, rises in the rate of inflation tend to penalize suppliers of financial capital, who receive repayment in dollars that are worth less because of inflation, while demanders of financial capital end up better off, because they can repay their loans in dollars that are worth less than originally expected.

The unintended redistributions of buying power that inflation causes may have a broader effect on society. America's widespread acceptance of market forces rests on a perception that people's actions have a reasonable connection to market outcomes. When inflation causes a retiree who built up a pension or invested at a fixed interest rate to suffer, however, while someone who borrowed at a fixed interest rate benefits from inflation, it is hard to believe that this outcome was deserved in any way. Similarly, when homeowners benefit from inflation because the price of their homes rises, while renters suffer because they are paying higher rent, it is hard to see any useful incentive effects. One of the reasons that the general public dislikes inflation is a sense that it makes economic rewards and penalties more arbitrary—and therefore likely to be perceived as unfair – even dangerous, as the next Clear It Up feature shows.

Clear It Up

Is there a connection between German hyperinflation and Hitler's rise to power?

Germany suffered an intense hyperinflation of its currency, the Mark, in the years after World War I, when the Weimar Republic in Germany resorted to printing money to pay its bills and the onset of the Great Depression created the social turmoil that Adolf Hitler was able to take advantage of in his rise to power. Shiller described the connection this way in a National Bureau of Economic Research 1996 Working Paper:

> A fact that is probably little known to young people today, even in Germany, is that the final collapse of the Mark in 1923, the time when the Mark's inflation reached astronomical levels (inflation of 35,974.9% in November 1923 alone, for an annual rate that month of 4.69×10^{28}%), came in the same month as did Hitler's Beer Hall Putsch, his Nazi Party's armed attempt to overthrow the German government. This failed putsch resulted in Hitler's imprisonment, at which time he wrote his book *Mein Kampf*, setting forth an inspirational plan for Germany's future, suggesting plans for world domination. . .
>
> . . . Most people in Germany today probably do not clearly remember these events; this lack of attention to it may be because its memory is blurred by the more dramatic events that succeeded it (the Nazi seizure of power and World War II). However, to someone living through these historical events in sequence . . . [the putsch] may have been remembered as vivid evidence of the potential effects of inflation.

Blurred Price Signals

Prices are the messengers in a market economy, conveying information about conditions of demand and supply. Inflation blurs those price messages. Inflation means that we perceive price signals more vaguely, like a radio program received with considerable static. If the static becomes severe, it is hard to tell what is happening.

In Israel, when inflation accelerated to an annual rate of 500% in 1985, some stores stopped posting prices directly on items, since they would have had to put new labels on the items or shelves every few days to reflect inflation. Instead, a shopper just took items from a shelf and went up to the checkout register to find out the price for that day. Obviously, this situation makes comparing prices and shopping for the best deal rather difficult. When the levels and changes of prices become uncertain, businesses and individuals find it harder to react to economic signals. In a world where inflation is at a high rate, but bouncing up and down to some extent, does a higher price of a good mean that inflation has risen, or that supply of that good has decreased, or that demand for that good has increased? Should a buyer of the good take the higher prices as an economic hint to start substituting other products—or have the prices of the substitutes risen by an equal amount? Should a seller of the good take a higher price as a reason to increase production—or is the higher price only a sign of a general inflation in which the prices of all inputs to production are rising as well? The true story will presumably become clear over time, but at a given moment, who can say?

High and variable inflation means that the incentives in the economy to adjust in response to changes in prices are weaker. Markets will adjust toward their equilibrium prices and quantities more erratically and slowly, and many individual markets will experience a greater chance of surpluses and shortages.

Problems of Long-Term Planning

Inflation can make long-term planning difficult. In discussing unintended redistributions, we considered the case of someone trying to plan for retirement with a pension that is fixed in nominal terms and a high rate of inflation. Similar problems arise for all people trying to save for retirement, because they must consider what their money will really buy several decades in the future when we cannot know the rate of future inflation.

Inflation, especially at moderate or high levels, will pose substantial planning problems for businesses, too. A firm can make money from inflation—for example, by paying bills and wages as late as possible so that it can pay in inflated dollars, while collecting revenues as soon as possible. A firm can also suffer losses from inflation, as in the case of a retail business that gets stuck holding too much cash, only to see inflation eroding the value of that cash. However, when a business spends its time focusing on how to profit by inflation, or at least how to avoid suffering from it, an inevitable tradeoff strikes: less time is spent on improving products and services or on figuring out how to make existing products and services more cheaply. An economy with high inflation rewards businesses that have found clever ways of profiting from inflation, which are not necessarily the businesses that excel at productivity, innovation, or quality of service.

In the short term, low or moderate levels of inflation may not pose an overwhelming difficulty for business planning, because costs of doing business and sales revenues may rise at similar rates. If, however, inflation varies substantially over the short or medium term, then it may make sense for businesses to stick to shorter-term strategies. The evidence as to whether relatively low rates of inflation reduce productivity is controversial among economists. There is some evidence that if inflation can be held to moderate levels of less than 3% per year, it need not prevent a nation's real economy from growing at a healthy pace. For some countries that have experienced hyperinflation of several thousand percent per year, an annual inflation rate of 20–30% may feel basically the same as zero. However, several economists have pointed to the suggestive fact that when U.S. inflation heated up in the early 1970s—to 10%—U.S. growth in productivity slowed down, and when inflation slowed down in the 1980s, productivity edged up again not long thereafter, as **Figure 9.7** shows.

Figure 9.7 U.S. Inflation Rate and U.S. Labor Productivity, 1961–2014 Over the last several decades in the United States, there have been times when rising inflation rates have been closely followed by lower productivity rates and lower inflation rates have corresponded to increasing productivity rates. As the graph shows, however, this correlation does not always exist.

Any Benefits of Inflation?

Although the economic effects of inflation are primarily negative, two countervailing points are worth noting. First,

the impact of inflation will differ considerably according to whether it is creeping up slowly at 0% to 2% per year, galloping along at 10% to 20% per year, or racing to the point of hyperinflation at, say, 40% per month. Hyperinflation can rip an economy and a society apart. An annual inflation rate of 2%, 3%, or 4%, however, is a long way from a national crisis. Low inflation is also better than deflation which occurs with severe recessions.

Second, economists sometimes argue that moderate inflation may help the economy by making wages in labor markets more flexible. The discussion in **Unemployment** pointed out that wages tend to be sticky in their downward movements and that unemployment can result. A little inflation could nibble away at real wages, and thus help real wages to decline if necessary. In this way, even if a moderate or high rate of inflation may act as sand in the gears of the economy, perhaps a low rate of inflation serves as oil for the gears of the labor market. This argument is controversial. A full analysis would have to account for all the effects of inflation. It does, however, offer another reason to believe that, all things considered, very low rates of inflation may not be especially harmful.

9.5 | Indexing and Its Limitations

By the end of this section, you will be able to:

- Explain the relationship between indexing and inflation
- Identify three ways the government can control inflation through macroeconomic policy

When a price, wage, or interest rate is adjusted automatically with inflation, economists use the term **indexed**. An indexed payment increases according to the index number that measures inflation. Those in private markets and government programs observe a wide range of indexing arrangements. Since the negative effects of inflation depend in large part on having inflation unexpectedly affect one part of the economy but not another—say, increasing the prices that people pay but not the wages that workers receive—indexing will take some of the sting out of inflation.

Indexing in Private Markets

In the 1970s and 1980s, labor unions commonly negotiated wage contracts that had **cost-of-living adjustments (COLAs)** which guaranteed that their wages would keep up with inflation. These contracts were sometimes written as, for example, COLA plus 3%. Thus, if inflation was 5%, the wage increase would automatically be 8%, but if inflation rose to 9%, the wage increase would automatically be 12%. COLAs are a form of indexing applied to wages.

Loans often have built-in inflation adjustments, too, so that if the inflation rate rises by two percentage points, then the interest rate that a financial institution charges on the loan rises by two percentage points as well. An **adjustable-rate mortgage (ARM)** is a type of loan that one can use to purchase a home in which the interest rate varies with the rate of inflation. Often, a borrower will be able receive a lower interest rate if borrowing with an ARM, compared to a fixed-rate loan. The reason is that with an ARM, the lender is protected against the risk that higher inflation will reduce the real loan payments, and so the risk premium part of the interest rate can be correspondingly lower.

A number of ongoing or long-term business contracts also have provisions that prices will adjust automatically according to inflation. Sellers like such contracts because they are not locked into a low nominal selling price if inflation turns out higher than expected. Buyers like such contracts because they are not locked into a high buying price if inflation turns out to be lower than expected. A contract with automatic adjustments for inflation in effect agrees on a real price for the borrower to pay, rather than a nominal price.

Indexing in Government Programs

Many government programs are indexed to inflation. The U.S. income tax code is designed so that as a person's income rises above certain levels, the tax rate on the marginal income earned rises as well. That is what the expression "move into a higher tax bracket" means. For example, according to the basic tax tables from the Internal Revenue Service, in 2017 a single person owed 10% of all taxable income from $0 to $9,325; 15% of all income from $9,326 to $37,950; 25% of all taxable income from $37,951 to $91,900; 28% of all taxable income from $91,901 to $191,650; 33% of all taxable income from $191,651 to $416,700; 35% of all taxable income from $416,701 to $418,400; and 39.6% of all income from $418,401 and above.

Because of the many complex provisions in the rest of the tax code, it is difficult to determine exactly the taxes an individual owes the government based on these numbers, but the numbers illustrate the basic theme that tax rates rise as the marginal dollar of income rises. Until the late 1970s, if nominal wages increased along with inflation, people were moved into higher tax brackets and owed a higher proportion of their income in taxes, even though their real

income had not risen. In 1981, the government eliminated this “bracket creep”. Now, the income levels where higher tax rates kick in are indexed to rise automatically with inflation.

The Social Security program offers two examples of indexing. Since the passage of the Social Security Indexing Act of 1972, the level of Social Security benefits increases each year along with the Consumer Price Index. Also, Social Security is funded by payroll taxes, which the government imposes on the income earned up to a certain amount—$117,000 in 2014. The government adjusts this level of income upward each year according to the rate of inflation, so that an indexed increase in the Social Security tax base accompanies the indexed rise in the benefit level.

As yet another example of a government program affected by indexing, in 1996 the U.S., government began offering indexed bonds. Bonds are means by which the U.S. government (and many private-sector companies as well) borrows money; that is, investors buy the bonds, and then the government repays the money with interest. Traditionally, government bonds have paid a fixed rate of interest. This policy gave a government that had borrowed an incentive to encourage inflation, because it could then repay its past borrowing in inflated dollars at a lower real interest rate. However, indexed bonds promise to pay a certain real rate of interest above whatever inflation rate occurs. In the case of a retiree trying to plan for the long term and worried about the risk of inflation, for example, indexed bonds that guarantee a rate of return higher than inflation—no matter the level of inflation—can be a very comforting investment.

Might Indexing Reduce Concern over Inflation?

Indexing may seem like an obviously useful step. After all, when individuals, firms, and government programs are indexed against inflation, then people can worry less about arbitrary redistributions and other effects of inflation.

However, some of the fiercest opponents of inflation express grave concern about indexing. They point out that indexing is always partial. Not every employer will provide COLAs for workers. Not all companies can assume that costs and revenues will rise in lockstep with the general rates of inflation. Not all interest rates for borrowers and savers will change to match inflation exactly. However, as partial inflation indexing spreads, the political opposition to inflation may diminish. After all, older people whose Social Security benefits are protected against inflation, or banks that have loaned their money with adjustable-rate loans, no longer have as much reason to care whether inflation heats up. In a world where some people are indexed against inflation and some are not, financially savvy businesses and investors may seek out ways to be protected against inflation, while the financially unsophisticated and small businesses may suffer from it most.

A Preview of Policy Discussions of Inflation

This chapter has focused on how economists measure inflation, historical experience with inflation, how to adjust nominal variables into real ones, how inflation affects the economy, and how indexing works. We have barely hinted at the causes of inflation, and we have not addressed government policies to deal with inflation. We will examine these issues in depth in other chapters. However, it is useful to offer a preview here.

We can sum up the cause of inflation in one phrase: Too many dollars chasing too few goods. The great surges of inflation early in the twentieth century came after wars, which are a time when government spending is very high, but consumers have little to buy, because production is going to the war effort. Governments also commonly impose price controls during wartime. After the war, the price controls end and pent-up buying power surges forth, driving up inflation. Otherwise, if too few dollars are chasing too many goods, then inflation will decline or even turn into deflation. Therefore, we typically associate slowdowns in economic activity, as in major recessions and the Great Depression, with a reduction in inflation or even outright deflation.

The policy implications are clear. If we are to avoid inflation, the amount of purchasing power in the economy must grow at roughly the same rate as the production of goods. Macroeconomic policies that the government can use to affect the amount of purchasing power—through taxes, spending, and regulation of interest rates and credit—can thus cause inflation to rise or reduce inflation to lower levels.

Bring it Home

A $550 Million Loaf of Bread?

As we will learn in **Money and Banking**, the existence of money provides enormous benefits to an economy. In a real sense, money is the lubrication that enhances the workings of markets. Money makes transactions easier. It allows people to find employment producing one product, then use the money earned to purchase the other products they need to live. However, too much money in circulation can lead to inflation. Extreme cases of governments recklessly printing money lead to hyperinflation. Inflation reduces the value of money. Hyperinflation, because money loses value so quickly, ultimately results in people no longer using money. The economy reverts to barter, or it adopts another country's more stable currency, like U.S. dollars. In the meantime, the economy literally falls apart as people leave jobs and fend for themselves because it is not worth the time to work for money that will be worthless in a few days.

Only national governments have the power to cause hyperinflation. Hyperinflation typically happens when government faces extraordinary demands for spending, which it cannot finance by taxes or borrowing. The only option is to print money—more and more of it. With more money in circulation chasing the same amount (or even fewer) goods and services, the only result is increasingly higher prices until the economy and/or the government collapses. This is why economists are generally wary of letting inflation spiral out of control.

KEY TERMS

adjustable-rate mortgage (ARM) a loan a borrower uses to purchase a home in which the interest rate varies with market interest rates

base year arbitrary year whose value as an index number economists define as 100; inflation from the base year to other years can easily be seen by comparing the index number in the other year to the index number in the base year—for example, 100; so, if the index number for a year is 105, then there has been exactly 5% inflation between that year and the base year

basket of goods and services a hypothetical group of different items, with specified quantities of each one meant to represent a "typical" set of consumer purchases, used as a basis for calculating how the price level changes over time

Consumer Price Index (CPI) a measure of inflation that U.S. government statisticians calculate based on the price level from a fixed basket of goods and services that represents the average consumer's purchases

core inflation index a measure of inflation typically calculated by taking the CPI and excluding volatile economic variables such as food and energy prices to better measure the underlying and persistent trend in long-term prices

cost-of-living adjustments (COLAs) a contractual provision that wage increases will keep up with inflation

deflation negative inflation; most prices in the economy are falling

Employment Cost Index a measure of inflation based on wages paid in the labor market

GDP deflator a measure of inflation based on the prices of all the GDP components

hyperinflation an outburst of high inflation that often occurs (although not exclusively) when economies shift from a controlled economy to a market-oriented economy

index number a unit-free number derived from the price level over a number of years, which makes computing inflation rates easier, since the index number has values around 100

indexed a price, wage, or interest rate is adjusted automatically for inflation

inflation a general and ongoing rise in price levels in an economy

International Price Index a measure of inflation based on the prices of merchandise that is exported or imported

Producer Price Index (PPI) a measure of inflation based on prices paid for supplies and inputs by producers of goods and services

quality/new goods bias inflation calculated using a fixed basket of goods over time tends to overstate the true rise in cost of living, because it does not account for improvements in the quality of existing goods or the invention of new goods

substitution bias an inflation rate calculated using a fixed basket of goods over time tends to overstate the true rise in the cost of living, because it does not take into account that the person can substitute away from goods whose prices rise considerably

KEY CONCEPTS AND SUMMARY

9.1 Tracking Inflation

Economists measure the price level by using a basket of goods and services and calculating how the total cost of

buying that basket of goods will increase over time. Economists often express the price level in terms of index numbers, which transform the cost of buying the basket of goods and services into a series of numbers in the same proportion to each other, but with an arbitrary base year of 100. We measure the inflation rate as the percentage change between price levels or index numbers over time.

9.2 How to Measure Changes in the Cost of Living

Measuring price levels with a fixed basket of goods will always have two problems: the substitution bias, by which a fixed basket of goods does not allow for buying more of what becomes relatively less expensive and less of what becomes relatively more expensive; and the quality/new goods bias, by which a fixed basket cannot account for improvements in quality and the advent of new goods. These problems can be reduced in degree—for example, by allowing the basket of goods to evolve over time—but we cannot totally eliminate them. The most commonly cited measure of inflation is the Consumer Price Index (CPI), which is based on a basket of goods representing what the typical consumer buys. The Core Inflation Index further breaks down the CPI by excluding volatile economic commodities. Several price indices are not based on baskets of consumer goods. The GDP deflator is based on all GDP components. The Producer Price Index is based on prices of supplies and inputs bought by producers of goods and services. An Employment Cost Index measures wage inflation in the labor market. An International Price Index is based on the prices of merchandise that is exported or imported.

9.3 How the U.S. and Other Countries Experience Inflation

In the U.S. economy, the annual inflation rate in the last two decades has typically been around 2% to 4%. The periods of highest inflation in the United States in the twentieth century occurred during the years after World Wars I and II, and in the 1970s. The period of lowest inflation—actually, with deflation—was the 1930s Great Depression.

9.4 The Confusion Over Inflation

Unexpected inflation will tend to hurt those whose money received, in terms of wages and interest payments, does not rise with inflation. In contrast, inflation can help those who owe money that they can pay in less valuable, inflated dollars. Low rates of inflation have relatively little economic impact over the short term. Over the medium and the long term, even low rates of inflation can complicate future planning. High rates of inflation can muddle price signals in the short term and prevent market forces from operating efficiently, and can vastly complicate long-term savings and investment decisions.

9.5 Indexing and Its Limitations

A payment is indexed if it is automatically adjusted for inflation. Examples of indexing in the private sector include wage contracts with cost-of-living adjustments (COLAs) and loan agreements like adjustable-rate mortgages (ARMs). Examples of indexing in the public sector include tax brackets and Social Security payments.

SELF-CHECK QUESTIONS

1. **Table 9.4** shows the fruit prices that the typical college student purchased from 2001 to 2004. What is the amount spent each year on the "basket" of fruit with the quantities shown in column 2?

Items	Qty	(2001) Price	(2001) Amount Spent	(2002) Price	(2002) Amount Spent	(2003) Price	(2003) Amount Spent	(2004) Price	(2004) Amount Spent
Apples	10	$0.50		$0.75		$0.85		$0.88	
Bananas	12	$0.20		$0.25		$0.25		$0.29	
Grapes	2	$0.65		$0.70		$0.90		$0.95	
Raspberries	1	$2.00		$1.90		$2.05		$2.13	$2.13
Total									

Table 9.4

2. Construct the price index for a "fruit basket" in each year using 2003 as the base year.

3. Compute the inflation rate for fruit prices from 2001 to 2004.

4. Edna is living in a retirement home where most of her needs are taken care of, but she has some discretionary spending. Based on the basket of goods in **Table 9.5**, by what percentage does Edna's cost of living increase between time 1 and time 2?

Items	Quantity	(Time 1) Price	(Time 2) Price
Gifts for grandchildren	12	$50	$60
Pizza delivery	24	$15	$16
Blouses	6	$60	$50
Vacation trips	2	$400	$420

Table 9.5

5. **How to Measure Changes in the Cost of Living** introduced a number of different price indices. Which price index would be best to use to adjust your paycheck for inflation?

6. The Consumer Price Index is subject to the substitution bias and the quality/new goods bias. Are the Producer Price Index and the GDP Deflator also subject to these biases? Why or why not?

7. Go to this **website (http://www.measuringworth.com/ppowerus/)** for the Purchasing Power Calculator at MeasuringWorth.com. How much money would it take today to purchase what one dollar would have bought in the year of your birth?

8. If inflation rises unexpectedly by 5%, would a state government that had recently borrowed money to pay for a new highway benefit or lose?

9. How should an increase in inflation affect the interest rate on an adjustable-rate mortgage?

10. A fixed-rate mortgage has the same interest rate over the life of the loan, whether the mortgage is for 15 or 30 years. By contrast, an adjustable-rate mortgage changes with market interest rates over the life of the mortgage. If inflation falls unexpectedly by 3%, what would likely happen to a homeowner with an adjustable-rate mortgage?

REVIEW QUESTIONS

11. How do economists use a basket of goods and services to measure the price level?

12. Why do economists use index numbers to measure the price level rather than dollar value of goods?

13. What is the difference between the price level and the rate of inflation?

14. Why does "substitution bias" arise if we calculate the inflation rate based on a fixed basket of goods?

15. Why does the "quality/new goods bias" arise if we calculate the inflation rate based on a fixed basket of goods?

16. What has been a typical range of inflation in the U.S. economy in the last decade or so?

17. Over the last century, during what periods was the U.S. inflation rate highest and lowest?

18. What is deflation?

19. Identify several parties likely to be helped and hurt by inflation.

20. What is indexing?

21. Name several forms of indexing in the private and public sector.

CRITICAL THINKING QUESTIONS

22. Inflation rates, like most statistics, are imperfect measures. Can you identify some ways that the inflation rate for fruit does not perfectly capture the rising price of fruit?

23. Given the federal budget deficit in recent years, some economists have argued that by adjusting Social Security payments for inflation using the CPI, Social Security is overpaying recipients. What is their argument, and do you agree or disagree with it?

24. Why is the GDP deflator not an accurate measure of inflation as it impacts a household?

25. Imagine that the government statisticians who calculate the inflation rate have been updating the basic basket of goods once every 10 years, but now they decide to update it every five years. How will this change affect the amount of substitution bias and quality/new goods bias?

26. Describe a situation, either a government policy situation, an economic problem, or a private sector situation, where using the CPI to convert from nominal to real would be more appropriate than using the GDP deflator.

27. Describe a situation, either a government policy situation, an economic problem, or a private sector situation, where using the GDP deflator to convert from nominal to real would be more appropriate than using the CPI.

28. Why do you think the U.S. experience with inflation over the last 50 years has been so much milder than in many other countries?

29. If, over time, wages and salaries on average rise at least as fast as inflation, why do people worry about how inflation affects incomes?

30. Who in an economy is the big winner from inflation?

31. If a government gains from unexpected inflation when it borrows, why would it choose to offer indexed bonds?

32. Do you think perfect indexing is possible? Why or why not?

PROBLEMS

33. The index number representing the price level changes from 110 to 115 in one year, and then from 115 to 120 the next year. Since the index number increases by five each year, is five the inflation rate each year? Is the inflation rate the same each year? Explain your answer.

34. The total price of purchasing a basket of goods in the United Kingdom over four years is: year 1=£940, year 2=£970, year 3=£1000, and year 4=£1070. Calculate two price indices, one using year 1 as the base year (set equal to 100) and the other using year 4 as the base year (set equal to 100). Then, calculate the inflation rate based on the first price index. If you had used the other price index, would you get a different inflation rate? If you are unsure, do the calculation and find out.

35. Within 1 or 2 percentage points, what has the U.S. inflation rate been during the last 20 years? Draw a graph to show the data.

36. If inflation rises unexpectedly by 5%, indicate for each of the following whether the economic actor is helped, hurt, or unaffected:

a. A union member with a COLA wage contract
b. Someone with a large stash of cash in a safe deposit box
c. A bank lending money at a fixed rate of interest
d. A person who is not due to receive a pay raise for another 11 months

37. Rosalie the Retiree knows that when she retires in 16 years, her company will give her a one-time payment of $20,000. However, if the inflation rate is 6% per year, how much buying power will that $20,000 have when measured in today's dollars? *Hint*: Start by calculating the rise in the price level over the 16 years.

10 | The International Trade and Capital Flows

Figure 10.1 A World of Money We are all part of the global financial system, which includes many different currencies. (Credit: modification of work by epSos.de/Flickr Creative Commons)

Bring it Home

More than Meets the Eye in the Congo

How much do you interact with the global financial system? Do you think not much? Think again. Suppose you take out a student loan, or you deposit money into your bank account. You just affected domestic savings and borrowing. Now say you are at the mall and buy two T-shirts “made in China,” and later contribute to a charity that helps refugees. What is the impact? You affected how much money flows into and out of the United States. If you open an IRA savings account and put money in an international mutual fund, you are involved in the flow of money overseas. While your involvement may not seem as influential as that of someone like the president, who can increase or decrease foreign aid and, thereby, have a huge impact on money flows in and out of the country, you do interact with the global financial system on a daily basis.

The balance of payments—a term you will meet soon—seems like a huge topic, but once you learn the specific components of trade and money, it all makes sense. Along the way, you may have to give up some common misunderstandings about trade and answer some questions: If a country is running a trade deficit, is that bad? Is a trade surplus good? For example, look at the Democratic Republic of the Congo (often referred to as “Congo”), a large country in Central Africa. In 2013, it ran a trade surplus of $1 billion, so it must be doing well, right? In contrast, the trade deficit in the United States was $508 billion in 2013. Do these figures suggest that the United States economy is performing worse than the Congolese economy? Not necessarily. The U.S. trade deficit tends to worsen as the economy strengthens. In contrast, high poverty rates in the Congo persist, and these rates are not going down even with the positive trade balance. Clearly, it is more complicated than

simply asserting that running a trade deficit is bad for the economy. You will learn more about these issues and others in this chapter.

Introduction to International Trade and Capital Flows

In this chapter, you will learn about:

- Measuring Trade Balances
- Trade Balances in Historical and International Context
- Trade Balances and Flows of Financial Capital
- The National Saving and Investment Identity
- The Pros and Cons of Trade Deficits and Surpluses
- The Difference between Level of Trade and the Trade Balance

The **balance of trade** (or trade balance) is any gap between a nation's dollar value of its exports, or what its producers sell abroad, and a nation's dollar value of imports, or the foreign-made products and services that households and businesses purchase. Recall from **The Macroeconomic Perspective** that if exports exceed imports, the economy has a trade surplus. If imports exceed exports, the economy has a trade deficit. If exports and imports are equal, then trade is balanced, but what happens when trade is out of balance and large trade surpluses or deficits exist?

Germany, for example, has had substantial trade surpluses in recent decades, in which exports have greatly exceeded imports. According to the Central Intelligence Agency's The World Factbook, in 2016, Germany ran a trade surplus of $295 billion. In contrast, the U.S. economy in recent decades has experienced large trade deficits, in which imports have considerably exceeded exports. In 2016, for example, U.S. imports exceeded exports by $502 billion.

A series of financial crises triggered by unbalanced trade can lead economies into deep recessions. These crises begin with large trade deficits. At some point, foreign investors become pessimistic about the economy and move their money to other countries. The economy then drops into deep recession, with real GDP often falling up to 10% or more in a single year. This happened to Mexico in 1995 when their GDP fell 8.1%. A number of countries in East Asia—Thailand, South Korea, Malaysia, and Indonesia—succumbed to the same economic illness in 1997–1998 (called the Asian Financial Crisis). In the late 1990s and into the early 2000s, Russia and Argentina had the identical experience. What are the connections between imbalances of trade in goods and services and the flows of international financial capital that set off these economic avalanches?

We will start by examining the balance of trade in more detail, by looking at some patterns of trade balances in the United States and around the world. Then we will examine the intimate connection between international flows of goods and services and international flows of financial capital, which to economists are really just two sides of the same coin. People often assume that trade surpluses like those in Germany must be a positive sign for an economy, while trade deficits like those in the United States must be harmful. As it turns out, both trade surpluses and deficits can be either good or bad. We will see why in this chapter.

10.1 | Measuring Trade Balances

By the end of this section, you will be able to:

- Explain merchandise trade balance, current account balance, and unilateral transfers
- Identify components of the U.S. current account balance
- Calculate the merchandise trade balance and current account balance using import and export data for a country

A few decades ago, it was common to track the solid or physical items that planes, trains, and trucks transported between countries as a way of measuring the balance of trade. Economists call this measurement is called the

merchandise trade balance. In most high-income economies, including the United States, goods comprise less than half of a country's total production, while services comprise more than half. The last two decades have seen a surge in international trade in services, powered by technological advances in telecommunications and computers that have made it possible to export or import customer services, finance, law, advertising, management consulting, software, construction engineering, and product design. Most global trade still takes the form of goods rather than services, and the government announces and the media prominently report the merchandise trade balance. Old habits are hard to break. Economists, however, typically rely on broader measures such as the balance of trade or the **current account balance** which includes other international flows of income and foreign aid.

Components of the U.S. Current Account Balance

Table 10.1 breaks down the four main components of the U.S. current account balance for the last quarter of 2015 (seasonally adjusted). The first line shows the merchandise trade balance; that is, exports and imports of goods. Because imports exceed exports, the trade balance in the final column is negative, showing a merchandise trade deficit. We can explain how the government collects this trade information in the following Clear It Up feature.

	Value of Exports (money flowing into the United States)	Value of Imports (money flowing out of the United States)	Balance
Goods	$410.0	$595.5	–$185.3
Services	$180.4	$122.3	$58.1
Income receipts and payments	$203.0	$152.4	$50.6
Unilateral transfers	$27.3	$64.4	–$37.1
Current account balance	$820.7	$934.4	–$113.7

Table 10.1 Components of the U.S. Current Account Balance for 2015 (in billions)

Clear It Up

How does the U.S. government collect trade statistics?

Do not confuse the balance of trade (which tracks imports and exports), with the current account balance, which includes not just exports and imports, but also income from investment and transfers.

The Bureau of Economic Analysis (BEA) within the U.S. Department of Commerce compiles statistics on the balance of trade using a variety of different sources. Merchandise importers and exporters must file monthly documents with the Census Bureau, which provides the basic data for tracking trade. To measure international trade in services—which can happen over a telephone line or computer network without shipping any physical goods—the BEA carries out a set of surveys. Another set of BEA surveys tracks investment flows, and there are even specific surveys to collect travel information from U.S. residents visiting Canada and Mexico. For measuring unilateral transfers, the BEA has access to official U.S. government spending on aid, and then also carries out a survey of charitable organizations that make foreign donations.

The BEA then cross-checks this information on international flows of goods and capital against other available data. For example, the Census Bureau also collects data from the shipping industry, which it can use to check the data on trade in goods. All companies involved in international flows of capital—including banks

and companies making financial investments like stocks—must file reports, which the U.S. Department of the Treasury ultimately checks. The BEA also can cross check information on foreign trade by looking at data collected by other countries on their foreign trade with the United States, and also at the data collected by various international organizations. Take these data sources, stir carefully, and you have the U.S. balance of trade statistics. Much of the statistics that we cite in this chapter come from these sources.

The second row of **Table 10.1** provides data on trade in services. Here, the U.S. economy is running a surplus. Although the level of trade in services is still relatively small compared to trade in goods, the importance of services has expanded substantially over the last few decades. For example, U.S. exports of services were equal to about one-half of U.S. exports of goods in 2015, compared to one-fifth in 1980.

The third component of the current account balance, labeled "income payments," refers to money that U.S. financial investors received on their foreign investments (money flowing into the United States) and payments to foreign investors who had invested their funds here (money flowing out of the United States). The reason for including this money on foreign investment in the overall measure of trade, along with goods and services, is that, from an economic perspective, income is just as much an economic transaction as car, wheat, or oil shipments: it is just trade that is happening in the financial capital market.

The final category of the current account balance is **unilateral transfers**, which are payments that government, private charities, or individuals make in which they send money abroad without receiving any direct good or service. Economic or military assistance from the U.S. government to other countries fits into this category, as does spending abroad by charities to address poverty or social inequalities. When an individual in the United States sends money overseas, as is the case with some immigrants, it is also counted in this category. The current account balance treats these unilateral payments like imports, because they also involve a stream of payments leaving the country. For the U.S. economy, unilateral transfers are almost always negative. This pattern, however, does not always hold. In 1991, for example, when the United States led an international coalition against Saddam Hussein's Iraq in the Gulf War, many other nations agreed that they would make payments to the United States to offset the U.S. war expenses. These payments were large enough that, in 1991, the overall U.S. balance on unilateral transfers was a positive $10 billion.

The following Work It Out feature steps you through the process of using the values for goods, services, and income payments to calculate the merchandise balance and the current account balance.

Work It Out

Calculating the Merchandise Balance and the Current Account Balance

	Exports (in $ billions)	Imports (in $ billions)	Balance
Goods			
Services			
Income payments			
Unilateral transfers			
Current account balance			

Table 10.2 Calculating Merchandise Balance and Current Account Balance

Use the information given below to fill in **Table 10.2**, and then calculate:

- The merchandise balance
- The current account balance

Known information:

- Unilateral transfers: $130
- Exports in goods: $1,046
- Exports in services: $509
- Imports in goods: $1,562
- Imports in services: $371
- Income received by U.S. investors on foreign stocks and bonds: $561
- Income received by foreign investors on U.S. assets: $472

Step 1. Focus on goods and services first. Enter the dollar amount of exports of both goods and services under the Export column.

Step 2. Enter imports of goods and services under the Import column.

Step 3. Under the Export column and in the row for Income payments, enter the financial flows of money coming back to the United States. U.S. investors are earning this income from abroad.

Step 4. Under the Import column and in the row for Income payments, enter the financial flows of money going out of the United States to foreign investors. Foreign investors are earning this money on U.S. assets, like stocks.

Step 5. Unilateral transfers are money flowing out of the United States in the form of, for example, military aid, foreign aid, and global charities. Because the money leaves the country, enter it under Imports and in the final column as well, as a negative.

Step 6. Calculate the trade balance by subtracting imports from exports in both goods and services. Enter this in the final Balance column. This can be positive or negative.

Step 7. Subtract the income payments flowing out of the country (under Imports) from the money coming back to the United States (under Exports) and enter this amount under the Balance column.

Step 8. Enter unilateral transfers as a negative amount under the Balance column.

Step 9. The merchandise trade balance is the difference between exports of goods and imports of goods—the first number under Balance.

Step 10. Now sum up your columns for Exports, Imports, and Balance. The final balance number is the current account balance.

The merchandise balance of trade is the difference between exports and imports. In this case, it is equal to $1,046 – $1,562, a trade deficit of –$516 billion. The current account balance is –$419 billion. See the completed **Table 10.3**.

	Value of Exports (money flowing into the United States)	Value of Imports (money flowing out of the United States)	Balance
Goods	$1,510.3	$2,272.9	–$762.6
Services	$750.9	$488.7	$262.2
Income receipts and payments	$782.9	$600.5	$182.4
Unilateral transfers	$128.6	$273.6	–$145.0

	Value of Exports (money flowing into the United States)	Value of Imports (money flowing out of the United States)	Balance
Current account balance	$3,172.7	$3,635.7	–$463.0

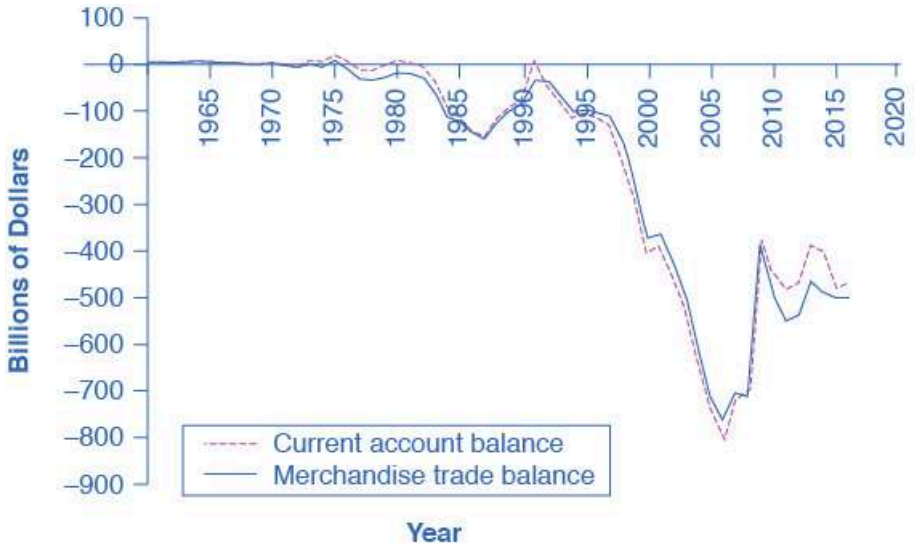
Table 10.3 Completed Merchandise Balance and Current Account Balance

10.2 | Trade Balances in Historical and International Context

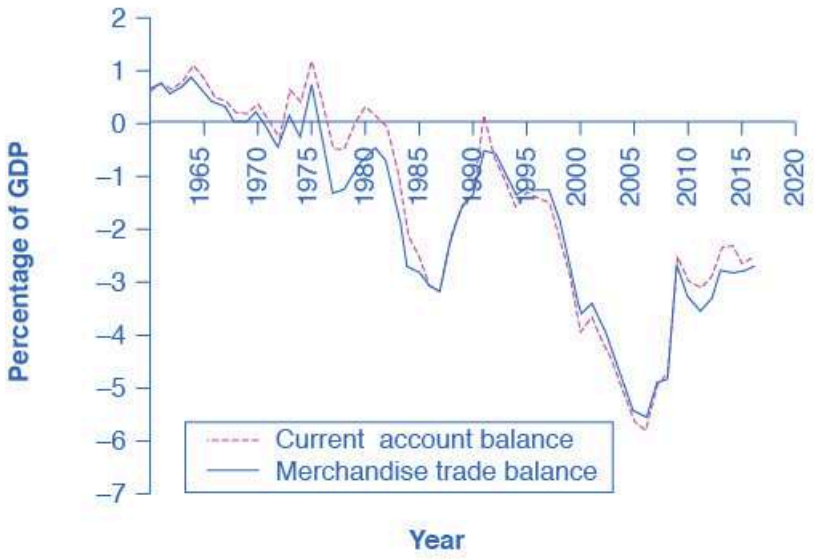
By the end of this section, you will be able to:

- Analyze graphs of the current account balance and the merchandise trade balance
- Identify patterns in U.S. trade surpluses and deficits
- Compare the U.S. trade surpluses and deficits to other countries' trade surpluses and deficits

We present the history of the U.S. current account balance in recent decades in several different ways. **Figure 10.2** (a) shows the current account balance and the merchandise trade balance in dollar terms. **Figure 10.2** (b) shows the current account balance and merchandise account balance yet again, this time as a share of the GDP for that year. By dividing the trade deficit in each year by GDP in that year, **Figure 10.2** (b) factors out both inflation and growth in the real economy.



(a) The current account and merchandise trade balance in nominal dollars



(b) The current account and merchandise trade balance as a percentage of GDP

Figure 10.2 Current Account Balance and Merchandise Trade Balance, 1960–2015 (a) The current account balance and the merchandise trade balance in billions of dollars from 1960 to 2015. If the lines are above zero dollars, the United States was running a positive trade balance and current account balance. If the lines fall below zero dollars, the United States is running a trade deficit and a deficit in its current account balance. (b) This shows the same items—trade balance and current account balance—in relationship to the size of the U.S. economy, or GDP, from 1960 to 2015.

By either measure, the U.S. balance of trade pattern is clear. From the 1960s into the 1970s, the U.S. economy had mostly small trade surpluses—that is, the graphs in **Figure 10.2** show positive numbers. However, starting in the 1980s, the trade deficit increased rapidly, and after a tiny surplus in 1991, the current account trade deficit became even larger in the late 1990s and into the mid-2000s. However, the trade deficit declined in 2009 after the recession had taken hold, then rebounded partially in 2010 and has remained stable up through 2016.

Table 10.4 shows the U.S. trade picture in 2013 compared with some other economies from around the world. While the U.S. economy has consistently run trade deficits in recent years, Japan and many European nations, among them France and Germany, have consistently run trade surpluses. Some of the other countries listed include Brazil, the largest economy in Latin America; Nigeria, along with South Africa competing to be the largest economy in Africa; and China, India, and Korea. The first column offers one measure of an economy's globalization: **exports of goods and services as a percentage of GDP**. The second column shows the trade balance. Usually, most countries have trade surpluses or deficits that are less than 5% of GDP. As you can see, the U.S. current account balance is –2.6% of GDP, while Germany's is 8.4% of GDP.

	Exports of Goods and Services	Current Account Balance
United States	17.6%	–2.6%
Japan	16.2%	3.1%
Germany	46.8%	8.4%
United Kingdom	27.2%	–5.4%
Canada	31.5%	–3.2%
Sweden	45.6%	5.2%
Korea	45.9%	7.7%
Mexico	35.4%	–2.9%
Brazil	13.0%	–3.3%
China	22.1%	3.0%
India	19.9%	–1.1%
Nigeria	10.7%	-3.3%
World	-	0.0%

Table 10.4 Level and Balance of Trade in 2015 (figures as a percentage of GDP, Source: http://data.worldbank.org/indicator/BN.CAB.XOKA.GD.ZS)

10.3 | Trade Balances and Flows of Financial Capital

By the end of this section, you will be able to:

- Explain the connection between trade balances and financial capital flows
- Calculate comparative advantage
- Explain balanced trade in terms of investment and capital flows

As economists see it, trade surpluses can be either good or bad, depending on circumstances, and trade deficits can be good or bad, too. The challenge is to understand how the international flows of goods and services are connected with international flows of **financial capital**. In this module we will illustrate the intimate connection between trade balances and flows of financial capital in two ways: a parable of trade between Robinson Crusoe and Friday, and a circular flow diagram representing flows of trade and payments.

A Two-Person Economy: Robinson Crusoe and Friday

To understand how economists view trade deficits and surpluses, consider a parable based on the story of Robinson Crusoe. Crusoe, as you may remember from the classic novel by Daniel Defoe first published in 1719, was

shipwrecked on a desert island. After living alone for some time, he is joined by a second person, whom he names Friday. Think about the balance of trade in a two-person economy like that of Robinson and Friday.

Robinson and Friday trade goods and services. Perhaps Robinson catches fish and trades them to Friday for coconuts, or Friday weaves a hat out of tree fronds and trades it to Robinson for help in carrying water. For a period of time, each individual trade is self-contained and complete. Because each trade is voluntary, both Robinson and Friday must feel that they are receiving fair value for what they are giving. As a result, each person's exports are always equal to his imports, and trade is always in balance between the two. Neither person experiences either a trade deficit or a trade surplus.

However, one day Robinson approaches Friday with a proposition. Robinson wants to dig ditches for an irrigation system for his garden, but he knows that if he starts this project, he will not have much time left to fish and gather coconuts to feed himself each day. He proposes that Friday supply him with a certain number of fish and coconuts for several months, and then after that time, he promises to repay Friday out of the extra produce that he will be able to grow in his irrigated garden. If Friday accepts this offer, then a trade imbalance comes into being. For several months, Friday will have a trade surplus: that is, he is exporting to Robinson more than he is importing. More precisely, he is giving Robinson fish and coconuts, and at least for the moment, he is receiving nothing in return. Conversely, Robinson will have a trade deficit, because he is importing more from Friday than he is exporting.

This parable raises several useful issues in thinking about what a trade deficit and a trade surplus really mean in economic terms. The first issue that this story of Robinson and Friday raises is this: Is it better to have a trade surplus or a trade deficit? The answer, as in any voluntary market interaction, is that if both parties agree to the transaction, then they may both be better off. Over time, if Robinson's irrigated garden is a success, it is certainly possible that both Robinson and Friday can benefit from this agreement.

The parable raises a second issue: What can go wrong? Robinson's proposal to Friday introduces an element of uncertainty. Friday is, in effect, making a loan of fish and coconuts to Robinson, and Friday's happiness with this arrangement will depend on whether Robinson repays that loan as planned, in full and on time. Perhaps Robinson spends several months loafing and never builds the irrigation system, or perhaps Robinson has been too optimistic about how much he will be able to grow with the new irrigation system, which turns out not to be very productive. Perhaps, after building the irrigation system, Robinson decides that he does not want to repay Friday as much as he previously agreed. Any of these developments will prompt a new round of negotiations between Friday and Robinson. Why the repayment failed is likely to shape Friday's attitude toward these renegotiations. If Robinson worked very hard and the irrigation system just did not increase production as intended, Friday may have some sympathy. If Robinson loafed or if he just refuses to pay, Friday may become irritated.

A third issue that the parable raises is that an intimate relationship exists between a trade deficit and international borrowing, and between a trade surplus and international lending. The size of Friday's trade surplus is exactly how much he is lending to Robinson. The size of Robinson's trade deficit is exactly how much he is borrowing from Friday. To economists, a trade surplus literally means the same thing as an outflow of financial capital, and a trade deficit literally means the same thing as an inflow of financial capital. This last insight is worth exploring in greater detail, which we will do in the following section.

The story of Robinson and Friday also provides a good opportunity to consider the law of comparative advantage, which you learn more about in the **International Trade** chapter. The following Work It Out feature steps you through calculating comparative advantage for the wheat and cloth traded between the United States and Great Britain in the 1800s.

Work It Out

Calculating Comparative Advantage

In the 1800s, the United States and Britain traded wheat and cloth. **Table 10.5** shows the varying hours of labor per unit of output.

	Wheat (in bushels)	Cloth (in yards)	Relative labor cost of wheat (Pw/Pc)	Relative labor cost of cloth (Pc/Pw)
United States	8	9	8/9	9/8
Britain	4	3	4/3	3/4

Table 10.5

Step 1. Observe from **Table 10.5** that, in the United States, it takes eight hours to supply a bushel of wheat and nine hours to supply a yard of cloth. In contrast, it takes four hours to supply a bushel of wheat and three hours to supply a yard of cloth in Britain.

Step 2. Recognize the difference between absolute advantage and comparative advantage. Britain has an absolute advantage (lowest cost) in each good, since it takes a lower amount of labor to make each good in Britain. Britain also has a comparative advantage in the production of cloth (lower opportunity cost in cloth (3/4 versus 9/8)). The United States has a comparative advantage in wheat production (lower opportunity cost of 8/9 versus 4/3).

Step 3. Determine the relative price of one good in terms of the other good. The price of wheat, in this example, is the amount of cloth you have to give up. To find this price, convert the hours per unit of wheat and cloth into units per hour. To do so, observe that in the United States it takes eight hours to make a bushel of wheat, so workers can process 1/8 of a bushel of wheat in an hour. It takes nine hours to make a yard of cloth in the United States, so workers can produce 1/9 of a yard of cloth in an hour. If you divide the amount of cloth (1/9 of a yard) by the amount of wheat you give up (1/8 of a bushel) in an hour, you find the price (8/9) of one good (wheat) in terms of the other (cloth).

The Balance of Trade as the Balance of Payments

The connection between trade balances and international flows of financial capital is so close that economists sometimes describe the balance of trade as the balance of payments. Each category of the current account balance involves a corresponding flow of payments between a given country and the rest of the world economy.

Figure 10.3 shows the flow of goods and services and payments between one country—the United States in this example—and the rest of the world. The top line shows U.S. exports of goods and services, while the second line shows financial payments from purchasers in other countries back to the U.S. economy. The third line then shows U.S. imports of goods, services, and investment, and the fourth line shows payments from the home economy to the rest of the world. Flow of goods and services (lines one and three) show up in the current account, while we find flow of funds (lines two and four) in the financial account.

The bottom four lines in **Figure 10.3** show the flows of investment income. In the first of the bottom lines, we see investments made abroad with funds flowing from the home country to the rest of the world. Investment income stemming from an investment abroad then runs in the other direction from the rest of the world to the home country. Similarly, we see on the bottom third line, an investment from the rest of the world into the home country and investment income (bottom fourth line) flowing from the home country to the rest of the world. We find the investment income (bottom lines two and four) in the current account, while investment to the rest of the world or into the home country (lines one and three) is in the financial account. This figure does not show unilateral transfers, the fourth item in the current account.

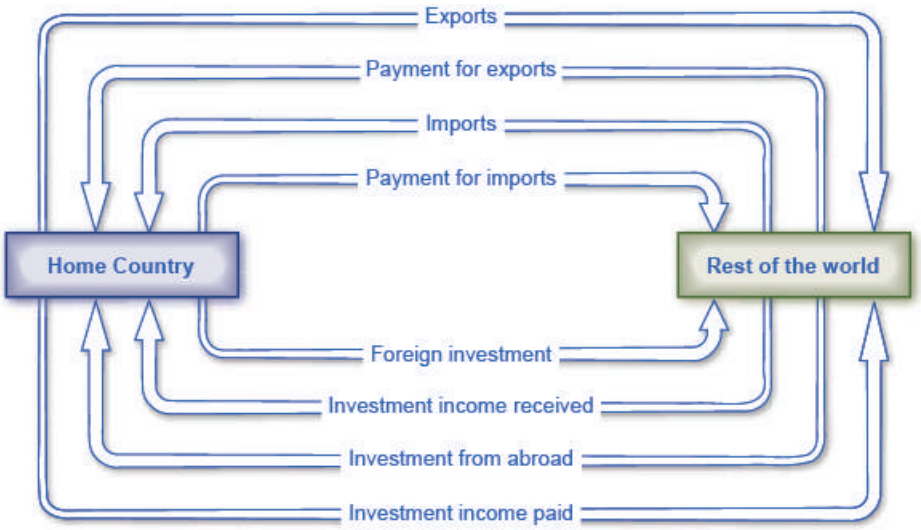


Figure 10.3 Flow of Investment Goods and Capital Each element of the current account balance involves a flow of financial payments between countries. The top line shows exports of goods and services leaving the home country; the second line shows the money that the home country receives for those exports. The third line shows imports that the home country receives; the fourth line shows the payments that the home country sent abroad in exchange for these imports.

A current account deficit means that, the country is a net borrower from abroad. Conversely, a positive current account balance means a country is a net lender to the rest of the world. Just like the parable of Robinson and Friday, the lesson is that a trade surplus means an overall outflow of financial investment capital, as domestic investors put their funds abroad, while a deficit in the current account balance is exactly equal to the overall or net inflow of foreign investment capital from abroad.

It is important to recognize that an inflow and outflow of foreign capital does not necessarily refer to a debt that governments owe to other governments, although government debt may be part of the picture. Instead, these international flows of financial capital refer to all of the ways in which private investors in one country may invest in another country—by buying real estate, companies, and financial investments like stocks and bonds.

10.4 | The National Saving and Investment Identity

By the end of this section, you will be able to:

- Explain the determinants of trade and current account balance
- Identify and calculate supply and demand for financial capital
- Explain how a nation's own level of domestic saving and investment determines a nation's balance of trade
- Predict the rising and falling of trade deficits based on a nation's saving and investment identity

The close connection between trade balances and international flows of savings and investments leads to a macroeconomic analysis. This approach views trade balances—and their associated flows of financial capital—in the context of the overall levels of savings and financial investment in the economy.

Understanding the Determinants of the Trade and Current Account Balance

The **national saving and investment identity** provides a useful way to understand the determinants of the trade and current account balance. In a nation’s financial capital market, the quantity of financial capital supplied at any given time must equal the quantity of financial capital demanded for purposes of making investments. What is on the supply and demand sides of financial capital? See the following Clear It Up feature for the answer to this question.

Clear It Up

What comprises the supply and demand of financial capital?

A country's national savings is the total of its domestic savings by household and companies (private savings) as well as the government (public savings). If a country is running a trade deficit, it means money from abroad is entering the country and the government considers it part of the supply of financial capital.

The demand for financial capital (money) represents groups that are borrowing the money. Businesses need to borrow to finance their investments in factories, materials, and personnel. When the federal government runs a budget deficit, it is also borrowing money from investors by selling Treasury bonds. Therefore, both business investment and the federal government can demand (or borrow) the supply of savings.

There are two main sources for the supply of financial capital in the U.S. economy: saving by individuals and firms, called S, and the inflow of financial capital from foreign investors, which is equal to the trade deficit (M – X), or imports minus exports. There are also two main sources of demand for financial capital in the U.S. economy: private sector investment, I, and government borrowing, where the government needs to borrow when government spending, G, is higher than the taxes collected, T. We can express this national savings and investment identity in algebraic terms:

$$\begin{aligned} \text{Supply of financial capital} &= \text{Demand for financial capital} \\ S + (M - X) &= I + (G - T) \end{aligned}$$

Again, in this equation, S is private savings, T is taxes, G is government spending, M is imports, X is exports, and I is investment. This relationship is true as a matter of definition because, for the macro economy, the quantity supplied of financial capital must be equal to the quantity demanded.

However, certain components of the national savings and investment identity can switch between the supply side and the demand side. Some countries, like the United States in most years since the 1970s, have budget deficits, which mean the government is spending more than it collects in taxes, and so the government needs to borrow funds. In this case, the government term would be $G - T > 0$, showing that spending is larger than taxes, and the government would be a demander of financial capital on the left-hand side of the equation (that is, a borrower), not a supplier of financial capital on the right-hand side. However, if the government runs a budget surplus so that the taxes exceed spending, as the U.S. government did from 1998 to 2001, then the government in that year was contributing to the supply of financial capital ($T - G > 0$), and would appear on the left (saving) side of the national savings and investment identity.

Similarly, if a national economy runs a trade surplus, the trade sector will involve an outflow of financial capital to other countries. A trade surplus means that the domestic financial capital is in surplus within a country and can be invested in other countries.

The fundamental notion that total quantity of financial capital demanded equals total quantity of financial capital supplied must always remain true. Domestic savings will always appear as part of the supply of financial capital and domestic investment will always appear as part of the demand for financial capital. However, the government and trade balance elements of the equation can move back and forth as either suppliers or demanders of financial capital, depending on whether government budgets and the trade balance are in surplus or deficit.

Domestic Saving and Investment Determine the Trade Balance

One insight from the national saving and investment identity is that a nation's own levels of domestic saving and investment determine a nation's balance of trade. To understand this point, rearrange the identity to put the balance of trade all by itself on one side of the equation. Consider first the situation with a trade deficit, and then the situation with a trade surplus.

In the case of a trade deficit, the national saving and investment identity can be rewritten as:

$$\begin{aligned} \text{Trade deficit} &= \text{Domestic investment} - \text{Private domestic saving} - \text{Government (or public) savings} \\ (M - X) &= I - S - (T - G) \end{aligned}$$

In this case, domestic investment is higher than domestic saving, including both private and government saving. The

only way that domestic investment can exceed domestic saving is if capital is flowing into a country from abroad. After all, that extra financial capital for investment has to come from someplace.

Now consider a trade *surplus* from the standpoint of the national saving and investment identity:

$$\begin{aligned} \text{Trade surplus} &= \text{Private domestic saving} + \text{Public saving} - \text{Domestic investment} \\ (X - M) &= S + (T - G) - I \end{aligned}$$

In this case, domestic savings (both private and public) is higher than domestic investment. That extra financial capital will be invested abroad.

This connection of domestic saving and investment to the trade balance explains why economists view the balance of trade as a fundamentally macroeconomic phenomenon. As the national saving and investment identity shows, the performance of certain sectors of an economy, like cars or steel, do not determine the trade balance. Further, whether the nation's trade laws and regulations encourage free trade or protectionism also does not determine the trade balance (see **Globalization and Protectionism**).

Exploring Trade Balances One Factor at a Time

The national saving and investment identity also provides a framework for thinking about what will cause trade deficits to rise or fall. Begin with the version of the identity that has domestic savings and investment on the left and the trade deficit on the right:

$$\begin{aligned} \text{Domestic investment} - \text{Private domestic savings} - \text{Public domestic savings} &= \text{Trade deficit} \\ I - S - (T - G) &= (M - X) \end{aligned}$$

Now, consider the factors on the left-hand side of the equation one at a time, while holding the other factors constant.

As a first example, assume that the level of domestic investment in a country rises, while the level of private and public saving remains unchanged. **Table 10.6** shows the result in the first row under the equation. Since the equality of the national savings and investment identity must continue to hold—it is, after all, an identity that must be true by definition—the rise in domestic investment will mean a higher trade deficit. This situation occurred in the U.S. economy in the late 1990s. Because of the surge of new information and communications technologies that became available, business investment increased substantially. A fall in private saving during this time and a rise in government saving more or less offset each other. As a result, the financial capital to fund that business investment came from abroad, which is one reason for the very high U.S. trade deficits of the late 1990s and early 2000s.

Domestic Investment	–	Private Domestic Savings	–	Public Domestic Savings	=	Trade Deficit
I	**–**	**S**	**–**	**(T – G)**	**=**	**(M – X)**
Up		No change		No change		Then M – X must rise
No change		Up		No change		Then M – X must fall
No change		No change		Down		Then M – X must rise

Table 10.6 Causes of a Changing Trade Balance

As a second scenario, assume that the level of domestic savings rises, while the level of domestic investment and public savings remain unchanged. In this case, the trade deficit would decline. As domestic savings rises, there would be less need for foreign financial capital to meet investment needs. For this reason, a policy proposal often made for reducing the U.S. trade deficit is to increase private saving—although exactly how to increase the overall rate of saving has proven controversial.

As a third scenario, imagine that the government budget deficit increased dramatically, while domestic investment

and private savings remained unchanged. This scenario occurred in the U.S. economy in the mid-1980s. The federal budget deficit increased from $79 billion in 1981 to $221 billion in 1986—an increase in the demand for financial capital of $142 billion. The current account balance collapsed from a surplus of $5 billion in 1981 to a deficit of $147 billion in 1986—an increase in the supply of financial capital from abroad of $152 billion. The connection at that time is clear: a sharp increase in government borrowing increased the U.S. economy's demand for financial capital, and foreign investors through the trade deficit primarily supplied that increase. The following Work It Out feature walks you through a scenario in which private domestic savings has to rise by a certain amount to reduce a trade deficit.

Work It Out

Solving Problems with the Saving and Investment Identity

Use the saving and investment identity to answer the following question: Country A has a trade deficit of $200 billion, private domestic savings of $500 billion, a government deficit of $200 billion, and private domestic investment of $500 billion. To reduce the $200 billion trade deficit by $100 billion, by how much does private domestic savings have to increase?

Step 1. Write out the savings investment formula solving for the trade deficit or surplus on the left:

$$(X - M) = S + (T - G) - I$$

Step 2. In the formula, put the amount for the trade deficit in as a negative number (X – M). The left side of your formula is now:

$$-200 = S + (T - G) - I$$

Step 3. Enter the private domestic savings (S) of $500 in the formula:

$$-200 = 500 + (T - G) - I$$

Step 4. Enter the private domestic investment (I) of $500 into the formula:

$$-200 = 500 + (T - G) - 500$$

Step 5. The government budget surplus or balance is represented by (T – G). Enter a budget deficit amount for (T – G) of –200:

$$-200 = 500 + (-200) - 500$$

Step 6. Your formula now is:

$$\begin{aligned}(X - M) &= S + (T - G) - I \\ -200 &= 500 + (-200) - 500\end{aligned}$$

The question is: To reduce your trade deficit (X – M) of –200 to –100 (in billions of dollars), by how much will savings have to rise?

$$\begin{aligned}(X - M) &= S + (T - G) - I \\ -100 &= S + (-200) - 500 \\ 600 &= S\end{aligned}$$

Step 7. Summarize the answer: Private domestic savings needs to rise by $100 billion, to a total of $600 billion, for the two sides of the equation to remain equal (–100 = –100).

Short-Term Movements in the Business Cycle and the Trade Balance

In the short run, whether an economy is in a recession or on the upswing can affect trade imbalances. A recession tends to make a trade deficit smaller, or a trade surplus larger, while a period of strong economic growth tends to make a trade deficit larger, or a trade surplus smaller.

As an example, note in **Figure 10.2** that the U.S. trade deficit declined by almost half from 2006 to 2009. One primary reason for this change is that during the recession, as the U.S. economy slowed down, it purchased fewer of all goods, including fewer imports from abroad. However, buying power abroad fell less, and so U.S. exports did not

fall by as much.

Conversely, in the mid-2000s, when the U.S. trade deficit became very large, a contributing short-term reason is that the U.S. economy was growing. As a result, there was considerable aggressive buying in the U.S. economy, including the buying of imports. Thus, a trade deficit (or a much lower trade surplus) often accompanies a rapidly growing domestic economy, while a trade surplus (or a much lower trade deficit) accompanies a slowing or recessionary domestic economy.

When the trade deficit rises, it necessarily means a greater net inflow of foreign financial capital. The national saving and investment identity teaches that the rest of the economy can absorb this inflow of foreign financial capital in several different ways. For example, reduced private savings could offset the additional inflow of financial capital from abroad, leaving domestic investment and public saving unchanged. Alternatively, the inflow of foreign financial capital could result in higher domestic investment, leaving private and public saving unchanged. Yet another possibility is that greater government borrowing could absorb the inflow of foreign financial capital, leaving domestic saving and investment unchanged. The national saving and investment identity does not specify which of these scenarios, alone or in combination, will occur—only that one of them must occur.

10.5 | The Pros and Cons of Trade Deficits and Surpluses

By the end of this section, you will be able to:

- Identify three ways in which borrowing money or running a trade deficit can result in a healthy economy
- Identify three ways in which borrowing money or running a trade deficit can result in a weaker economy

Because flows of trade always involve flows of financial payments, flows of international trade are actually the same as flows of international financial capital. The question of whether trade deficits or surpluses are good or bad for an economy is, in economic terms, exactly the same question as whether it is a good idea for an economy to rely on net inflows of financial capital from abroad or to make net investments of financial capital abroad. Conventional wisdom often holds that borrowing money is foolhardy, and that a prudent country, like a prudent person, should always rely on its own resources. While it is certainly possible to borrow too much—as anyone with an overloaded credit card can testify—borrowing at certain times can also make sound economic sense. For both individuals and countries, there is no economic merit in a policy of abstaining from participation in financial capital markets.

It makes economic sense to borrow when you are buying something with a long-run payoff; that is, when you are making an investment. For this reason, it can make economic sense to borrow for a college education, because the education will typically allow you to earn higher wages, and so to repay the loan and still come out ahead. It can also make sense for a business to borrow in order to purchase a machine that will last 10 years, as long as the machine will increase output and profits by more than enough to repay the loan. Similarly, it can make economic sense for a national economy to borrow from abroad, as long as it wisely invests the money in ways that will tend to raise the nation’s economic growth over time. Then, it will be possible for the national economy to repay the borrowed money over time and still end up better off than before.

One vivid example of a country that borrowed heavily from abroad, invested wisely, and did perfectly well is the United States during the nineteenth century. The United States ran a trade deficit in 40 of the 45 years from 1831 to 1875, which meant that it was importing capital from abroad over that time. However, that financial capital was mostly invested in projects like railroads that brought a substantial economic payoff. (See the following Clear It Up feature for more on this.)

A more recent example along these lines is the experience of South Korea, which had trade deficits during much of the 1970s—and so was an importer of capital over that time. However, South Korea also had high rates of investment in physical plant and equipment, and its economy grew rapidly. From the mid-1980s into the mid-1990s, South Korea often had trade surpluses—that is, it was repaying its past borrowing by sending capital abroad.

In contrast, some countries have run large trade deficits, borrowed heavily in global capital markets, and ended up in all kinds of trouble. Two specific sorts of trouble are worth examining. First, a borrower nation can find itself in a bind if it does not invest the incoming funds from abroad in a way that leads to increased productivity. Several of

Latin America's large economies, including Mexico and Brazil, ran large trade deficits and borrowed heavily from abroad in the 1970s, but the inflow of financial capital did not boost productivity sufficiently, which meant that these countries faced enormous troubles repaying the money borrowed when economic conditions shifted during the 1980s. Similarly, it appears that a number of African nations that borrowed foreign funds in the 1970s and 1980s did not invest in productive economic assets. As a result, several of those countries later faced large interest payments, with no economic growth to show for the borrowed funds.

Clear It Up

Are trade deficits always harmful?

For most years of the nineteenth century, U.S. imports exceeded exports and the U.S. economy had a trade deficit. Yet the string of trade deficits did not hold back the economy at all. Instead, the trade deficits contributed to the strong economic growth that gave the U.S. economy the highest per capita GDP in the world by around 1900.

The U.S. trade deficits meant that the U.S. economy was receiving a net inflow of foreign capital from abroad. Much of that foreign capital flowed into two areas of investment—railroads and public infrastructure like roads, water systems, and schools—which were important to helping the U.S. economy grow.

We should not overstate the effect of foreign investment capital on U.S. economic growth. In most years the foreign financial capital represented no more than 6–10% of the funds that the government used for overall physical investment in the economy. Nonetheless, the trade deficit and the accompanying investment funds from abroad were clearly a help, not a hindrance, to the U.S. economy in the nineteenth century.

A second “trouble” is: What happens if the foreign money flows in, and then suddenly flows out again? We raised this scenario at the start of the chapter. In the mid-1990s, a number of countries in East Asia—Thailand, Indonesia, Malaysia, and South Korea—ran large trade deficits and imported capital from abroad. However, in 1997 and 1998 many foreign investors became concerned about the health of these economies, and quickly pulled their money out of stock and bond markets, real estate, and banks. The extremely rapid departure of that foreign capital staggered the banking systems and economies of these countries, plunging them into deep recession. We investigate and discuss the links between international capital flows, banks, and recession in **The Impacts of Government Borrowing**.

While a trade deficit is not always harmful, there is no guarantee that running a trade surplus will bring robust economic health. For example, Germany and Japan ran substantial trade surpluses for most of the last three decades. Regardless of their persistent trade surpluses, both countries have experienced occasional recessions and neither country has had especially robust annual growth in recent years. Read more about Japan’s trade surplus in the next Clear It Up feature.

Link It Up

Watch this video (http://openstaxcollege.org/l/tradedeficit) on whether or not trade deficit is good for the economy.

The sheer size and persistence of the U.S. trade deficits and inflows of foreign capital since the 1980s are a legitimate cause for concern. The huge U.S. economy will not be destabilized by an outflow of international capital as easily as, say, the comparatively tiny economies of Thailand and Indonesia were in 1997–1998. Even an economy that is not knocked down, however, can still be shaken. American policymakers should certainly be paying attention to those cases where a pattern of extensive and sustained current account deficits and foreign borrowing has gone badly—if only as a cautionary tale.

Clear It Up

Are trade surpluses always beneficial? Considering Japan since the 1990s.

Perhaps no economy around the world is better known for its trade surpluses than Japan. Since 1990, the size of these surpluses has often been near $100 billion per year. When Japan's economy was growing vigorously in the 1960s and 1970s, many, especially non-economists, described its large trade surpluses either a cause or a result of its robust economic health. However, from a standpoint of economic growth, Japan's economy has been teetering in and out of recession since 1990, with real GDP growth averaging only about 1% per year, and an unemployment rate that has been creeping higher. Clearly, a whopping trade surplus is no guarantee of economic good health.

Instead, Japan's trade surplus reflects that Japan has a very high rate of domestic savings, more than the Japanese economy can invest domestically, and so it invests the extra funds abroad. In Japan's slow economy, consumption of imports is relatively low, and the growth of consumption is relatively slow. Thus, Japan's exports continually exceed its imports, leaving the trade surplus continually high. Recently, Japan's trade surpluses began to deteriorate. In 2013, Japan ran a trade deficit due to the high cost of imported oil. By 2015, Japan again had a surplus.

10.6 | The Difference between Level of Trade and the Trade Balance

By the end of this section, you will be able to:

- Identify three factors that influence a country's level of trade
- Differentiate between balance of trade and level of trade

A nation's *level* of trade may at first sound like much the same issue as the *balance* of trade, but these two are actually quite separate. It is perfectly possible for a country to have a very high level of trade—measured by its exports of goods and services as a share of its GDP—while it also has a near-balance between exports and imports. A high level of trade indicates that the nation exports a good portion of its production. It is also possible for a country's trade to be a relatively low share of GDP, relative to global averages, but for the imbalance between its exports and its imports to be quite large. We emphasized this general theme earlier in **Measuring Trade Balances**, which offered some illustrative figures on trade levels and balances.

A country's level of trade tells how much of its production it exports. We measure this by the percent of exports out of GDP. It indicates the degree of an economy's globalization. Some countries, such as Germany, have a high level of trade—they export almost 50% of their total production. The balance of trade tells us if the country is running a trade surplus or trade deficit. A country can have a low level of trade but a high trade deficit. (For example, the United States only exports 13% of GDP, but it has a trade deficit of over $500 billion.)

Three factors strongly influence a nation's level of trade: the size of its economy, its geographic location, and its history of trade. Large economies like the United States can do much of their trading internally, while small economies like Sweden have less ability to provide what they want internally and tend to have higher ratios of exports and imports to GDP. Nations that are neighbors tend to trade more, since costs of transportation and communication

are lower. Moreover, some nations have long and established patterns of international trade, while others do not.

Consequently, a relatively small economy like Sweden, with many nearby trading partners across Europe and a long history of foreign trade, has a high level of trade. Brazil and India, which are fairly large economies that have often sought to inhibit trade in recent decades, have lower levels of trade; whereas, the United States and Japan are extremely large economies that have comparatively few nearby trading partners. Both countries actually have quite low levels of trade by world standards. The ratio of exports to GDP in either the United States or in Japan is about half of the world average.

The balance of trade is a separate issue from the level of trade. The United States has a low level of trade, but had enormous trade deficits for most years from the mid-1980s into the 2000s. Japan has a low level of trade by world standards, but has typically shown large trade surpluses in recent decades. Nations like Germany and the United Kingdom have medium to high levels of trade by world standards, but Germany had a moderate trade surplus in 2015, while the United Kingdom had a moderate trade deficit. Their trade picture was roughly in balance in the late 1990s. Sweden had a high level of trade and a moderate trade surplus in 2015, while Mexico had a high level of trade and a moderate trade deficit that same year.

In short, it is quite possible for nations with a relatively low level of trade, expressed as a percentage of GDP, to have relatively large trade deficits. It is also quite possible for nations with a near balance between exports and imports to worry about the consequences of high levels of trade for the economy. It is not inconsistent to believe that a high level of trade is potentially beneficial to an economy, because of the way it allows nations to play to their comparative advantages, and to also be concerned about any macroeconomic instability caused by a long-term pattern of large trade deficits. The following Clear It Up feature discusses how this sort of dynamic played out in Colonial India.

Clear It Up

Are trade surpluses always beneficial? Considering Colonial India.

India was formally under British rule from 1858 to 1947. During that time, India consistently had trade surpluses with Great Britain. Anyone who believes that trade surpluses are a sign of economic strength and dominance while trade deficits are a sign of economic weakness must find this pattern odd, since it would mean that colonial India was successfully dominating and exploiting Great Britain for almost a century—which was not true.

Instead, India's trade surpluses with Great Britain meant that each year there was an overall flow of financial capital from India to Great Britain. In India, many heavily criticized this financial capital flow as the "drain," and they viewed eliminating the financial capital drain as one of the many reasons why India would benefit from achieving independence.

Final Thoughts about Trade Balances

Trade deficits can be a good or a bad sign for an economy, and trade surpluses can be a good or a bad sign. Even a trade balance of zero—which just means that a nation is neither a net borrower nor lender in the international economy—can be either a good or bad sign. The fundamental economic question is not whether a nation's economy is borrowing or lending at all, but whether the particular borrowing or lending in the particular economic conditions of that country makes sense.

It is interesting to reflect on how public attitudes toward trade deficits and surpluses might change if we could somehow change the labels that people and the news media affix to them. If we called a trade deficit "attracting foreign financial capital"—which accurately describes what a trade deficit means—then trade deficits might look more attractive. Conversely, if we called a trade surplus "shipping financial capital abroad"—which accurately captures what a trade surplus does—then trade surpluses might look less attractive. Either way, the key to understanding trade balances is to understand the relationships between flows of trade and flows of international payments, and what these relationships imply about the causes, benefits, and risks of different kinds of trade balances. The first step along this journey of understanding is to move beyond knee-jerk reactions to terms like "trade surplus,"

“trade balance,” and “trade deficit.”

Bring it Home

More than Meets the Eye in the Congo

Now that you see the big picture, you undoubtedly realize that all of the economic choices you make, such as depositing savings or investing in an international mutual fund, do influence the flow of goods and services as well as the flows of money around the world.

You now know that a trade surplus does not necessarily tell us whether an economy is performing well or not. The Democratic Republic of the Congo ran a trade surplus in 2013, as we learned in the beginning of the chapter. Yet its current account balance was –$2.8 billion. However, the return of political stability and the rebuilding in the aftermath of the civil war there has meant a flow of investment and financial capital into the country. In this case, a negative current account balance means the country is being rebuilt—and that is a good thing.

KEY TERMS

balance of trade (trade balance) the gap, if any, between a nation's exports and imports

current account balance a broad measure of the balance of trade that includes trade in goods and services, as well as international flows of income and foreign aid

exports of goods and services as a percentage of GDP the dollar value of exports divided by the dollar value of a country's GDP

financial capital the international flows of money that facilitates trade and investment

merchandise trade balance the balance of trade looking only at goods

national savings and investment identity the total of private savings and public savings (a government budget surplus)

unilateral transfers "one-way payments" that governments, private entities, or individuals make that they sent abroad with nothing received in return

KEY CONCEPTS AND SUMMARY

10.1 Measuring Trade Balances

The trade balance measures the gap between a country's exports and its imports. In most high-income economies, goods comprise less than half of a country's total production, while services comprise more than half. The last two decades have seen a surge in international trade in services; however, most global trade still takes the form of goods rather than services. The current account balance includes the trade in goods, services, and money flowing into and out of a country from investments and unilateral transfers.

10.2 Trade Balances in Historical and International Context

The United States developed large trade surpluses in the early 1980s, swung back to a tiny trade surplus in 1991, and then had even larger trade deficits in the late 1990s and early 2000s. As we will see below, a trade deficit necessarily means a net inflow of financial capital from abroad, while a trade surplus necessarily means a net outflow of financial capital from an economy to other countries.

10.3 Trade Balances and Flows of Financial Capital

International flows of goods and services are closely connected to the international flows of financial capital. A current account deficit means that, after taking all the flows of payments from goods, services, and income together, the country is a net borrower from the rest of the world. A current account surplus is the opposite and means the country is a net lender to the rest of the world.

10.4 The National Saving and Investment Identity

The national saving and investment identity is based on the relationship that the total quantity of financial capital supplied from all sources must equal the total quantity of financial capital demanded from all sources. If S is private saving, T is taxes, G is government spending, M is imports, X is exports, and I is investment, then for an economy with a current account deficit and a budget deficit:

$$\begin{aligned} \text{Supply of financial capital} &= \text{Demand for financial capital} \\ S + (M - X) &= I + (G - T) \end{aligned}$$

A recession tends to increase the trade balance (meaning a higher trade surplus or lower trade deficit), while economic boom will tend to decrease the trade balance (meaning a lower trade surplus or a larger trade deficit).

10.5 The Pros and Cons of Trade Deficits and Surpluses

Trade surpluses are no guarantee of economic health, and trade deficits are no guarantee of economic weakness. Either trade deficits or trade surpluses can work out well or poorly, depending on whether a government wisely invests the corresponding flows of financial capital.

10.6 The Difference between Level of Trade and the Trade Balance

There is a difference between the level of a country's trade and the balance of trade. The government measures its level of trade by the percentage of exports out of GDP, or the size of the economy. Small economies that have nearby trading partners and a history of international trade will tend to have higher levels of trade. Larger economies with few nearby trading partners and a limited history of international trade will tend to have lower levels of trade. The level of trade is different from the trade balance. The level of trade depends on a country's history of trade, its geography, and the size of its economy. A country's balance of trade is the dollar difference between its exports and imports.

Trade deficits and trade surpluses are not necessarily good or bad—it depends on the circumstances. Even if a country is borrowing, if it invests that money in productivity-boosting investments it can lead to an improvement in long-term economic growth.

SELF-CHECK QUESTIONS

1. If foreign investors buy more U.S. stocks and bonds, how would that show up in the current account balance?

2. If the trade deficit of the United States increases, how is the current account balance affected?

3. State whether each of the following events involves a financial flow to the Mexican economy or a financial flow out of the Mexican economy:

a. Mexico imports services from Japan
b. Mexico exports goods to Canada
c. U.S. investors receive a return from past financial investments in Mexico

4. In what way does comparing a country's exports to GDP reflect its degree of globalization?

5. At one point Canada's GDP was $1,800 billion and its exports were $542 billion. What was Canada's export ratio at this time?

6. The GDP for the United States is $18,036 billion and its current account balance is –$484 billion. What percent of GDP is the current account balance?

7. Why does the trade balance and the current account balance track so closely together over time?

8. State whether each of the following events involves a financial flow to the U.S. economy or away from the U.S. economy:

a. Export sales to Germany
b. Returns paid on past U.S. financial investments in Brazil
c. Foreign aid from the U.S. government to Egypt
d. Imported oil from the Russian Federation
e. Japanese investors buying U.S. real estate

9. How does the bottom portion of **Figure 10.3**, showing the international flow of investments and capital, differ from the upper portion?

10. Explain the relationship between a current account deficit or surplus and the flow of funds.

11. Using the national savings and investment identity, explain how each of the following changes (*ceteris paribus*) will increase or decrease the trade balance:

a. A lower domestic savings rate
b. The government changes from running a budget surplus to running a budget deficit
c. The rate of domestic investment surges

12. If a country is running a government budget surplus, why is (T – G) on the left side of the saving-investment identity?

13. What determines the size of a country's trade deficit?

14. If domestic investment increases, and there is no change in the amount of private and public saving, what must happen to the size of the trade deficit?

15. Why does a recession cause a trade deficit to increase?

16. Both the United States and global economies are booming. Will U.S. imports and/or exports increase?

17. For each of the following, indicate which type of government spending would justify a budget deficit and which would not.

a. Increased federal spending on Medicare
b. Increased spending on education
c. Increased spending on the space program
d. Increased spending on airports and air traffic control

18. How did large trade deficits hurt the East Asian countries in the mid 1980s? (Recall that trade deficits are equivalent to inflows of financial capital from abroad.)

19. Describe a scenario in which a trade surplus benefits an economy and one in which a trade surplus is occurring in an economy that performs poorly. What key factor or factors are making the difference in the outcome that results from a trade surplus?

20. The United States exports 14% of GDP while Germany exports about 50% of its GDP. Explain what that means.

21. Explain briefly whether each of the following would be more likely to lead to a higher level of trade for an economy, or a greater imbalance of trade for an economy.

a. Living in an especially large country
b. Having a domestic investment rate much higher than the domestic savings rate
c. Having many other large economies geographically nearby
d. Having an especially large budget deficit
e. Having countries with a tradition of strong protectionist legislation shutting out imports

REVIEW QUESTIONS

22. If imports exceed exports, is it a trade deficit or a trade surplus? What about if exports exceed imports?

23. What is included in the current account balance?

24. In recent decades, has the U.S. trade balance usually been in deficit, surplus, or balanced?

25. Does a trade surplus mean an overall inflow of financial capital to an economy, or an overall outflow of financial capital? What about a trade deficit?

26. What are the two main sides of the national savings and investment identity?

27. What are the main components of the national savings and investment identity?

28. When is a trade deficit likely to work out well for an economy? When is it likely to work out poorly?

29. Does a trade surplus help to guarantee strong economic growth?

30. What three factors will determine whether a nation has a higher or lower share of trade relative to its GDP?

31. What is the difference between trade deficits and balance of trade?

CRITICAL THINKING QUESTIONS

32. Occasionally, a government official will argue that a country should strive for both a trade surplus and a healthy inflow of capital from abroad. Explain why such a statement is economically impossible.

33. A government official announces a new policy. The country wishes to eliminate its trade deficit, but will strongly encourage financial investment from foreign firms. Explain why such a statement is contradictory.

34. If a country is a big exporter, is it more exposed to global financial crises?

35. If countries reduced trade barriers, would the international flows of money increase?

36. Is it better for your country to be an international lender or borrower?

37. Many think that the size of a trade deficit is due to a lack of competitiveness of domestic sectors, such as autos. Explain why this is not true.

38. If you observed a country with a rapidly growing trade surplus over a period of a year or so, would you be more likely to believe that the country's economy was in a period of recession or of rapid growth? Explain.

39. Occasionally, a government official will argue that a country should strive for both a trade surplus and a healthy inflow of capital from abroad. Is this possible?

40. What is more important, a country's current account balance or GDP growth? Why?

41. Will nations that are more involved in foreign trade tend to have higher trade imbalances, lower trade imbalances, or is the pattern unpredictable?

42. Some economists warn that the persistent trade deficits and a negative current account balance that the United States has run will be a problem in the long run. Do you agree or not? Explain your answer.

PROBLEMS

43. In 2001, the United Kingdom's economy exported goods worth £192 billion and services worth another £77 billion. It imported goods worth £225 billion and services worth £66 billion. Receipts of income from abroad were £140 billion while income payments going abroad were £131 billion. Government transfers from the United Kingdom to the rest of the world were £23 billion, while various U.K government agencies received payments of £16 billion from the rest of the world.

a. Calculate the U.K. merchandise trade deficit for 2001.
b. Calculate the current account balance for 2001.
c. Explain how you decided whether payments on foreign investment and government transfers counted on the positive or the negative side of the current account balance for the United Kingdom in 2001.

44. Imagine that the U.S. economy finds itself in the following situation: a government budget deficit of $100 billion, total domestic savings of $1,500 billion, and total domestic physical capital investment of $1,600 billion. According to the national saving and investment identity, what will be the current account balance? What will be the current account balance if investment rises by $50 billion, while the budget deficit and national savings remain the same?

45. **Table 10.7** provides some hypothetical data on macroeconomic accounts for three countries represented by A, B, and C and measured in billions of currency units. In **Table 10.7**, private household saving is SH, tax revenue is T, government spending is G, and investment spending is I.

	A	B	C
SH	700	500	600
T	00	500	500
G	600	350	650
I	800	400	450

Table 10.7 Macroeconomic Accounts

a. Calculate the trade balance and the net inflow of foreign saving for each country.
b. State whether each one has a trade surplus or deficit (or balanced trade).
c. State whether each is a net lender or borrower internationally and explain.

46. Imagine that the economy of Germany finds itself in the following situation: the government budget has a surplus of 1% of Germany's GDP; private savings is 20% of GDP; and physical investment is 18% of GDP.

a. Based on the national saving and investment identity, what is the current account balance?
b. If the government budget surplus falls to zero, how will this affect the current account balance?

11 | The Aggregate Demand/ Aggregate Supply Model

Figure 11.1 New Home Construction At the peak of the housing bubble, many people across the country were able to secure the loans necessary to build new houses. (Credit: modification of work by Tim Pierce/Flickr Creative Commons)

Bring it Home

From Housing Bubble to Housing Bust

The United States experienced rising home ownership rates for most of the last two decades. Between 1990 and 2006, the U.S. housing market grew. Homeownership rates grew from 64% to a high of over 69% between 2004 and 2005. For many people, this was a period in which they could either buy first homes or buy a larger and more expensive home. During this time mortgage values tripled. Housing became more accessible to Americans and was considered to be a safe financial investment. **Figure 11.2** shows how new single family home sales peaked in 2005 at 107,000 units.

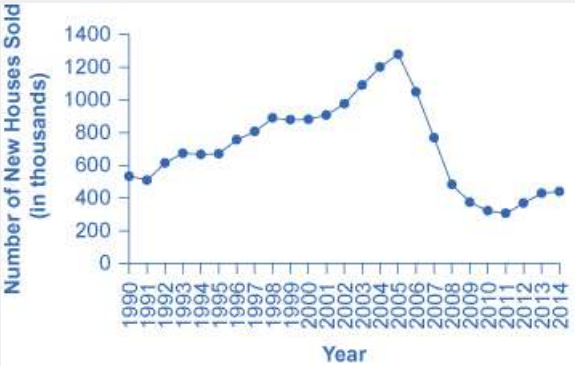


Figure 11.2 New Single Family Houses Sold From the early 1990s up through 2005, the number of new single family houses sold rose steadily. In 2006, the number dropped dramatically and this dramatic decline continued through 2011. By 2014, the number of new houses sold had begun to climb back up, but the levels are still lower than those of 1990. (Source: U.S. Census Bureau)

The housing bubble began to show signs of bursting in 2005, as delinquency and late payments began to grow and an oversupply of new homes on the market became apparent. Dropping home values contributed to a decrease in the overall wealth of the household sector and caused homeowners to pull back on spending. Several mortgage lenders were forced to file for bankruptcy because homeowners were not making their payments, and by 2008 the problem had spread throughout the financial markets. Lenders clamped down on credit and the housing bubble burst. Financial markets were now in crisis and unable or unwilling to even extend credit to credit-worthy customers.

The housing bubble and the crisis in the financial markets were major contributors to the Great Recession that led to unemployment rates over 10% and falling GDP. While the United States is still recovering from the impact of the Great Recession, it has made substantial progress in restoring financial market stability through implementing aggressive fiscal and monetary policy.

The economic history of the United States is cyclical in nature with recessions and expansions. Some of these fluctuations are severe, such as the economic downturn that occurred during the Great Depression in the 1930s which lasted several years. Why does the economy grow at different rates in different years? What are the causes of the cyclical behavior of the economy? This chapter will introduce an important model, the aggregate demand–aggregate supply model, to begin our understanding of why economies expand and contract over time.

Introduction to the Aggregate Supply–Aggregate Demand Model

In this chapter, you will learn about:

- Macroeconomic Perspectives on Demand and Supply
- Building a Model of Aggregate Supply and Aggregate Demand
- Shifts in Aggregate Supply
- Shifts in Aggregate Demand
- How the AS–AD Model Incorporates Growth, Unemployment, and Inflation
- Keynes' Law and Say's Law in the AS–AD Model

A key part of macroeconomics is the use of models to analyze macro issues and problems. How is the rate of economic growth connected to changes in the unemployment rate? Is there a reason why unemployment and inflation seem to move in opposite directions: lower unemployment and higher inflation from 1997 to 2000, higher unemployment and lower inflation in the early 2000s, lower unemployment and higher inflation in the mid-2000s,

and then higher unemployment and lower inflation in 2009? Why did the current account deficit rise so high, but then decline in 2009?

To analyze questions like these, we must move beyond discussing macroeconomic issues one at a time, and begin building economic models that will capture the relationships and interconnections between them. The next three chapters take up this task. This chapter introduces the macroeconomic model of aggregate supply and aggregate demand, how the two interact to reach a macroeconomic equilibrium, and how shifts in aggregate demand or aggregate supply will affect that equilibrium. This chapter also relates the model of aggregate supply and aggregate demand to the three goals of economic policy (growth, unemployment, and inflation), and provides a framework for thinking about many of the connections and tradeoffs between these goals. The chapter on **The Keynesian Perspective** focuses on the macroeconomy in the short run, where aggregate demand plays a crucial role. The chapter on **The Neoclassical Perspective** explores the macroeconomy in the long run, where aggregate supply plays a crucial role.

11.1 | Macroeconomic Perspectives on Demand and Supply

By the end of this section, you will be able to:

- Explain Say's Law and understand why it primarily applies in the long run
- Explain Keynes' Law and understand why it primarily applies in the short run

Macroeconomists over the last two centuries have often divided into two groups: those who argue that supply is the most important determinant of the size of the macroeconomy while demand just tags along, and those who argue that demand is the most important factor in the size of the macroeconomy while supply just tags along.

Say's Law and the Macroeconomics of Supply

Those economists who emphasize the role of supply in the macroeconomy often refer to the work of a famous early nineteenth century French economist named Jean-Baptiste Say (1767–1832). **Say's law** is: "Supply creates its own demand." As a matter of historical accuracy, it seems clear that Say never actually wrote down this law and that it oversimplifies his beliefs, but the law lives on as useful shorthand for summarizing a point of view.

The intuition behind Say's law is that each time a good or service is produced and sold, it generates income that is earned for someone: a worker, a manager, an owner, or those who are workers, managers, and owners at firms that supply inputs along the chain of production. We alluded to this earlier in our discussion of the National Income approach to measuring GDP. The forces of supply and demand in individual markets will cause prices to rise and fall. The bottom line remains, however, that every sale represents income to someone, and so, Say's law argues, a given value of supply must create an equivalent value of demand somewhere else in the economy. Because Jean-Baptiste Say, Adam Smith, and other economists writing around the turn of the nineteenth century who discussed this view were known as "classical" economists, modern economists who generally subscribe to the Say's law view on the importance of supply for determining the size of the macroeconomy are called **neoclassical economists**.

If supply always creates exactly enough demand at the macroeconomic level, then (as Say himself recognized) it is hard to understand why periods of recession and high unemployment should ever occur. To be sure, even if total supply always creates an equal amount of total demand, the economy could still experience a situation of some firms earning profits while other firms suffer losses. Nevertheless, a recession is not a situation where all business failures are exactly counterbalanced by an offsetting number of successes. A recession is a situation in which the economy as a whole is shrinking in size, business failures outnumber the remaining success stories, and many firms end up suffering losses and laying off workers.

Say's law that supply creates its own demand does seem a good approximation for the long run. Over periods of some years or decades, as the productive power of an economy to supply goods and services increases, total demand in the economy grows at roughly the same pace. However, over shorter time horizons of a few months or even years, recessions or even depressions occur in which firms, as a group, seem to face a lack of demand for their products.

Keynes' Law and the Macroeconomics of Demand

The alternative to Say's law, with its emphasis on supply, is **Keynes' law**: "Demand creates its own supply." As a

matter of historical accuracy, just as Jean-Baptiste Say never wrote down anything as simpleminded as Say's law, John Maynard Keynes never wrote down Keynes' law, but the law is a useful simplification that conveys a certain point of view.

When Keynes wrote his influential work *The General Theory of Employment, Interest, and Money* during the 1930s Great Depression, he pointed out that during the Depression, the economy's capacity to supply goods and services had not changed much. U.S. unemployment rates soared higher than 20% from 1933 to 1935, but the number of possible workers had not increased or decreased much. Factories closed, but machinery and equipment had not disappeared. Technologies that had been invented in the 1920s were not un-invented and forgotten in the 1930s. Thus, Keynes argued that the Great Depression—and many ordinary recessions as well—were not caused by a drop in the ability of the economy to supply goods as measured by labor, physical capital, or technology. He argued the economy often produced less than its full potential, not because it was technically impossible to produce more with the existing workers and machines, but because a lack of demand in the economy as a whole led to inadequate incentives for firms to produce. In such cases, he argued, the level of GDP in the economy was not primarily determined by the potential of what the economy could supply, but rather by the amount of total demand.

Keynes' law seems to apply fairly well in the short run of a few months to a few years, when many firms experience either a drop in demand for their output during a recession or so much demand that they have trouble producing enough during an economic boom. However, demand cannot tell the whole macroeconomic story, either. After all, if demand was all that mattered at the macroeconomic level, then the government could make the economy as large as it wanted just by pumping up total demand through a large increase in the government spending component or by legislating large tax cuts to push up the consumption component. Economies do, however, face genuine limits to how much they can produce, limits determined by the quantity of labor, physical capital, technology, and the institutional and market structures that bring these factors of production together. These constraints on what an economy can supply at the macroeconomic level do not disappear just because of an increase in demand.

Combining Supply and Demand in Macroeconomics

Two insights emerge from this overview of Say's law with its emphasis on macroeconomic supply and Keynes' law with its emphasis on macroeconomic demand. The first conclusion, which is not exactly a hot news flash, is that an economic approach focused only on the supply side or only on the demand side can be only a partial success. We need to take into account both supply and demand. The second conclusion is that since Keynes' law applies more accurately in the short run and Say's law applies more accurately in the long run, the tradeoffs and connections between the three goals of macroeconomics may be different in the short run and the long run.

11.2 | Building a Model of Aggregate Demand and Aggregate Supply

By the end of this section, you will be able to:

- Explain the aggregate supply curve and how it relates to real GDP and potential GDP
- Explain the aggregate demand curve and how it is influenced by price levels
- Interpret the aggregate demand/aggregate supply model
- Identify the point of equilibrium in the aggregate demand/aggregate supply model
- Define short run aggregate supply and long run aggregate supply

To build a useful macroeconomic model, we need a model that shows what determines total supply or total demand for the economy, and how total demand and total supply interact at the macroeconomic level. We call this the **aggregate demand/aggregate supply model**. This module will explain aggregate supply, aggregate demand, and the equilibrium between them. The following modules will discuss the causes of shifts in aggregate supply and aggregate demand.

The Aggregate Supply Curve and Potential GDP

Firms make decisions about what quantity to supply based on the profits they expect to earn. They determine profits, in turn, by the price of the outputs they sell and by the prices of the inputs, like labor or raw materials, that they need to buy. **Aggregate supply (AS)** refers to the total quantity of output (i.e. real GDP) firms will produce and sell. The

aggregate supply (AS) curve shows the total quantity of output (i.e. real GDP) that firms will produce and sell at each price level.

Figure 11.3 shows an aggregate supply curve. In the following paragraphs, we will walk through the elements of the diagram one at a time: the horizontal and vertical axes, the aggregate supply curve itself, and the meaning of the potential GDP vertical line.

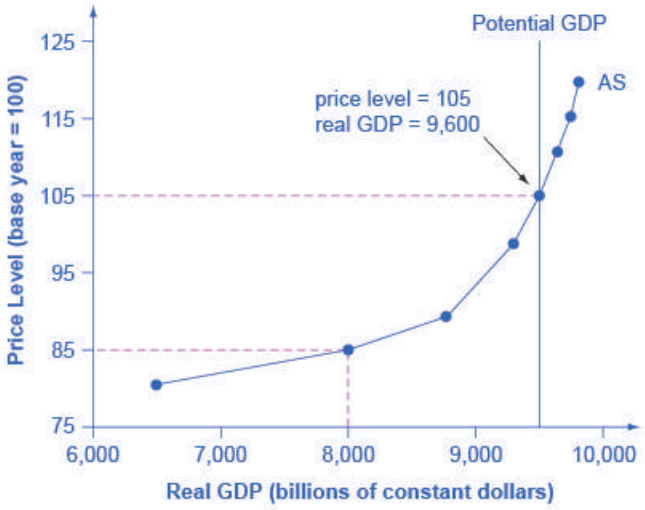


Figure 11.3 The Aggregate Supply Curve Aggregate supply (AS) slopes up, because as the price level for outputs rises, with the price of inputs remaining fixed, firms have an incentive to produce more to earn higher profits. The potential GDP line shows the maximum that the economy can produce with full employment of workers and physical capital.

The diagram's horizontal axis shows real GDP—that is, the level of GDP adjusted for inflation. The vertical axis shows the price level, which measures the average price of all goods and services produced in the economy. In other words, the price level in the AD-AS model is what we called the GDP Deflator in **The Macroeconomic Perspective**. Remember that the price level is different from the inflation rate. Visualize the price level as an index number, like the Consumer Price Index, while the inflation rate is the percentage change in the price level over time.

As the price level rises, real GDP rises as well. Why? The price level on the vertical axis represents prices for final goods or outputs bought in the economy—i.e. the GDP deflator—not the price level for intermediate goods and services that are inputs to production. Thus, the AS curve describes how suppliers will react to a higher price level for final outputs of goods and services, while holding the prices of inputs like labor and energy constant. If firms across the economy face a situation where the price level of what they produce and sell is rising, but their costs of production are not rising, then the lure of higher profits will induce them to expand production. In other words, an aggregate supply curve shows how producers as a group will respond to an increase in aggregate demand.

An AS curve's slope changes from nearly flat at its far left to nearly vertical at its far right. At the far left of the aggregate supply curve, the level of output in the economy is far below **potential GDP**, which we define as the amount of real GDP an economy can produce by fully employing its existing levels of labor, physical capital, and technology, in the context of its existing market and legal institutions. At these relatively low levels of output, levels of unemployment are high, and many factories are running only part-time, or have closed their doors. In this situation, a relatively small increase in the prices of the outputs that businesses sell—while assuming no rise in input prices—can encourage a considerable surge in the quantity of aggregate supply because so many workers and factories are ready to swing into production.

As the GDP increases, however, some firms and industries will start running into limits: perhaps nearly all of the expert workers in a certain industry will have jobs or factories in certain geographic areas or industries will be running at full speed. In the AS curve's intermediate area, a higher price level for outputs continues to encourage a greater quantity of output—but as the increasingly steep upward slope of the aggregate supply curve shows, the increase in real GDP in response to a given rise in the price level will not be as large. (Read the following Clear It Up feature to learn why the AS curve crosses potential GDP.)

Clear It Up

Why does AS cross potential GDP?

Economists typically draw the aggregate supply curve to cross the potential GDP line. This shape may seem puzzling: How can an economy produce at an output level which is higher than its "potential" or "full employment" GDP? The economic intuition here is that if prices for outputs were high enough, producers would make fanatical efforts to produce: all workers would be on double-overtime, all machines would run 24 hours a day, seven days a week. Such hyper-intense production would go beyond using potential labor and physical capital resources fully, to using them in a way that is not sustainable in the long term. Thus, it is possible for production to sprint above potential GDP, but only in the short run.

At the far right, the aggregate supply curve becomes nearly vertical. At this quantity, higher prices for outputs cannot encourage additional output, because even if firms want to expand output, the inputs of labor and machinery in the economy are fully employed. In this example, the vertical line in the exhibit shows that potential GDP occurs at a total output of 9,500. When an economy is operating at its potential GDP, machines and factories are running at capacity, and the unemployment rate is relatively low—at the natural rate of unemployment. For this reason, potential GDP is sometimes also called **full-employment GDP**.

The Aggregate Demand Curve

Aggregate demand (AD) refers to the amount of total spending on domestic goods and services in an economy. (Strictly speaking, AD is what economists call total planned expenditure. We will further explain this distinction in the appendix **The Expenditure-Output Model** . For now, just think of aggregate demand as total spending.) It includes all four components of demand: consumption, investment, government spending, and net exports (exports minus imports). This demand is determined by a number of factors, but one of them is the price level—recall though, that the price level is an index number such as the GDP deflator that measures the average price of the things we buy. The **aggregate demand (AD) curve** shows the total spending on domestic goods and services at each price level.

Figure 11.4 presents an aggregate demand (AD) curve. Just like the aggregate supply curve, the horizontal axis shows real GDP and the vertical axis shows the price level. The AD curve slopes down, which means that increases in the price level of outputs lead to a lower quantity of total spending. The reasons behind this shape are related to how changes in the price level affect the different components of aggregate demand. The following components comprise aggregate demand: consumption spending (C), investment spending (I), government spending (G), and spending on exports (X) minus imports (M): C + I + G + X – M.

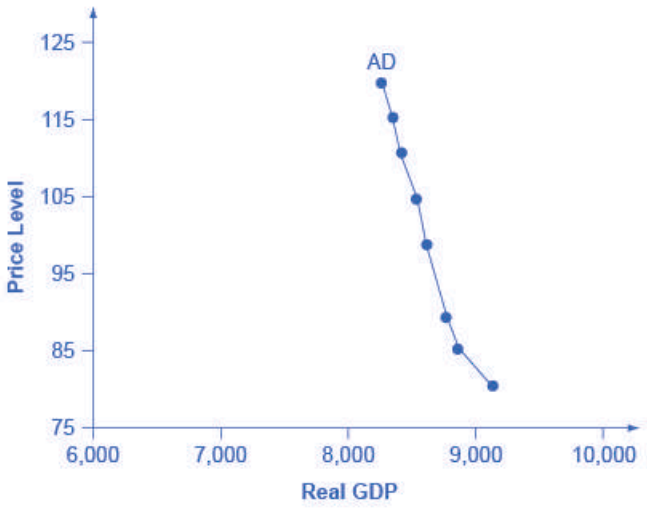


Figure 11.4 The Aggregate Demand Curve Aggregate demand (AD) slopes down, showing that, as the price level rises, the amount of total spending on domestic goods and services declines.

The wealth effect holds that as the price level increases, the buying power of savings that people have stored up in bank accounts and other assets will diminish, eaten away to some extent by inflation. Because a rise in the price level reduces people's wealth, consumption spending will fall as the price level rises.

The interest rate effect is that as prices for outputs rise, the same purchases will take more money or credit to accomplish. This additional demand for money and credit will push interest rates higher. In turn, higher interest rates will reduce borrowing by businesses for investment purposes and reduce borrowing by households for homes and cars—thus reducing consumption and investment spending.

The foreign price effect points out that if prices rise in the United States while remaining fixed in other countries, then goods in the United States will be relatively more expensive compared to goods in the rest of the world. U.S. exports will be relatively more expensive, and the quantity of exports sold will fall. U.S. imports from abroad will be relatively cheaper, so the quantity of imports will rise. Thus, a higher domestic price level, relative to price levels in other countries, will reduce net export expenditures.

Among economists all three of these effects are controversial, in part because they do not seem to be very large. For this reason, the aggregate demand curve in **Figure 11.4** slopes downward fairly steeply. The steep slope indicates that a higher price level for final outputs reduces aggregate demand for all three of these reasons, but that the change in the quantity of aggregate demand as a result of changes in price level is not very large.

Read the following Work It Out feature to learn how to interpret the AD/AS model. In this example, aggregate supply, aggregate demand, and the price level are given for the imaginary country of Xurbia.

Work It Out

Interpreting the AD/AS Model

Table 11.1 shows information on aggregate supply, aggregate demand, and the price level for the imaginary country of Xurbia. What information does **Table 11.1** tell you about the state of the Xurbia's economy? Where is the equilibrium price level and output level (this is the SR macroequilibrium)? Is Xurbia risking inflationary pressures or facing high unemployment? How can you tell?

Price Level	Aggregate Demand	Aggregate Supply
110	$700	$600
120	$690	$640
130	$680	$680
140	$670	$720
150	$660	$740
160	$650	$760
170	$640	$770

Table 11.1 Price Level: Aggregate Demand/Aggregate Supply

To begin to use the AD/AS model, it is important to plot the AS and AD curves from the data provided. What is the equilibrium?

Step 1. Draw your x- and y-axis. Label the x-axis Real GDP and the y-axis Price Level.

Step 2. Plot AD on your graph.

Step 3. Plot AS on your graph.

Step 4. Look at **Figure 11.5** which provides a visual to aid in your analysis.

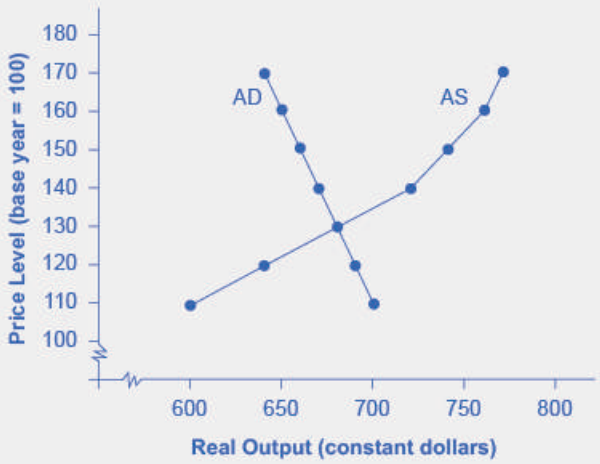


Figure 11.5 The AD/AS Curves AD and AS curves created from the data in **Table 11.1**.

Step 5. Determine where AD and AS intersect. This is the equilibrium with price level at 130 and real GDP at $680.

Step 6. Look at the graph to determine where equilibrium is located. We can see that this equilibrium is fairly far from where the AS curve becomes near-vertical (or at least quite steep) which seems to start at about $750 of real output. This implies that the economy is not close to potential GDP. Thus, unemployment will be high. In the relatively flat part of the AS curve, where the equilibrium occurs, changes in the price level will not be a major concern, since such changes are likely to be small.

Step 7. Determine what the steep portion of the AS curve indicates. Where the AS curve is steep, the economy is at or close to potential GDP.

Step 8. Draw conclusions from the given information:

- If equilibrium occurs in the flat range of AS, then economy is not close to potential GDP and will be experiencing unemployment, but stable price level.
- If equilibrium occurs in the steep range of AS, then the economy is close or at potential GDP and will be experiencing rising price levels or inflationary pressures, but will have a low unemployment rate.

Equilibrium in the Aggregate Demand/Aggregate Supply Model

The intersection of the aggregate supply and aggregate demand curves shows the equilibrium level of real GDP and the equilibrium price level in the economy. At a relatively low price level for output, firms have little incentive to produce, although consumers would be willing to purchase a large quantity of output. As the price level rises, aggregate supply rises and aggregate demand falls until the equilibrium point is reached.

Figure 11.6 combines the AS curve from **Figure 11.3** and the AD curve from **Figure 11.4** and places them both on a single diagram. In this example, the equilibrium point occurs at point E, at a price level of 90 and an output level of 8,800.

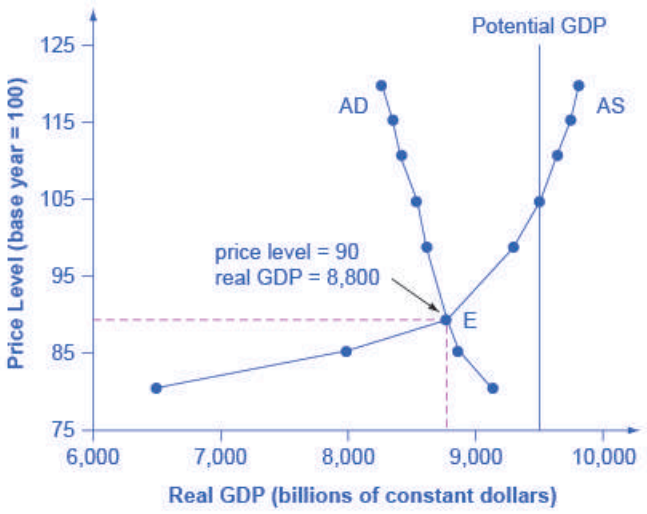


Figure 11.6 Aggregate Supply and Aggregate Demand The equilibrium, where aggregate supply (AS) equals aggregate demand (AD), occurs at a price level of 90 and an output level of 8,800.

Confusion sometimes arises between the aggregate supply and aggregate demand model and the microeconomic analysis of demand and supply in particular markets for goods, services, labor, and capital. Read the following Clear It Up feature to gain an understanding of whether AS and AD are macro or micro.

Clear It Up

Are AS and AD macro or micro?

These aggregate supply and demand models and the microeconomic analysis of demand and supply in particular markets for goods, services, labor, and capital have a superficial resemblance, but they also have many underlying differences.

For example, the vertical and horizontal axes have distinctly different meanings in macroeconomic and microeconomic diagrams. The vertical axis of a microeconomic demand and supply diagram expresses a price (or wage or rate of return) for an individual good or service. This price is implicitly relative: it is intended to be compared with the prices of other products (for example, the price of pizza relative to the price of fried chicken). In contrast, the vertical axis of an aggregate supply and aggregate demand diagram expresses the level of a price index like the Consumer Price Index or the GDP deflator—combining a wide array of prices from across the economy. The price level is absolute: it is not intended to be compared to any other prices since it is essentially the average price of all products in an economy. The horizontal axis of a microeconomic supply and demand curve measures the quantity of a particular good or service. In contrast, the horizontal axis of the aggregate demand and aggregate supply diagram measures GDP, which is the sum of all the final goods and services produced in the economy, not the quantity in a specific market.

In addition, the economic reasons for the shapes of the curves in the macroeconomic model are different from the reasons behind the shapes of the curves in microeconomic models. Demand curves for individual goods or services slope down primarily because of the existence of substitute goods, not the wealth effects, interest rate, and foreign price effects associated with aggregate demand curves. The slopes of individual supply and demand curves can have a variety of different slopes, depending on the extent to which quantity demanded and quantity supplied react to price in that specific market, but the slopes of the AS and AD curves are much the same in every diagram (although as we shall see in later chapters, short-run and long-run perspectives will emphasize different parts of the AS curve).

In short, just because the AD/AS diagram has two lines that cross, do not assume that it is the same as every other diagram where two lines cross. The intuitions and meanings of the macro and micro diagrams are only

distant cousins from different branches of the economics family tree.

Defining SRAS and LRAS

In the Clear It Up feature titled "Why does AS cross potential GDP?" we differentiated between short run changes in aggregate supply which the AS curve shows and long run changes in aggregate supply which the vertical line at potential GDP defines. In the short run, if demand is too low (or too high), it is possible for producers to supply less GDP (or more GDP) than potential. In the long run, however, producers are limited to producing at potential GDP. For this reason, we may also refer to what we have been calling the AS curve as the **short run aggregate supply (SRAS) curve**. We may also refer to the vertical line at potential GDP as the **long run aggregate supply (LRAS) curve**.

11.3 | Shifts in Aggregate Supply

By the end of this section, you will be able to:

- Explain how productivity growth changes the aggregate supply curve
- Explain how changes in input prices change the aggregate supply curve

The original equilibrium in the AD/AS diagram will shift to a new equilibrium if the AS or AD curve shifts. When the aggregate supply curve shifts to the right, then at every price level, producers supply a greater quantity of real GDP. When the AS curve shifts to the left, then at every price level, producers supply a lower quantity of real GDP. This module discusses two of the most important factors that can lead to shifts in the AS curve: productivity growth and changes in input prices.

How Productivity Growth Shifts the AS Curve

In the long run, the most important factor shifting the AS curve is productivity growth. Productivity means how much output can be produced with a given quantity of labor. One measure of this is output per worker or GDP per capita. Over time, productivity grows so that the same quantity of labor can produce more output. Historically, the real growth in GDP per capita in an advanced economy like the United States has averaged about 2% to 3% per year, but productivity growth has been faster during certain extended periods like the 1960s and the late 1990s through the early 2000s, or slower during periods like the 1970s. A higher level of productivity shifts the AS curve to the right, because with improved productivity, firms can produce a greater quantity of output at every price level. **Figure 11.7** (a) shows an outward shift in productivity over two time periods. The AS curve shifts out from $SRAS_0$ to $SRAS_1$ to $SRAS_2$, and the equilibrium shifts from E_0 to E_1 to E_2. Note that with increased productivity, workers can produce more GDP. Thus, full employment corresponds to a higher level of potential GDP, which we show as a rightward shift in LRAS from $LRAS_0$ to $LRAS_1$ to $LRAS_2$.

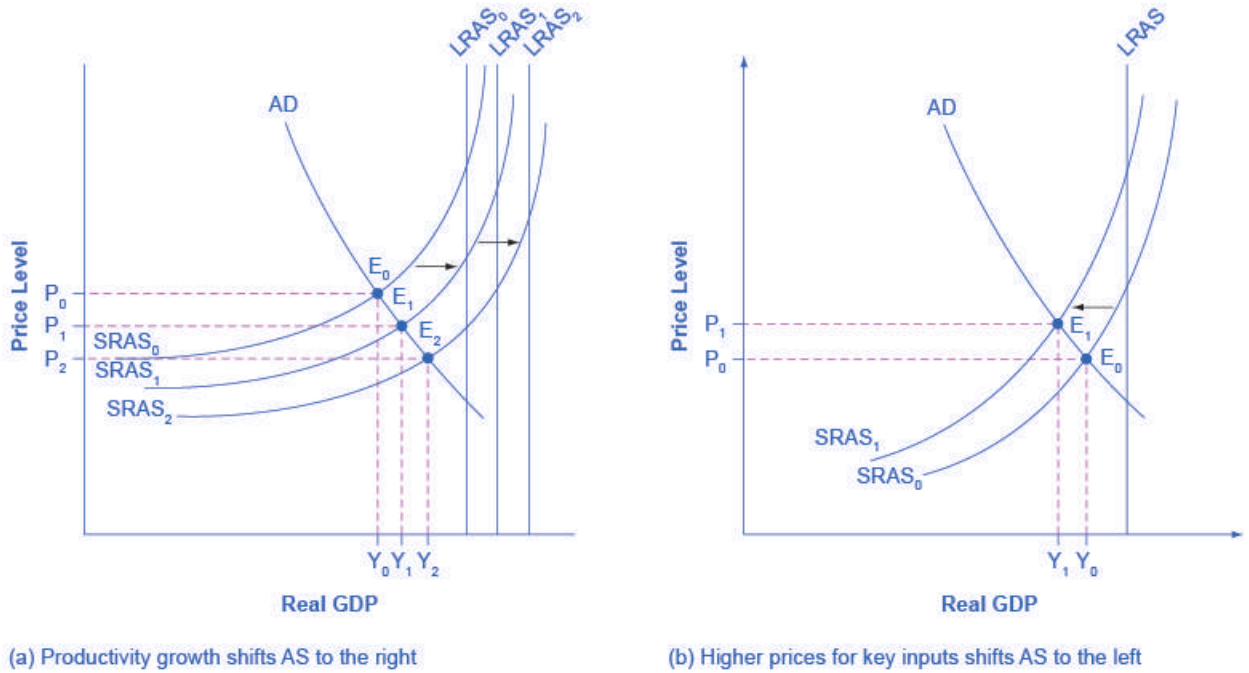


Figure 11.7 Shifts in Aggregate Supply (a) The rise in productivity causes the SRAS curve to shift to the right. The original equilibrium E_0 is at the intersection of AD and $SRAS_0$. When SRAS shifts right, then the new equilibrium E_1 is at the intersection of AD and $SRAS_1$, and then yet another equilibrium, E_2, is at the intersection of AD and $SRAS_2$. Shifts in SRAS to the right, lead to a greater level of output and to downward pressure on the price level. (b) A higher price for inputs means that at any given price level for outputs, a lower real GDP will be produced so aggregate supply will shift to the left from $SRAS_0$ to $SRAS_1$. The new equilibrium, E_1, has a reduced quantity of output and a higher price level than the original equilibrium (E_0).

A shift in the SRAS curve to the right will result in a greater real GDP and downward pressure on the price level, if aggregate demand remains unchanged. However, if this shift in SRAS results from gains in productivity growth, which we typically measure in terms of a few percentage points per year, the effect will be relatively small over a few months or even a couple of years. Recall how in **Choice in a World of Scarcity**, we said that a nation's production possibilities frontier is fixed in the short run, but shifts out in the long run? This is the same phenomenon using a different model.

How Changes in Input Prices Shift the AS Curve

Higher prices for inputs that are widely used across the entire economy can have a macroeconomic impact on aggregate supply. Examples of such widely used inputs include labor and energy products. Increases in the price of such inputs will cause the SRAS curve to shift to the left, which means that at each given price level for outputs, a higher price for inputs will discourage production because it will reduce the possibilities for earning profits. **Figure 11.7** (b) shows the aggregate supply curve shifting to the left, from $SRAS_0$ to $SRAS_1$, causing the equilibrium to move from E_0 to E_1. The movement from the original equilibrium of E_0 to the new equilibrium of E_1 will bring a nasty set of effects: reduced GDP or recession, higher unemployment because the economy is now further away from potential GDP, and an inflationary higher price level as well. For example, the U.S. economy experienced recessions in 1974–1975, 1980–1982, 1990–91, 2001, and 2007–2009 that were each preceded or accompanied by a rise in the key input of oil prices. In the 1970s, this pattern of a shift to the left in SRAS leading to a stagnant economy with high unemployment and inflation was nicknamed **stagflation**.

Conversely, a decline in the price of a key input like oil will shift the SRAS curve to the right, providing an incentive for more to be produced at every given price level for outputs. From 1985 to 1986, for example, the average price of crude oil fell by almost half, from $24 a barrel to $12 a barrel. Similarly, from 1997 to 1998, the price of a barrel of crude oil dropped from $17 per barrel to $11 per barrel. In both cases, the plummeting oil price led to a situation like that which we presented earlier in **Figure 11.7** (a), where the outward shift of SRAS to the right allowed the economy to expand, unemployment to fall, and inflation to decline.

Along with energy prices, two other key inputs that may shift the SRAS curve are the cost of labor, or wages, and the cost of imported goods that we use as inputs for other products. In these cases as well, the lesson is that lower prices for inputs cause SRAS to shift to the right, while higher prices cause it to shift back to the left. Note that, unlike changes in productivity, changes in input prices do not generally cause LRAS to shift, only SRAS.

Other Supply Shocks

The aggregate supply curve can also shift due to shocks to input goods or labor. For example, an unexpected early freeze could destroy a large number of agricultural crops, a shock that would shift the AS curve to the left since there would be fewer agricultural products available at any given price.

Similarly, shocks to the labor market can affect aggregate supply. An extreme example might be an overseas war that required a large number of workers to cease their ordinary production in order to go fight for their country. In this case, SRAS and LRAS would both shift to the left because there would be fewer workers available to produce goods at any given price.

11.4 | Shifts in Aggregate Demand

By the end of this section, you will be able to:

- Explain how imports influence aggregate demand
- Identify ways in which business confidence and consumer confidence can affect aggregate demand
- Explain how government policy can change aggregate demand
- Evaluate why economists disagree on the topic of tax cuts

As we mentioned previously, the components of aggregate demand are consumption spending (C), investment spending (I), government spending (G), and spending on exports (X) minus imports (M). (Read the following Clear It Up feature for explanation of why imports are subtracted from exports and what this means for aggregate demand.) A shift of the AD curve to the right means that at least one of these components increased so that a greater amount of total spending would occur at every price level. A shift of the AD curve to the left means that at least one of these components decreased so that a lesser amount of total spending would occur at every price level. **The Keynesian Perspective** will discuss the components of aggregate demand and the factors that affect them. Here, the discussion will sketch two broad categories that could cause AD curves to shift: changes in consumer or firm behavior and changes in government tax or spending policy.

Do imports diminish aggregate demand?

We have seen that the formula for aggregate demand is AD = C + I + G + X - M, where M is the total value of imported goods. Why is there a minus sign in front of imports? Does this mean that more imports will result in a lower level of aggregate demand? The short answer is yes, because aggregate demand is defined as total demand for domestically produced goods and services.

When an American buys a foreign product, for example, it gets counted along with all the other consumption. Thus, the income generated does not go to American producers, but rather to producers in another country. It would be wrong to count this as part of domestic demand. Therefore, imports added in consumption are subtracted back out in the M term of the equation.

Because of the way in which we write the demand equation, it is easy to make the mistake of thinking that imports are bad for the economy. Just keep in mind that every negative number in the M term has a corresponding positive number in the C or I or G term, and they always cancel out.

How Changes by Consumers and Firms Can Affect AD

When consumers feel more confident about the future of the economy, they tend to consume more. If business confidence is high, then firms tend to spend more on investment, believing that the future payoff from that investment will be substantial. Conversely, if consumer or business confidence drops, then consumption and investment spending decline.

The University of Michigan publishes a survey of consumer confidence and constructs an index of consumer confidence each month. The survey results are then reported at **http://www.sca.isr.umich.edu (http://www.sca.isr.umich.edu/)** , which break down the change in consumer confidence among different income levels. According to that index, consumer confidence averaged around 90 prior to the Great Recession, and then it fell to below 60 in late 2008, which was the lowest it had been since 1980. Since then, confidence has climbed from a 2011 low of 55.8 back to a level in the low 80s, which economists consider close to a healthy state.

The Organization for Economic Development and Cooperation (OECD) publishes one measure of business confidence: the "business tendency surveys". The OECD collects business opinion survey data for 21 countries on future selling prices and employment, among other business climate elements. After sharply declining during the Great Recession, the measure has risen above zero again and is back to long-term averages (the indicator dips below zero when business outlook is weaker than usual). Of course, either of these survey measures is not very precise. They can however, suggest when confidence is rising or falling, as well as when it is relatively high or low compared to the past.

Because economists associate a rise in confidence with higher consumption and investment demand, it will lead to an outward shift in the AD curve, and a move of the equilibrium, from E_0 to E_1, to a higher quantity of output and a higher price level, as **Figure 11.8** (a) shows.

Consumer and business confidence often reflect macroeconomic realities; for example, confidence is usually high when the economy is growing briskly and low during a recession. However, economic confidence can sometimes rise or fall for reasons that do not have a close connection to the immediate economy, like a risk of war, election results, foreign policy events, or a pessimistic prediction about the future by a prominent public figure. U.S. presidents, for example, must be careful in their public pronouncements about the economy. If they offer economic pessimism, they risk provoking a decline in confidence that reduces consumption and investment and shifts AD to the left, and in a self-fulfilling prophecy, contributes to causing the recession that the president warned against in the first place. **Figure 11.8** (b) shows a shift of AD to the left, and the corresponding movement of the equilibrium, from E_0 to E_1, to a lower quantity of output and a lower price level.

Link It Up

Visit this website (http://openstaxcollege.org/l/consumerconfid) for data on consumer confidence.

Link It Up

Visit this website (http://openstaxcollege.org/l/businessconfid) for data on business confidence.

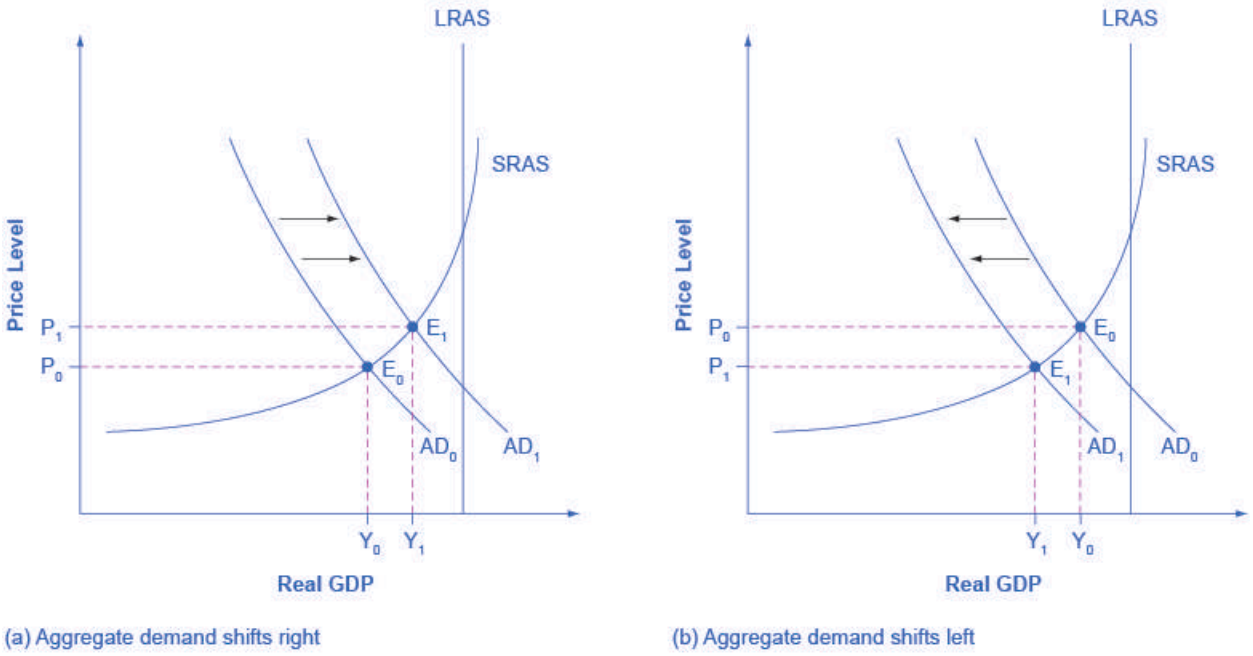


Figure 11.8 Shifts in Aggregate Demand (a) An increase in consumer confidence or business confidence can shift AD to the right, from AD_0 to AD_1. When AD shifts to the right, the new equilibrium (E_1) will have a higher quantity of output and also a higher price level compared with the original equilibrium (E_0). In this example, the new equilibrium (E_1) is also closer to potential GDP. An increase in government spending or a cut in taxes that leads to a rise in consumer spending can also shift AD to the right. (b) A decrease in consumer confidence or business confidence can shift AD to the left, from AD_0 to AD_1. When AD shifts to the left, the new equilibrium (E_1) will have a lower quantity of output and also a lower price level compared with the original equilibrium (E_0). In this example, the new equilibrium (E_1) is also farther below potential GDP. A decrease in government spending or higher taxes that leads to a fall in consumer spending can also shift AD to the left.

How Government Macroeconomic Policy Choices Can Shift AD

Government spending is one component of AD. Thus, higher government spending will cause AD to shift to the right, as in **Figure 11.8** (a), while lower government spending will cause AD to shift to the left, as in **Figure 11.8** (b). For example, in the United States, government spending declined by 3.2% of GDP during the 1990s, from 21% of GDP in 1991, and to 17.8% of GDP in 1998. However, from 2005 to 2009, the peak of the Great Recession, government spending increased from 19% of GDP to 21.4% of GDP. If changes of a few percentage points of GDP seem small to you, remember that since GDP was about $14.4 trillion in 2009, a seemingly small change of 2% of GDP is equal to close to $300 billion.

Tax policy can affect consumption and investment spending, too. Tax cuts for individuals will tend to increase consumption demand, while tax increases will tend to diminish it. Tax policy can also pump up investment demand by offering lower tax rates for corporations or tax reductions that benefit specific kinds of investment. Shifting C or I will shift the AD curve as a whole.

During a recession, when unemployment is high and many businesses are suffering low profits or even losses, the U.S. Congress often passes tax cuts. During the 2001 recession, for example, the U.S. Congress enacted a tax cut into law. At such times, the political rhetoric often focuses on how people experiencing hard times need relief from taxes. The aggregate supply and aggregate demand framework, however, offers a complementary rationale, as **Figure 11.9** illustrates. The original equilibrium during a recession is at point E_0, relatively far from the full employment level of output. The tax cut, by increasing consumption, shifts the AD curve to the right. At the new equilibrium (E_1), real GDP rises and unemployment falls and, because in this diagram the economy has not yet reached its potential or full employment level of GDP, any rise in the price level remains muted. Read the following Clear It Up feature to consider the question of whether economists favor tax cuts or oppose them.

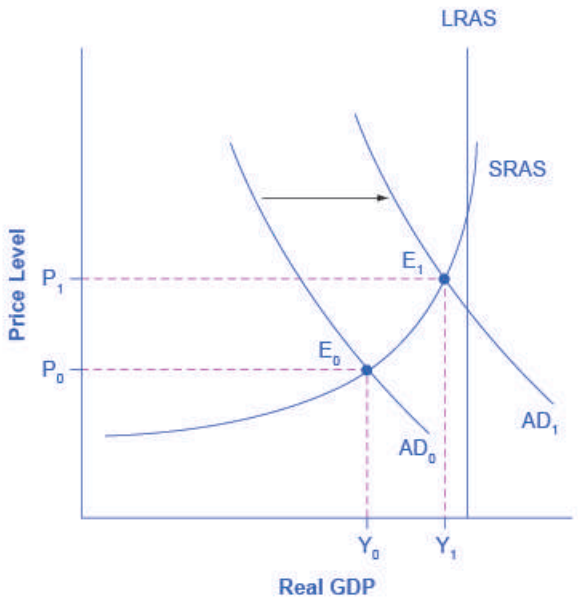


Figure 11.9 Recession and Full Employment in the AD/AS Model Whether the economy is in a recession is illustrated in the AD/AS model by how close the equilibrium is to the potential GDP line as indicated by the vertical LRAS line. In this example, the level of output Y_0 at the equilibrium E_0 is relatively far from the potential GDP line, so it can represent an economy in recession, well below the full employment level of GDP. In contrast, the level of output Y_1 at the equilibrium E_1 is relatively close to potential GDP, and so it would represent an economy with a lower unemployment rate.

Clear It Up

Do economists favor tax cuts or oppose them?

One of the most fundamental divisions in American politics over the last few decades has been between those who believe that the government should cut taxes substantially and those who disagree. Ronald Reagan rode into the presidency in 1980 partly because of his promise, soon carried out, to enact a substantial tax cut. George Bush lost his bid for reelection against Bill Clinton in 1992 partly because he had broken his 1988 promise: “Read my lips! No new taxes!” In the 2000 presidential election, both George W. Bush and Al Gore advocated substantial tax cuts and Bush succeeded in pushing a tax cut package through Congress early in 2001. More recently in 2017, Donald Trump has pushed for tax cuts to stimulate the economy. Disputes over tax cuts often ignite at the state and local level as well.

What side do economists take? Do they support broad tax cuts or oppose them? The answer, unsatisfying to zealots on both sides, is that it depends. One issue is whether equally large government spending cuts accompany the tax cuts. Economists differ, as does any broad cross-section of the public, on how large government spending should be and what programs the government might cut back. A second issue, more

relevant to the discussion in this chapter, concerns how close the economy is to the full employment output level. In a recession, when the AD and AS curves intersect far below the full employment level, tax cuts can make sense as a way of shifting AD to the right. However, when the economy is already performing extremely well, tax cuts may shift AD so far to the right as to generate inflationary pressures, with little gain to GDP.

With the AD/AS framework in mind, many economists might readily believe that the 1981 Reagan tax cuts, which took effect just after two serious recessions, were beneficial economic policy. Similarly, Congress enacted the 2001 Bush tax cuts and the 2009 Obama tax cuts during recessions. However, some of the same economists who favor tax cuts during recession would be much more dubious about identical tax cuts at a time the economy is performing well and cyclical unemployment is low.

Government spending and tax rate changes can be useful tools to affect aggregate demand. We will discuss these in greater detail in the **Government Budgets and Fiscal Policy** chapter and **The Impacts of Government Borrowing**. Other policy tools can shift the aggregate demand curve as well. For example, as we will discus in the **Monetary Policy and Bank Regulation** chapter, the Federal Reserve can affect interest rates and credit availability. Higher interest rates tend to discourage borrowing and thus reduce both household spending on big-ticket items like houses and cars and investment spending by business. Conversely, lower interest rates will stimulate consumption and investment demand. Interest rates can also affect exchange rates, which in turn will have effects on the export and import components of aggregate demand.

Clarifying the details of these alternative policies and how they affect the components of aggregate demand can wait for **The Keynesian Perspective** chapter. Here, the key lesson is that a shift of the aggregate demand curve to the right leads to a greater real GDP and to upward pressure on the price level. Conversely, a shift of aggregate demand to the left leads to a lower real GDP and a lower price level. Whether these changes in output and price level are relatively large or relatively small, and how the change in equilibrium relates to potential GDP, depends on whether the shift in the AD curve is happening in the AS curve's relatively flat or relatively steep portion.

11.5 | How the AD/AS Model Incorporates Growth, Unemployment, and Inflation

By the end of this section, you will be able to:

- Use the aggregate demand/aggregate supply model to show periods of economic growth and recession
- Explain how unemployment and inflation impact the aggregate demand/aggregate supply model
- Evaluate the importance of the aggregate demand/aggregate supply model

The AD/AS model can convey a number of interlocking relationships between the three macroeconomic goals of growth, unemployment, and low inflation. Moreover, the AD/AS framework is flexible enough to accommodate both the Keynes' law approach that focuses on aggregate demand and the short run, while also including the Say's law approach that focuses on aggregate supply and the long run. These advantages are considerable. Every model is a simplified version of the deeper reality and, in the context of the AD/AS model, the three macroeconomic goals arise in ways that are sometimes indirect or incomplete. In this module, we consider how the AD/AS model illustrates the three macroeconomic goals of economic growth, low unemployment, and low inflation.

Growth and Recession in the AD/AS Diagram

In the AD/AS diagram, long-run economic growth due to productivity increases over time will be represented by a gradual shift to the right of aggregate supply. The vertical line representing potential GDP (or the "full employment level of GDP") will gradually shift to the right over time as well. Earlier **Figure 11.7** (a) showed a pattern of economic growth over three years, with the AS curve shifting slightly out to the right each year. However, the factors that determine the speed of this long-term economic growth rate—like investment in physical and human capital, technology, and whether an economy can take advantage of catch-up growth—do not appear directly in the AD/AS diagram.

In the short run, GDP falls and rises in every economy, as the economy dips into recession or expands out of recession.

The AD/AS diagram illustrates recessions when the equilibrium level of real GDP is substantially below potential GDP, as we see at the equilibrium point E_0 in **Figure 11.9**. From another standpoint, in years of resurgent economic growth the equilibrium will typically be close to potential GDP, as equilibrium point E_1 in that earlier figure shows.

Unemployment in the AD/AS Diagram

We described two types of unemployment in the **Unemployment** chapter. Short run variations in unemployment (cyclical unemployment) are caused by the business cycle as the economy expands and contracts. Over the long run, in the United States, the unemployment rate typically hovers around 5% (give or take one percentage point or so), when the economy is healthy. In many of the national economies across Europe, the unemployment rate in recent decades has only dropped to about 10% or a bit lower, even in good economic years. We call this baseline level of unemployment that occurs year-in and year-out the natural rate of unemployment and we determine it by how well the structures of market and government institutions in the economy lead to a matching of workers and employers in the labor market. Potential GDP can imply different unemployment rates in different economies, depending on the natural rate of unemployment for that economy.

The AD/AS diagram shows cyclical unemployment by how close the economy is to the potential or full GDP employment level. Returning to **Figure 11.9**, relatively low cyclical unemployment for an economy occurs when the level of output is close to potential GDP, as in the equilibrium point E_1. Conversely, high cyclical unemployment arises when the output is substantially to the left of potential GDP on the AD/AS diagram, as at the equilibrium point E_0. Although we do not show the factors that determine the natural rate of unemployment separately in the AD/AS model, they are implicitly part of what determines potential GDP or full employment GDP in a given economy.

Inflationary Pressures in the AD/AS Diagram

Inflation fluctuates in the short run. Higher inflation rates have typically occurred either during or just after economic booms: for example, the biggest spurts of inflation in the U.S. economy during the twentieth century followed the wartime booms of World War I and World War II. Conversely, rates of inflation generally decline during recessions. As an extreme example, inflation actually became negative—a situation called “deflation”—during the Great Depression. Even during the relatively short 1991-1992 recession, the inflation rate declined from 5.4% in 1990 to 3.0% in 1992. During the relatively short 2001 recession, the rate of inflation declined from 3.4% in 2000 to 1.6% in 2002. During the deep recession of 2007–2009, the inflation rate declined from 3.8% in 2008 to –0.4% in 2009. Some countries have experienced bouts of high inflation that lasted for years. In the U.S. economy since the mid–1980s, inflation does not seem to have had any long-term trend to be substantially higher. Instead, it has stayed in the 1–5% range annually.

The AD/AS framework implies two ways that inflationary pressures may arise. One possible trigger is if aggregate demand continues to shift to the right when the economy is already at or near potential GDP and full employment, thus pushing the macroeconomic equilibrium into the AS curve's steep portion. In **Figure 11.10** (a), there is a shift of aggregate demand to the right. The new equilibrium E_1 is clearly at a higher price level than the original equilibrium E_0. In this situation, the aggregate demand in the economy has soared so high that firms in the economy are not capable of producing additional goods, because labor and physical capital are fully employed, and so additional increases in aggregate demand can only result in a rise in the price level.

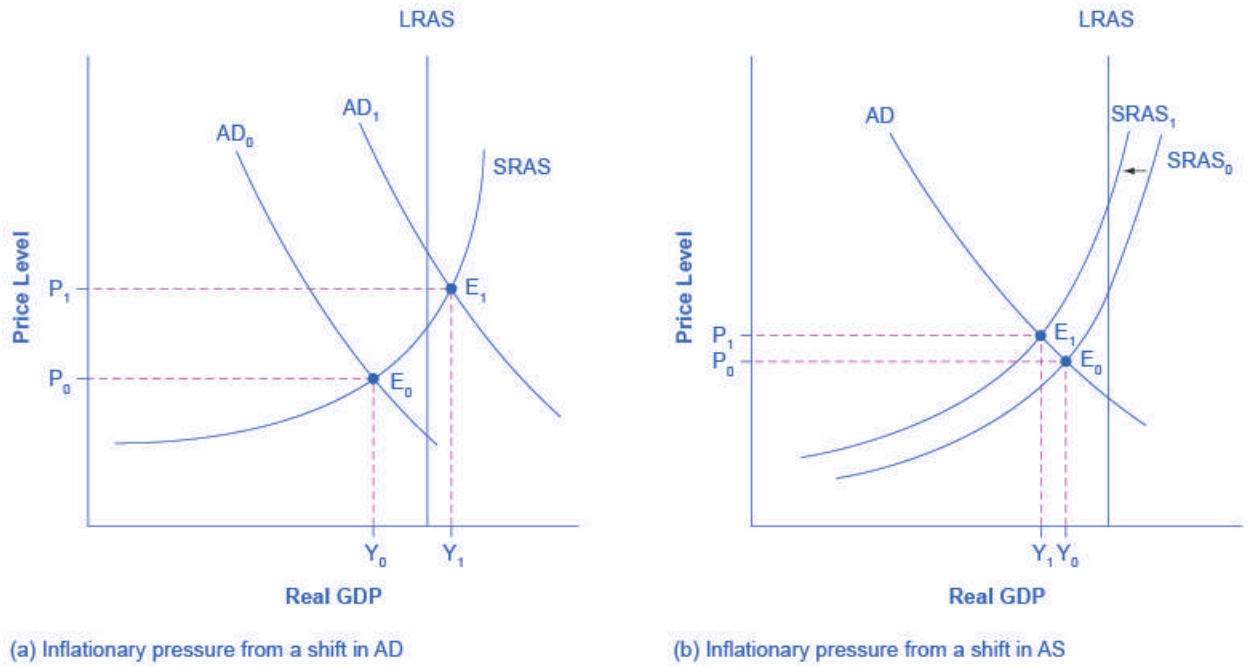


Figure 11.10 Sources of Inflationary Pressure in the AD/AS Model (a) A shift in aggregate demand, from AD_0 to AD_1, when it happens in the area of the SRAS curve that is near potential GDP, will lead to a higher price level and to pressure for a higher price level and inflation. The new equilibrium (E1) is at a higher price level (P1) than the original equilibrium. (b) A shift in aggregate supply, from $SRAS_0$ to $SRAS_1$, will lead to a lower real GDP and to pressure for a higher price level and inflation. The new equilibrium (E_1) is at a higher price level (P_1), while the original equilibrium (E_0) is at the lower price level (P_0).

An alternative source of inflationary pressures can occur due to a rise in input prices that affects many or most firms across the economy—perhaps an important input to production like oil or labor—and causes the aggregate supply curve to shift back to the left. In **Figure 11.10** (b), the SRAS curve's shift to the left also increases the price level from P_0 at the original equilibrium (E_0) to a higher price level of P_1 at the new equilibrium (E_1). In effect, the rise in input prices ends up, after the final output is produced and sold, passing along in the form of a higher price level for outputs.

The AD/AS diagram shows only a one-time shift in the price level. It does not address the question of what would cause inflation either to vanish after a year, or to sustain itself for several years. There are two explanations for why inflation may persist over time. One way that continual inflationary price increases can occur is if the government continually attempts to stimulate aggregate demand in a way that keeps pushing the AD curve when it is already in the SRAS curve's steep portion. A second possibility is that, if inflation has been occurring for several years, people might begin to expect a certain level of inflation. If they do, then these expectations will cause prices, wages and interest rates to increase annually by the amount of the inflation expected. These two reasons are interrelated, because if a government fosters a macroeconomic environment with inflationary pressures, then people will grow to expect inflation. However, the AD/AS diagram does not show these patterns of ongoing or expected inflation in a direct way.

Importance of the Aggregate Demand/Aggregate Supply Model

Macroeconomics takes an overall view of the economy, which means that it needs to juggle many different concepts. For example, start with the three macroeconomic goals of growth, low inflation, and low unemployment. Aggregate demand has four elements: consumption, investment, government spending, and exports less imports. Aggregate supply reveals how businesses throughout the economy will react to a higher price level for outputs. Finally, a wide array of economic events and policy decisions can affect aggregate demand and aggregate supply, including government tax and spending decisions; consumer and business confidence; changes in prices of key inputs like oil; and technology that brings higher levels of productivity.

The aggregate demand/aggregate supply model is one of the fundamental diagrams in this course (like the budget constraint diagram that we introduced in the **Choice in a World of Scarcity** chapter and the supply and demand

diagram in the **Demand and Supply** chapter) because it provides an overall framework for bringing these factors together in one diagram. Some version of the AD/AS model will appear in every chapter in the rest of this book.

11.6 | Keynes' Law and Say's Law in the AD/AS Model

By the end of this section, you will be able to:

- Identify the neoclassical zone, the intermediate zone, and the Keynesian zone in the aggregate demand/aggregate supply model
- Use an aggregate demand/aggregate supply model as a diagnostic test to understand the current state of the economy

We can use the AD/AS model to illustrate both Say's law that supply creates its own demand and Keynes' law that demand creates its own supply. Consider the SRAS curve's three zones which **Figure 11.11** identifies: the Keynesian zone, the neoclassical zone, and the intermediate zone.

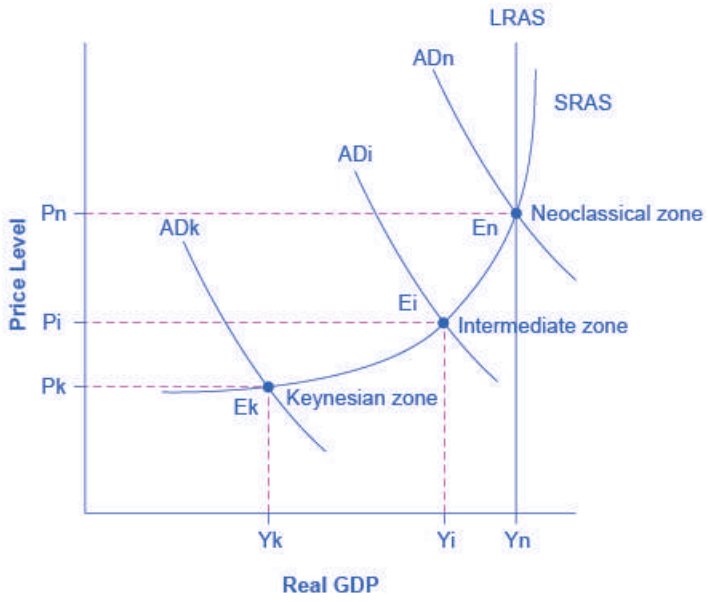


Figure 11.11 Keynes, Neoclassical, and Intermediate Zones in the Aggregate Supply Curve Near the equilibrium Ek, in the Keynesian zone at the far left of the SRAS curve, small shifts in AD, either to the right or the left, will affect the output level Yk, but will not much affect the price level. In the Keynesian zone, AD largely determines the quantity of output. Near the equilibrium En, in the neoclassical zone at the SRAS curve's far right, small shifts in AD, either to the right or the left, will have relatively little effect on the output level Yn, but instead will have a greater effect on the price level. In the neoclassical zone, the near-vertical SRAS curve close to the level of potential GDP largely determines the quantity of output. In the intermediate zone around equilibrium Ei, movement in AD to the right will increase both the output level and the price level, while a movement in AD to the left would decrease both the output level and the price level.

Focus first on the **Keynesian zone**, that portion of the SRAS curve on the far left which is relatively flat. If the AD curve crosses this portion of the SRAS curve at an equilibrium point like Ek, then certain statements about the economic situation will follow. In the Keynesian zone, the equilibrium level of real GDP is far below potential GDP, the economy is in recession, and cyclical unemployment is high. If aggregate demand shifted to the right or left in the Keynesian zone, it will determine the resulting level of output (and thus unemployment). However, inflationary price pressure is not much of a worry in the Keynesian zone, since the price level does not vary much in this zone.

Now, focus your attention on the **neoclassical zone** of the SRAS curve, which is the near-vertical portion on the right-hand side. If the AD curve crosses this portion of the SRAS curve at an equilibrium point like En where output is at or near potential GDP, then the size of potential GDP pretty much determines the level of output in the economy.

Since the equilibrium is near potential GDP, cyclical unemployment is low in this economy, although structural unemployment may remain an issue. In the neoclassical zone, shifts of aggregate demand to the right or the left have little effect on the level of output or employment. The only way to increase the size of the real GDP in the neoclassical zone is for AS to shift to the right. However, shifts in AD in the neoclassical zone will create pressures to change the price level.

Finally, consider the SRAS curve's **intermediate zone** in **Figure 11.11**. If the AD curve crosses this portion of the SRAS curve at an equilibrium point like Ei, then we might expect unemployment and inflation to move in opposing directions. For instance, a shift of AD to the right will move output closer to potential GDP and thus reduce unemployment, but will also lead to a higher price level and upward pressure on inflation. Conversely, a shift of AD to the left will move output further from potential GDP and raise unemployment, but will also lead to a lower price level and downward pressure on inflation.

This approach of dividing the SRAS curve into different zones works as a diagnostic test that we can apply to an economy, like a doctor checking a patient for symptoms. First, figure out in what zone the economy is. This will clarify the economic issues, tradeoffs, and policy choices. Some economists believe that the economy is strongly predisposed to be in one zone or another. Thus, hard-line Keynesian economists believe that the economies are in the Keynesian zone most of the time, and so they view the neoclassical zone as a theoretical abstraction. Conversely, hard-line neoclassical economists argue that economies are in the neoclassical zone most of the time and that the Keynesian zone is a distraction. **The Keynesian Perspective** and **The Neoclassical Perspective** should help to clarify the underpinnings and consequences of these contrasting views of the macroeconomy.

Bring it Home

From Housing Bubble to Housing Bust

We can explain economic fluctuations, whether those experienced during the 1930s Great Depression, the 1970s stagflation, or the 2008-2009 Great Recession, can be explained using the AD/AS diagram. Short-run fluctuations in output occur due to shifts of the SRAS curve, the AD curve, or both. In the case of the housing bubble, rising home values caused the AD curve to shift to the right as more people felt that rising home values increased their overall wealth. Many homeowners took on mortgages that exceeded their ability to pay because, as home values continued to rise, the increased value would pay off any debt outstanding. Increased wealth due to rising home values lead to increased home equity loans and increased spending. All these activities pushed AD to the right, contributing to low unemployment rates and economic growth in the United States. When the housing bubble burst, overall wealth dropped dramatically, wiping out the recent gains. This drop in home values was a demand shock to the U.S. economy because of its impact directly on the wealth of the household sector, and its contagion into the financial that essentially locked up new credit. The AD curve shifted to the left as evidenced by the Great Recession's rising unemployment.

Understanding the source of these macroeconomic fluctuations provided monetary and fiscal policy makers with insight about what policy actions to take to mitigate the impact of the housing crisis. From a monetary policy perspective, the Federal Reserve lowered short-term interest rates to between 0% and 0.25 %, to loosen up credit throughout the financial system. Discretionary fiscal policy measures included the passage of the Emergency Economic Stabilization Act of 2008 that allowed for the purchase of troubled assets, such as mortgages, from financial institutions and the American Recovery and Reinvestment Act of 2009 that increased government spending on infrastructure, provided for tax cuts, and increased transfer payments. In combination, both monetary and fiscal policy measures were designed to help stimulate aggregate demand in the U.S. economy, pushing the AD curve to the right.

While most economists agree on the usefulness of the AD/AS diagram in analyzing the sources of these fluctuations, there is still some disagreement about the effectiveness of policy decisions that are useful in stabilizing these fluctuations. We discuss the possible policy actions and the differences among economists about their effectiveness in more detail in **The Keynesian Perspective**, **Monetary Policy and Bank Regulation**, and **Government Budgets and Fiscal Policy**.

KEY TERMS

aggregate demand (AD) the amount of total spending on domestic goods and services in an economy

aggregate demand (AD) curve the total spending on domestic goods and services at each price level

aggregate demand/aggregate supply model a model that shows what determines total supply or total demand for the economy, and how total demand and total supply interact at the macroeconomic level

aggregate supply (AS) the total quantity of output (i.e. real GDP) firms will produce and sell

aggregate supply (AS) curve the total quantity of output (i.e. real GDP) that firms will produce and sell at each price level

full-employment GDP another name for potential GDP, when the economy is producing at its potential and unemployment is at the natural rate of unemployment

intermediate zone portion of the SRAS curve where GDP is below potential but not so far below as in the Keynesian zone; the SRAS curve is upward-sloping, but not vertical in the intermediate zone

Keynesian zone portion of the SRAS curve where GDP is far below potential and the SRAS curve is flat

Keynes' law "demand creates its own supply"

long run aggregate supply (LRAS) curve vertical line at potential GDP showing no relationship between the price level for output and real GDP in the long run

neoclassical economists economists who generally emphasize the importance of aggregate supply in determining the size of the macroeconomy over the long run

neoclassical zone portion of the SRAS curve where GDP is at or near potential output where the SRAS curve is steep

potential GDP the maximum quantity that an economy can produce given full employment of its existing levels of labor, physical capital, technology, and institutions

Say's law "supply creates its own demand"

short run aggregate supply (SRAS) curve positive short run relationship between the price level for output and real GDP, holding the prices of inputs fixed

stagflation an economy experiences stagnant growth and high inflation at the same time

KEY CONCEPTS AND SUMMARY

11.1 Macroeconomic Perspectives on Demand and Supply

Neoclassical economists emphasize Say's law, which holds that supply creates its own demand. Keynesian economists emphasize Keynes' law, which holds that demand creates its own supply. Many mainstream economists take a Keynesian perspective, emphasizing the importance of aggregate demand, for the short run, and a neoclassical perspective, emphasizing the importance of aggregate supply, for the long run.

11.2 Building a Model of Aggregate Demand and Aggregate Supply

The upward-sloping short run aggregate supply (SRAS) curve shows the positive relationship between the price level and the level of real GDP in the short run. Aggregate supply slopes up because when the price level for outputs increases, while the price level of inputs remains fixed, the opportunity for additional profits encourages more production. The aggregate supply curve is near-horizontal on the left and near-vertical on the right. In the long run,

we show the aggregate supply by a vertical line at the level of potential output, which is the maximum level of output the economy can produce with its existing levels of workers, physical capital, technology, and economic institutions.

The downward-sloping aggregate demand (AD) curve shows the relationship between the price level for outputs and the quantity of total spending in the economy. It slopes down because of: (a) the wealth effect, which means that a higher price level leads to lower real wealth, which reduces the level of consumption; (b) the interest rate effect, which holds that a higher price level will mean a greater demand for money, which will tend to drive up interest rates and reduce investment spending; and (c) the foreign price effect, which holds that a rise in the price level will make domestic goods relatively more expensive, discouraging exports and encouraging imports.

11.3 Shifts in Aggregate Supply

The aggregate demand/aggregate supply (AD/AS) diagram shows how AD and AS interact. The intersection of the AD and AS curves shows the equilibrium output and price level in the economy. Movements of either AS or AD will result in a different equilibrium output and price level. The aggregate supply curve will shift out to the right as productivity increases. It will shift back to the left as the price of key inputs rises, and will shift out to the right if the price of key inputs falls. If the AS curve shifts back to the left, the combination of lower output, higher unemployment, and higher inflation, called stagflation, occurs. If AS shifts out to the right, a combination of lower inflation, higher output, and lower unemployment is possible.

11.4 Shifts in Aggregate Demand

The AD curve will shift out as the components of aggregate demand—C, I, G, and X–M—rise. It will shift back to the left as these components fall. These factors can change because of different personal choices, like those resulting from consumer or business confidence, or from policy choices like changes in government spending and taxes. If the AD curve shifts to the right, then the equilibrium quantity of output and the price level will rise. If the AD curve shifts to the left, then the equilibrium quantity of output and the price level will fall. Whether equilibrium output changes relatively more than the price level or whether the price level changes relatively more than output is determined by where the AD curve intersects with the AS curve.

The AD/AS diagram superficially resembles the microeconomic supply and demand diagram on the surface, but in reality, what is on the horizontal and vertical axes and the underlying economic reasons for the shapes of the curves are very different. We can illustrate long-term economic growth in the AD/AS framework by a gradual shift of the aggregate supply curve to the right. We illustrate a recession when the intersection of AD and AS is substantially below potential GDP, while we illustrate an expanding economy when the intersection of AS and AD is near potential GDP.

11.5 How the AD/AS Model Incorporates Growth, Unemployment, and Inflation

Cyclical unemployment is relatively large in the AD/AS framework when the equilibrium is substantially below potential GDP. Cyclical unemployment is small in the AD/AS framework when the equilibrium is near potential GDP. The natural rate of unemployment, as determined by the labor market institutions of the economy, is built into what economists mean by potential GDP, but does not otherwise appear in an AD/AS diagram. The AD/AS framework shows pressures for inflation to rise or fall when the movement from one equilibrium to another causes the price level to rise or to fall. The balance of trade does not appear directly in the AD/AS diagram, but it appears indirectly in several ways. Increases in exports or declines in imports can cause shifts in AD. Changes in the price of key imported inputs to production, like oil, can cause shifts in AS. The AD/AS model is the key model we use in this book to understand macroeconomic issues.

11.6 Keynes' Law and Say's Law in the AD/AS Model

We can divide the SRAS curve into three zones. Keynes' law says demand creates its own supply, so that changes in aggregate demand cause changes in real GDP and employment. We can show Keynes' law on the horizontal Keynesian zone of the aggregate supply curve. The Keynesian zone occurs at the left of the SRAS curve where it is fairly flat, so movements in AD will affect output, but have little effect on the price level. Say's law says supply creates its own demand. Changes in aggregate demand have no effect on real GDP and employment, only on the price level. We can show Say's law on the vertical neoclassical zone of the aggregate supply curve. The neoclassical zone occurs at the right of the SRAS curve where it is fairly vertical, and so movements in AD will affect the price level, but have little impact on output. The intermediate zone in the middle of the SRAS curve is upward-sloping, so a rise in AD will cause higher output and price level, while a fall in AD will lead to a lower output and price level.

SELF-CHECK QUESTIONS

1. Describe the mechanism by which supply creates its own demand.

2. Describe the mechanism by which demand creates its own supply.

3. The short run aggregate supply curve was constructed assuming that as the price of outputs increases, the price of inputs stays the same. How would an increase in the prices of important inputs, like energy, affect aggregate supply?

4. In the AD/AS model, what prevents the economy from achieving equilibrium at potential output?

5. Suppose the U.S. Congress passes significant immigration reform that makes it more difficult for foreigners to come to the United States to work. Use the AD/AS model to explain how this would affect the equilibrium level of GDP and the price level.

6. Suppose concerns about the size of the federal budget deficit lead the U.S. Congress to cut all funding for research and development for ten years. Assuming this has an impact on technology growth, what does the AD/AS model predict would be the likely effect on equilibrium GDP and the price level?

7. How would a dramatic increase in the value of the stock market shift the AD curve? What effect would the shift have on the equilibrium level of GDP and the price level?

8. Suppose Mexico, one of our largest trading partners and purchaser of a large quantity of our exports, goes into a recession. Use the AD/AS model to determine the likely impact on our equilibrium GDP and price level.

9. A policymaker claims that tax cuts led the economy out of a recession. Can we use the AD/AS diagram to show this?

10. Many financial analysts and economists eagerly await the press releases for the reports on the home price index and consumer confidence index. What would be the effects of a negative report on both of these? What about a positive report?

11. What impact would a decrease in the size of the labor force have on GDP and the price level according to the AD/AS model?

12. Suppose, after five years of sluggish growth, the European Union's economy picks up speed. What would be the likely impact on the U.S. trade balance, GDP, and employment?

13. Suppose the Federal Reserve begins to increase the supply of money at an increasing rate. What impact would that have on GDP, unemployment, and inflation?

14. If the economy is operating in the neoclassical zone of the SRAS curve and aggregate demand falls, what is likely to happen to real GDP?

15. If the economy is operating in the Keynesian zone of the SRAS curve and aggregate demand falls, what is likely to happen to real GDP?

REVIEW QUESTIONS

16. What is Say’s law?

17. What is Keynes’ law?

18. Do neoclassical economists believe in Keynes’ law or Say’s law?

19. Does Say’s law apply more accurately in the long run or the short run? What about Keynes’ law?

20. What is on the horizontal axis of the AD/AS diagram? What is on the vertical axis?

21. What is the economic reason why the SRAS curve slopes up?

22. What are the components of the aggregate demand (AD) curve?

23. What are the economic reasons why the AD curve slopes down?

24. Briefly explain the reason for the near-horizontal shape of the SRAS curve on its far left.

25. Briefly explain the reason for the near-vertical shape of the SRAS curve on its far right.

26. What is potential GDP?

27. Name some factors that could cause the SRAS curve to shift, and say whether they would shift SRAS to the right or to the left.

28. Will the shift of SRAS to the right tend to make the equilibrium quantity and price level higher or lower? What about a shift of SRAS to the left?

29. What is stagflation?

30. Name some factors that could cause AD to shift, and say whether they would shift AD to the right or to the left.

31. Would a shift of AD to the right tend to make the equilibrium quantity and price level higher or lower? What about a shift of AD to the left?

32. How is long-term growth illustrated in an AD/AS model?

33. How is recession illustrated in an AD/AS model?

34. How is cyclical unemployment illustrated in an AD/AS model?

35. How is the natural rate of unemployment illustrated in an AD/AS model?

36. How is pressure for inflationary price increases shown in an AD/AS model?

37. What are some of the ways in which exports and imports can affect the AD/AS model?

38. What is the Keynesian zone of the SRAS curve? How much is the price level likely to change in the Keynesian zone?

39. What is the neoclassical zone of the SRAS curve? How much is the output level likely to change in the neoclassical zone?

40. What is the intermediate zone of the SRAS curve? Will a rise in output be accompanied by a rise or a fall in the price level in this zone?

CRITICAL THINKING QUESTIONS

41. Why would an economist choose either the neoclassical perspective or the Keynesian perspective, but not both?

42. On a microeconomic demand curve, a decrease in price causes an increase in quantity demanded because the product in question is now relatively less expensive than substitute products. Explain why aggregate demand does not increase for the same reason in response to a decrease in the aggregate price level. In other words, what causes total spending to increase if it is not because goods are now cheaper?

43. Economists expect that as the labor market continues to tighten going into the latter part of 2015 that workers should begin to expect wage increases in 2015 and 2016. Assuming this occurs and it was the only development in the labor market that year, how would this affect the AS curve? What if it was also accompanied by an increase in worker productivity?

44. If new government regulations require firms to use a cleaner technology that is also less efficient than what they previously used, what would the effect be on output, the price level, and employment using the AD/AS diagram?

45. During spring 2016 the Midwestern United States, which has a large agricultural base, experiences above-average rainfall. Using the AD/AS diagram, what is the effect on output, the price level, and employment?

46. Hydraulic fracturing (fracking) has the potential to significantly increase the amount of natural gas produced in the United States. If a large percentage of factories and utility companies use natural gas, what will happen to output, the price level, and employment as fracking becomes more widely used?

47. Some politicians have suggested tying the minimum wage to the consumer price index (CPI). Using the AD/AS diagram, what effects would this policy most likely have on output, the price level, and employment?

48. If households decide to save a larger portion of their income, what effect would this have on the output, employment, and price level in the short run? What about the long run?

49. If firms become more optimistic about the future of the economy and, at the same time, innovation in 3-D printing makes most workers more productive, what is the combined effect on output, employment, and the price-level?

50. If Congress cuts taxes at the same time that businesses become more pessimistic about the economy, what is the combined effect on output, the price level, and employment using the AD/AS diagram?

51. Suppose the level of structural unemployment increases. How would you illustrate the increase in structural unemployment in the AD/AS model? *Hint*: How does structural unemployment affect potential GDP?

52. If foreign wealth-holders decide that the United States is the safest place to invest their savings, what would the effect be on the economy here? Show graphically using the AD/AS model.

53. The AD/AS model is static. It shows a snapshot of the economy at a given point in time. Both economic growth and inflation are dynamic phenomena. Suppose economic growth is 3% per year and aggregate demand is growing at the same rate. What does the AD/AS model say the inflation rate should be?

54. Explain why the short-run aggregate supply curve might be fairly flat in the Keynesian zone of the SRAS curve. How might we tell if we are in the Keynesian zone of the AS?

55. Explain why the short-run aggregate supply curve might be vertical in the neoclassical zone of the SRAS curve. How might we tell if we are in the neoclassical zone of the AS?

56. Why might it be important for policymakers to know which in zone of the SRAS curve the economy is?

57. In your view, is the economy currently operating in the Keynesian, intermediate or neoclassical portion of the economy's aggregate supply curve?

58. Are Say's law and Keynes' law necessarily mutually exclusive?

PROBLEMS

59. Review the problem in the **Work It Out** titled "Interpreting the AD/AS Model." Like the information provided in that feature, **Table 11.2** shows information on aggregate supply, aggregate demand, and the price level for the imaginary country of Xurbia.

Price Level	AD	AS
110	700	600
120	690	640
130	680	680
140	670	720
150	660	740
160	650	760
170	640	770

Table 11.2 Price Level: AD/AS

a. Plot the AD/AS diagram from the data. Identify the equilibrium.
b. Imagine that, as a result of a government tax cut, aggregate demand becomes higher by 50 at every price level. Identify the new equilibrium.
c. How will the new equilibrium alter output? How will it alter the price level? What do you think will happen to employment?

60. The imaginary country of Harris Island has the aggregate supply and aggregate demand curves as **Table 11.3** shows.

Price Level	AD	AS
100	700	200
120	600	325
140	500	500
160	400	570
180	300	620

Table 11.3 Price Level: AD/AS

a. Plot the AD/AS diagram. Identify the equilibrium.
b. Would you expect unemployment in this economy to be relatively high or low?
c. Would you expect concern about inflation in this economy to be relatively high or low?
d. Imagine that consumers begin to lose confidence about the state of the economy, and so AD becomes lower by 275 at every price level. Identify the new aggregate equilibrium.
e. How will the shift in AD affect the original output, price level, and employment?

61. **Table 11.4** describes Santher's economy.

Price Level	AD	AS
50	1,000	250
60	950	580
70	900	750
80	850	850
90	800	900

Table 11.4 Price Level: AD/AS

a. Plot the AD/AS curves and identify the equilibrium.
b. Would you expect unemployment in this economy to be relatively high or low?
c. Would you expect prices to be a relatively large or small concern for this economy?
d. Imagine that input prices fall and so AS shifts to the right by 150 units. Identify the new equilibrium.
e. How will the shift in AS affect the original output, price level, and employment?