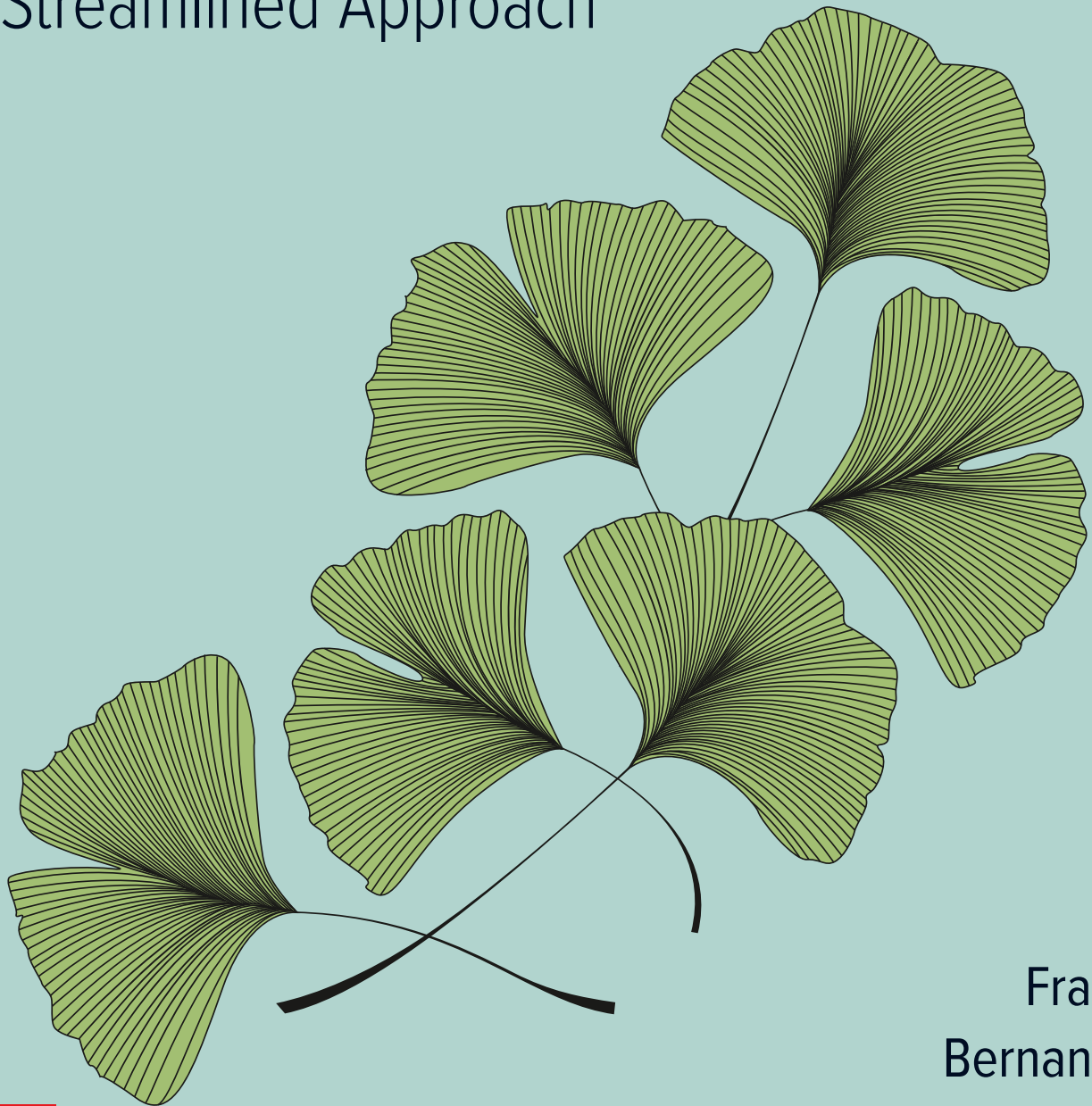


This International Student Edition is for use outside of the U.S.

fourth edition

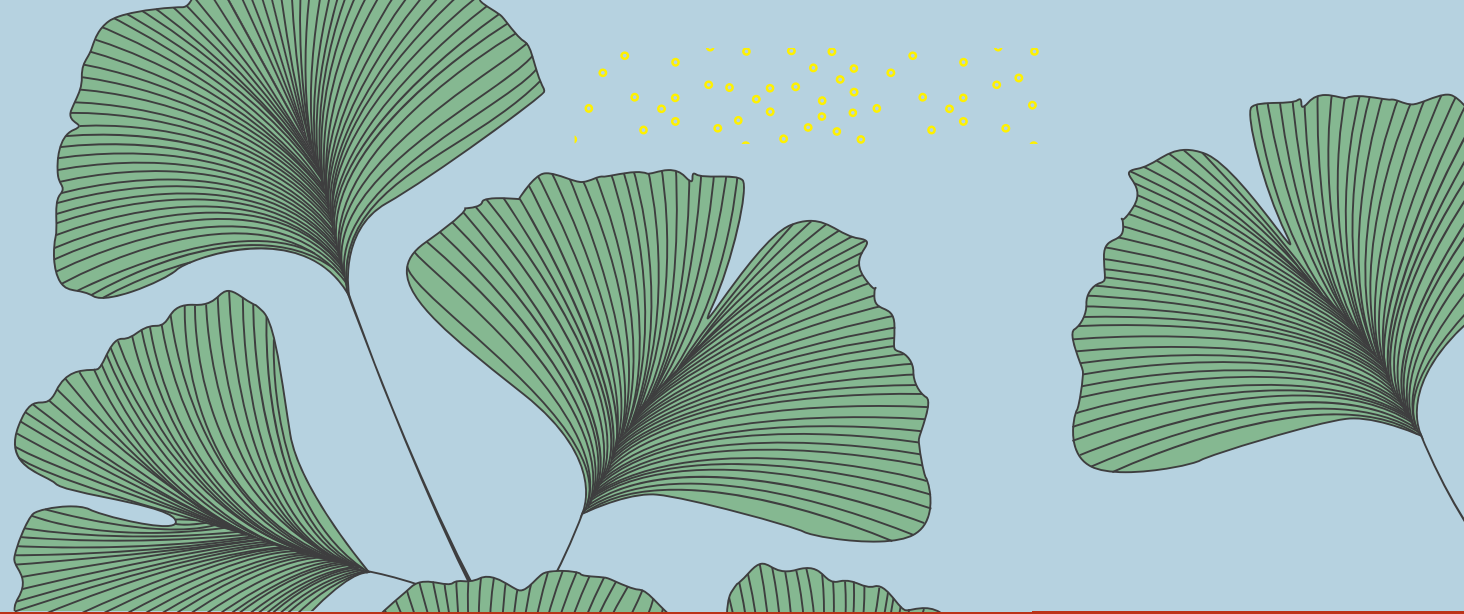
Principles of Microeconomics

A Streamlined Approach



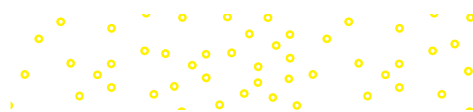
**Mc
Graw
Hill**

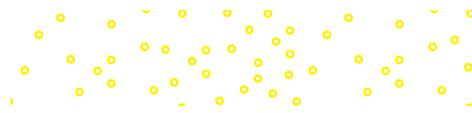
Frank
Bernanke
Antonovics
Heffetz



FOURTH EDITION

Principles of
MICROECONOMICS
A STREAMLINED APPROACH





THE MCGRAW HILL SERIES IN ECONOMICS

SURVEY OF ECONOMICS

Brue, McConnell, and Flynn
Essentials of Economics
Fourth Edition

Guell
Issues in Economics Today
Ninth Edition

Mandel
M: Economics, The Basics
Fourth Edition

Schiller and Gebhardt
Essentials of Economics
Eleventh Edition

PRINCIPLES OF ECONOMICS

Asarta and Butters
Connect Master Principles of
Economics

Colander
Economics, Microeconomics, and
Macroeconomics
Eleventh Edition

**Frank, Bernanke, Antonovics,
and Heffetz**
Principles of Economics, Principles of
Microeconomics, Principles of
Macroeconomics
Eighth Edition

**Frank, Bernanke, Antonovics,
and Heffetz**
A Streamlined Approach for:
Principles of Economics, Principles of
Microeconomics, and Principles of
Macroeconomics
Fourth Edition

Karlan and Morduch
Economics, Microeconomics, and
Macroeconomics
Third Edition

McConnell, Brue, and Flynn
Economics, Microeconomics, and
Macroeconomics
Twenty-Second Edition

McConnell, Brue, and Flynn
Brief Editions: Microeconomics and
Macroeconomics
Third Edition

Samuelson and Nordhaus
Economics, Microeconomics, and
Macroeconomics
Nineteenth Edition

Schiller and Gebhardt
The Economy Today, The Micro
Economy Today, and
The Macro Economy Today
Sixteenth Edition

Slavin
Economics, Microeconomics, and
Macroeconomics
Twelfth Edition

ECONOMICS OF SOCIAL ISSUES

Guell
Issues in Economics Today
Ninth Edition

Register and Grimes
Economics of Social Issues
Twenty-First Edition

DATA ANALYTICS FOR ECONOMICS

Prince
Predictive Analytics for Business
Strategy
First Edition

MANAGERIAL ECONOMICS

Baye and Prince
Managerial Economics and
Business Strategy
Tenth Edition

**Brickley, Smith, and
Zimmerman**
Managerial Economics and
Organizational
Architecture
Seventh Edition

Thomas and Maurice
Managerial Economics
Thirteenth Edition

INTERMEDIATE ECONOMICS

Bernheim and Whinston
Microeconomics
Second Edition

Dornbusch, Fischer, and Startz
Macroeconomics
Thirteenth Edition

Frank
Microeconomics and Behavior
Tenth Edition

ADVANCED ECONOMICS

Romer
Advanced Macroeconomics
Fifth Edition

MONEY AND BANKING

Cecchetti and Schoenholtz
Money, Banking, and Financial
Markets
Sixth Edition

URBAN ECONOMICS

O'Sullivan
Urban Economics
Ninth Edition

LABOR ECONOMICS

Borjas
Labor Economics
Eighth Edition

McConnell, Brue, and Macpherson
Contemporary Labor Economics
Twelfth Edition

PUBLIC FINANCE

Rosen and Gayer
Public Finance
Tenth Edition

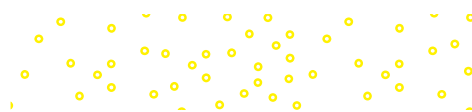
ENVIRONMENTAL ECONOMICS

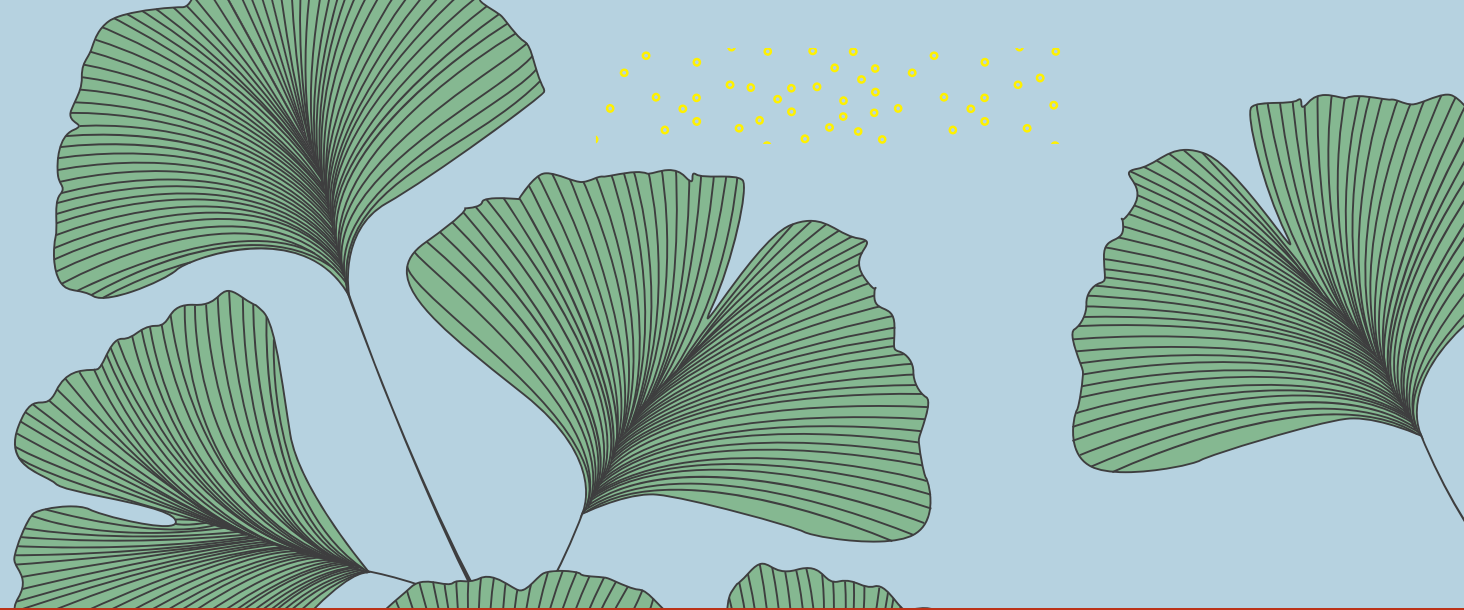
Field and Field
Environmental Economics: An
Introduction
Eighth Edition

INTERNATIONAL ECONOMICS

Appleyard and Field
International Economics
Ninth Edition

Pugel
International Economics
Seventeenth Edition





FOURTH EDITION

Principles of **MICROECONOMICS**

A STREAMLINED APPROACH

ROBERT H. FRANK

Cornell University

BEN S. BERNANKE

Brookings Institution [affiliated]

Former Chair, Board of Governors of the Federal Reserve System

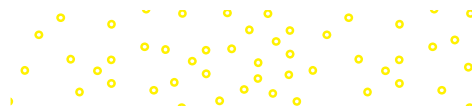
KATE ANTONOVICS

University of California, San Diego

ORI HEFFETZ

Cornell University and Hebrew University of Jerusalem

**Mc
Graw
Hill**



PRINCIPLES OF MICROECONOMICS: A STREAMLINED APPROACH

Published by McGraw Hill LLC, 1325 Avenue of the Americas, New York, NY 10121. Copyright ©2022 by McGraw Hill LLC. All rights reserved. Printed in the United States of America. No part of this publication may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without the prior written consent of McGraw Hill LLC, including, but not limited to, in any network or other electronic storage or transmission, or broadcast for distance learning.

Some ancillaries, including electronic and print components, may not be available to customers outside the United States.

This book is printed on acid-free paper.

1 2 3 4 5 6 7 8 9 LWI 26 25 24 23 22 21

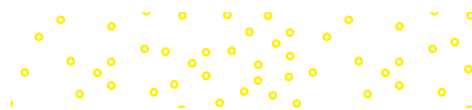
ISBN 978-1-264-37060-3

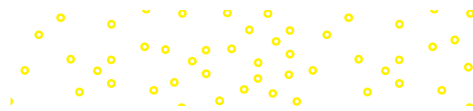
MHID 1-264-37060-1

Cover Image: *Miroslava Hlavacova/Shutterstock*

All credits appearing on page or at the end of the book are considered to be an extension of the copyright page.

The Internet addresses listed in the text were accurate at the time of publication. The inclusion of a website does not indicate an endorsement by the authors or McGraw Hill LLC, and McGraw Hill LLC does not guarantee the accuracy of the information presented at these sites.





DEDICATION

For Ellen

R. H. F.

For Anna

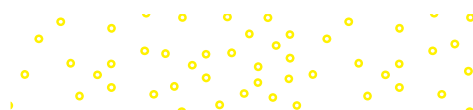
B. S. B.

For Fiona and Henry

K. A.

For Katrina, Eleanor, Daniel, and Amalia

O. H.



ABOUT THE AUTHORS

ROBERT H. FRANK



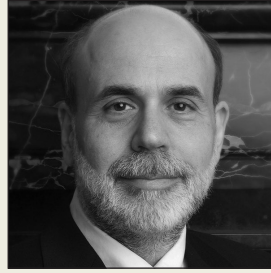
©Robert H. Frank

Robert H. Frank is the H. J. Louis Professor of Management and Professor of Economics, emeritus, at Cornell's Johnson School of Management, where he taught from 1972 to 2020. After receiving his B.S. from Georgia Tech in 1966, he taught math and science for two years as a Peace Corps Volunteer in rural Nepal. He received his M.A. in statistics in 1971 and his Ph.D. in economics in 1972 from The University of California at Berkeley. He also holds honorary doctorate degrees from the University of St. Gallen and Dalhousie University. During leaves of absence from Cornell, he has served as chief economist for the Civil Aeronautics Board (1978–1980), a Fellow at the Center for Advanced Study in the Behavioral Sciences (1992–1993), Professor of American Civilization at l'École des Hautes Études en Sciences Sociales in Paris (2000–2001), and the Peter and Charlotte Schoenfeld Visiting Faculty Fellow at the NYU Stern School of Business in 2008–2009. His papers have appeared in the *American Economic Review*, *Econometrica*, the *Journal of Political Economy*, and other leading professional journals, and for more than two decades, his economics columns appeared regularly in *The New York Times*.

Professor Frank is the author of a best-selling intermediate economics textbook *Microeconomics and Behavior*, Tenth Edition (McGraw Hill, 2021). His research has focused on rivalry and cooperation in economic and social behavior. His books on these themes include *Choosing the Right Pond* (Oxford, 1985), *Passions Within Reason* (W. W. Norton, 1988), *What Price the Moral High Ground?* (Princeton, 2004), *Falling Behind* (University of California Press, 2007), *The Economic Naturalist* (Basic Books, 2007), *The Economic Naturalist's Field Guide* (Basic Books, 2009), *The Darwin Economy* (Princeton, 2011), *Success and Luck* (Princeton, 2016), and *Under the Influence* (Princeton, 2020), which have been translated into 24 languages. *The Winner-Take-All Society* (The Free Press, 1995), co-authored with Philip Cook, received a Critic's Choice Award, was named a Notable Book of the Year by *The New York Times*, and was included in *BusinessWeek's* list of the 10 best books of 1995. *Luxury Fever* (The Free Press, 1999) was named to the Knight-Ridder Best Books list for 1999.

Professor Frank is a co-recipient of the 2004 Leontief Prize for Advancing the Frontiers of Economic Thought. He was awarded the Johnson School's Stephen Russell Distinguished Teaching Award in 2004, 2010, 2012, and 2018, and the School's Apple Distinguished Teaching Award in 2005. His introductory microeconomics course has graduated more than 7,000 enthusiastic economic naturalists over the years.

BEN S. BERNANKE



©Ben S. Bernanke

Professor Bernanke received his B.A. in economics from Harvard University in 1975 and his Ph.D. in economics from MIT in 1979. He taught at the Stanford Graduate School of Business from 1979 to 1985 and moved to Princeton University in 1985, where he was named the Howard Harrison and Gabrielle Snyder Beck Professor of Economics and Public Affairs and where he served as chair of the Economics Department. Professor Bernanke is currently a Distinguished Fellow in Residence with the Economic Studies Program at the Brookings Institution.

Professor Bernanke was sworn in on February 1, 2006, as chair and a member of the Board of Governors of the Federal Reserve System—his second term expired January 31, 2014. Professor Bernanke also served as chair of the Federal Open Market Committee, the Fed's principal monetary policymaking body. Professor Bernanke was also chair of the President's Council of Economic Advisers from June 2005 to January 2006.

Professor Bernanke's intermediate textbook, with Andrew Abel and Dean Croushore, *Macroeconomics*, Ninth Edition (Addison-Wesley, 2017), is a best seller in its field. He has authored numerous scholarly publications in macroeconomics, macroeconomic history, and finance. He has done significant research on the causes of the Great Depression, the role of financial markets and institutions in the business cycle, and measurement of the effects of monetary policy on the economy.

Professor Bernanke has held a Guggenheim Fellowship and a Sloan Fellowship, and he is a Fellow of the Econometric Society and of the American Academy of Arts and Sciences. He served as director of the Monetary Economics Program of the National Bureau of Economic Research (NBER) and as a member of the NBER's Business Cycle Dating Committee. From 2001 to 2004 he served as editor of the *American Economic Review*, and served as the president of the American Economic Association in 2019. Professor Bernanke's work with civic and professional groups includes having served two terms as a member of the Montgomery Township (N.J.) Board of Education.

KATE ANTONOVICS



©Kate Antonovics

Professor Antonovics received her B.A. from Brown University in 1993 and her Ph.D. in economics from the University of Wisconsin in 2000. Shortly thereafter, she joined the faculty in the Economics Department at the University of California, San

Diego. Professor Antonovics is also currently serving as the Provost of UC San Diego's Seventh College.

Professor Antonovics is known for her excellence in teaching and her innovative use of technology in the classroom. Her popular introductory-level microeconomics courses have regularly enrolled over 900 students each fall. She also teaches labor economics at both the undergraduate and graduate level. She has received numerous teaching awards, including the UCSD Department of Economics award for Best Undergraduate Teaching, the UCSD Academic Senate Distinguished Teaching Award, and the UCSD Chancellor's Associates Faculty Excellence Award in Undergraduate Teaching.

Professor Antonovics's research has focused on racial discrimination, gender discrimination, affirmative action, intergenerational income mobility, learning, and wage dynamics. Her papers have appeared in the *American Economic Review*, the *Review of Economics and Statistics*, the *Journal of Labor Economics*, and the *Journal of Human Resources*. She is a member of both the American Economic Association and the Society of Labor Economists.

ORI HEFFETZ



©Ori Heffetz

Professor Heffetz received his B.A. in physics and philosophy from Tel Aviv University in 1999 and his Ph.D. in economics from Princeton University in 2005. He is an Associate Professor of Economics at the Samuel Curtis Johnson Graduate School of Management at Cornell University, and at the Economics Department at the

Hebrew University of Jerusalem.

Bringing the real world into the classroom, Professor Heffetz has created a unique macroeconomics course that introduces basic concepts and tools from economic theory and applies them to current news and global events. His popular classes are taken by hundreds of students every year on Cornell's Ithaca and New York City campuses and via live videoconferencing in dozens of cities across the United States, Canada, and Latin America.

Professor Heffetz's research studies the social and cultural aspects of economic behavior, focusing on the mechanisms that drive consumers' choices and on the links between economic choices, individual well-being, and policymaking. He has published scholarly work on household consumption patterns, individual economic decision making, and survey methodology and measurement. He was a visiting researcher at the Bank of Israel during 2011, is currently a Research Associate at the National Bureau of Economic Research (NBER), and serves on the editorial board of *Social Choice and Welfare*.

A LESS IS MORE FOCUS

Our fourth streamlined edition arrives in the midst of some of the most dramatic upheavals ever witnessed, both in the economy generally and in higher education in particular. The COVID-19 pandemic has produced levels of unemployment not seen since the Great Depression and has created dramatic changes in the ways we teach across educational institutions at every level.

These developments have reinforced our confidence in the instructional philosophy that motivated us to produce our first edition—the need to strip away clutter and focus more intensively on central concepts. This approach, we believe, is especially well suited for the new environment.

In earlier editions, we noted that although many millions of dollars are spent each year on introductory economics instruction in American colleges and universities, the return on this investment has been disturbingly low. Studies have shown, for example, that several months after having taken a principles of economics course, former students are no better able to answer simple economics questions than others who never even took the course. Most students, it seems, leave our introductory courses without having learned even the most important basic economic principles. Such dismal performance, never defensible, has become even more difficult to justify in the face of looming resource shortages in higher education.

The problem, in our view, has almost always been that courses try to teach students far too much. In the process, really important ideas get little more coverage than minor ones, and everything ends up going by in a blur. The human brain tends to ignore new information unless it comes up repeatedly. That's hardly surprising, since only a tiny fraction of the terabytes of information that bombard us each day is likely to be relevant for anything we care about. Only when something comes up a third or fourth time does the brain start laying down new circuits for dealing with it. Yet when planning their lectures, many instructors ask themselves, "How much can I cover today?" And because modern electronic media enable them to click through upwards of 100 PowerPoint slides in an hour, they feel they better serve their students when they put more information before them. But that's not the way learning works! Professors should instead be asking, "How much can my students absorb?"



Our approach to this text was inspired by our conviction that students will learn far more if we attempt to cover much less. Our basic premise is that a small number of basic ideas do most of the heavy lifting in economics, and that if we focus narrowly and repeatedly on those ideas, illustrating and applying them in numerous familiar contexts, students can actually master them in just a single semester. The enthusiastic reactions of users of previous editions of our textbook affirm the validity of this premise. Our emphasis throughout is on active learning. We ask students to apply basic economic ideas themselves to answer related questions, exercises, and problems.

ADAPTING TO CLASSROOM TRENDS

Baumol's cost disease refers to the tendency for costs to rise more rapidly for goods and services for which growth in labor productivity is either slow or nonexistent. It is thus no surprise that the cost of traditional methods of delivering classroom instruction has been rising so much faster than the cost of producing most manufactured goods.

Largely as a result of Baumol's cost disease, tuition increases have far exceeded even the rapid growth in the cost of health care. This is what we would expect if the dominant teaching model remains as it was a century ago, in which a learned instructor stands in front of a class reciting truths cataloged in an assigned text. But as the late Herb Stein once remarked, "If something cannot go on forever, it will stop." And so it is with rising tuitions. Universities are already facing strong pressure to moderate their rates of tuition growth, pressure that has been greatly exacerbated by the COVID-19 pandemic.

One result has been that much of the content that professors have traditionally delivered in live lecture will instead be delivered electronically. Indeed, technological advances have given today's students an unparalleled ability to access information via the Internet, YouTube, and social media.

If early experience is any indication, the "flipped-classroom" model is one of the most promising adaptations to this new environment. In this approach, students are expected to study basic concepts before coming to class and then deepen their understanding of them through structured classroom exercises and discussion. The logic of the flipped classroom is compelling because under this approach, students have access to instructors precisely when students are engaged in those activities that students find the most challenging (for example, problem solving and policy evaluation). Indeed, numerous studies have found that the flipped-classroom

approach increases both student satisfaction and student learning.

A second adaptation, more pronounced in the wake of the COVID-19 lockdowns of March 2020, has been the move to remote instruction. The streamlined approach of this text is well suited for the goals of both the flipped-classroom and remote-instruction models. Rather than trying to bombard students with information they can easily access online, our book seeks to promote a deeper understanding of economics by focusing on core concepts. In addition, one of our central goals has been to create resources to help instructors promote student engagement outside the classroom. Some instructors may find these resources useful in completely overhauling the way they teach, while others may be interested in using them to make a few minor changes to their current courses.

In other words, this edition is intended to support a variety of teaching styles (and, indeed, our team of authors varies considerably in our pedagogical approach). The traditional approach has been to ask students to read the relevant sections from the textbook before coming to class. But instructors report that today's students are far less likely than their predecessors to complete such assignments. Stronger incentives can boost compliance. One effective approach assigns SmartBook chapters with adaptive questioning built in. Another administers brief tests at the start of class. These might involve two or three simple multiple-choice questions on the assigned material that are administered and graded electronically via audience response tools using smartphone apps.

Perhaps the biggest hurdle to effective implementation of the new teaching approaches has been a dearth of effective pre-class concept-delivery materials. To help fill this gap, we have created a library of short videos that focus on basic economic concepts. Many students have found these videos and animations engaging enough to watch even if they're not going to be tested on them, but we've also provided easily administered in-class questions that can boost compliance still further.

A big payoff in both the flipped-classroom and remote-instruction models comes from being able to use limited class time to discuss the concepts that students have studied before class. One approach begins by asking students to answer a multiple-choice question requiring application of a concept, and then reporting the frequencies with which students selected the various multiple-choice options. Students are then given a few moments to discuss the question—either with their neighbors in traditional classroom settings, or with fellow students in Zoom breakout rooms—before having an

opportunity to change the answers they originally submitted. Professors then call on students who've offered both correct and incorrect answers to the question to defend their answers to the class and lead the ensuing discussion. We've spent considerable effort drafting the kinds of questions that reliably provoke animated discussions of this sort.

In summary, here are the resources we have developed to support the flipped-classroom and remote-instruction approaches, all available within McGraw Hill Connect® specific to the fourth edition:

Before Class (Exposure)

- **SmartBook® Adaptive Reading Assignments:** SmartBook contains the same content as the print book, but actively tailors that content to the needs of the individual through adaptive probing and integrated learning resources. Instructions can assign SmartBook reading assignments for points to create incentives for students to come to class prepared.
- **Learning Glass Lecture Videos:** A collection of brief instructional videos featuring the authors, Kate Antonovics and Ori Heffetz, utilize exciting learning glass technology to provide students with an overview of important economic concepts. Perfect for an introduction to basic concepts before coming to class, or as a quick review, these videos can be accessed as resources within SmartBook, or are available as stand-alone assignments within Connect.

In Class (Engagement)

- **Clicker Questions:** Classroom-tested by the authors, these multiple-choice questions are designed to facilitate discussion and group work in class.
- **Economic Naturalist Application-Focused Videos:** A known hallmark of this franchise, the Economic Naturalist examples are now available as an expanded set of short, engaging video vignettes within Connect and SmartBook.

After Class (Reinforcement)

- **Connect Exercises:** All end-of-chapter homework exercises are available to be assigned within Connect. Many of these exercises include algorithmic variations and require students to interact with the graphing and tool within the platform. **Worked Problem Videos**, available as hints within Connect, work through these problems to aid in student understanding of core economic concepts and offer assistance with more challenging material.

- **Test Bank Assessment:** Hundreds of multiple-choice questions are available for summative assessments of the chapter content. Select problems are now offered as an algorithmic alternative, providing even more variation.

All of the above assets can be implemented by instructors as preferred in order to satisfy as much or as little of the flipped-classroom approach as is desired.

KEY THEMES AND FEATURES

Economic Naturalism

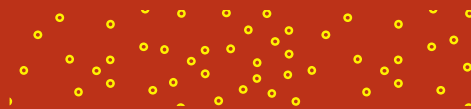
Relying on examples drawn from familiar contexts, we encourage students to become “economic naturalists,” people who employ basic economic principles to understand and explain what they observe in the world around them. An economic naturalist understands, for example, that infant safety seats are required in cars but not in airplanes because the marginal cost of space to accommodate these seats is typically zero in cars but often hundreds of dollars in airplanes. Scores of such examples are sprinkled throughout the text. Each one, we believe, poses a question that should make any curious person eager to learn the answer.

Our ultimate goal is to produce economic naturalists—people who see each human interaction as the result of an implicit or explicit cost-benefit calculation.

The economic naturalist sees mundane details of ordinary existence in a new light and becomes actively engaged in the attempt to understand them. Some representative examples follow:

- Why do movie theaters offer discount tickets to students?
- Why do we often see convenience stores located on adjacent street corners?
- Why do supermarket checkout lines all tend to be roughly the same length?

Economic Naturalist Video Series: We are very excited to offer an expanded video series based on Economic Naturalist examples. A series of videos covering some of our favorite micro- and macro-focused examples can be used as part of classroom presentations, or assigned for homework along with accompanying questions within McGraw Hill Connect®. These fascinating, fun, and thought-provoking applications of economics in everyday life encourage students to think like an economist. Refer to the distinguishing features pages of the preface for additional information. You can view one of these dynamic videos here: <http://econeveryday.com/why-do-cooked-rotisserie-chickens-cost-less-than-fresh-uncooked-chickens/>



Active Learning Stressed

The only way to learn to hit an overhead smash in tennis is through repeated practice. The same is true for learning economics. Accordingly, we consistently introduce new ideas in the context of simple examples and then follow them with applications showing how they work in familiar settings. At frequent intervals, we pose self-tests that both test and reinforce the understanding of these ideas. The end-of-chapter questions and problems are carefully crafted to help students internalize and extend basic concepts, and are available within Connect as assignable content so that instructors can require students to engage with this material. Experience with earlier editions confirms that this approach really does prepare students to apply basic economic ideas to solve economic puzzles drawn from the real world.

Both the Economic Naturalist and Learning Glass videos and accompanying multiple-choice questions that test students' understanding of the principles illustrated in the videos have become valued tools for instructors who incorporate elements of the flipped-classroom approach in their teaching, or those who are relying more heavily on other forms of remote learning. Our less-is-more approach to topic coverage is also uniquely well suited to these new instructional approaches.

Modern Microeconomics

- *The cost-benefit principle, which tells us to take only those actions whose benefits exceed their costs, is the core idea behind the economic way of thinking.* Introduced in Chapter 1 and employed repeatedly thereafter, this principle is more fully developed here than in any other text. It underlies the argument for economic efficiency as an important social goal. Rather than speak of trade-offs between efficiency and other goals, we stress that maximizing economic surplus—that is, taking those actions whose benefits exceed their costs—facilitates the achievement of every goal we care about.
- One of the biggest hurdles to the fruitful application of cost-benefit thinking is to recognize and measure the relevant costs and benefits. Common decision pitfalls identified by 2002 Nobel Laureate Daniel Kahneman and others—such as the tendency to ignore implicit costs, the tendency not to ignore sunk costs, and the tendency to confuse average and marginal costs and benefits—are introduced early in Chapter 1 and invoked repeatedly in subsequent chapters.

- There is perhaps no more exciting toolkit for the economic naturalist than a few principles of elementary game theory. In Chapter 7, *Games and Strategic Behavior*, we show how these principles enable students to answer a variety of strategic questions that arise in the marketplace and everyday life. In new Chapter 8, *An Introduction to Behavioral Economics*, we survey many of the most exciting developments in what has become the economics profession's most vibrant new field. We believe that the insights of the Nobel Laureate Ronald Coase are indispensable for understanding a host of familiar laws, customs, and social norms. In Chapter 9, *Externalities and Property Rights*, we show how such devices function to minimize misallocations that result from externalities.

CHANGES IN THE FOURTH EDITION

Changes Common to All Chapters

In all chapters, the narrative has been tightened. Many of the examples have been updated, with a focus on student-centered examples that connect to current topics such as the COVID-19 pandemic. The examples, self-tests, and the end-of-chapter material from the previous edition have been redesigned to provide more clarity and ease of use. Data have been updated throughout.

Chapter-by-Chapter Changes

Chapter 1

- Updated student-centered examples, such as Netflix, wireless keyboards, dogwalking, and Jeff Bezos
- New and updated end-of-chapter problems that reinforce the chapter's learning objectives
- Updated appendix on working with equations, graphs, and tables based on electric scooter rentals

Chapter 2

- Updated student-centered examples, such as digital versus print ads and Marvel Studio films
- New Economic Naturalist, "Why was there a shortage of toilet paper during the COVID-19 pandemic?"
- Three new end-of-chapter questions that reinforce the chapter's learning objectives, including a question related to the drop in crude oil prices during the COVID-19 pandemic

Chapter 3

- Minor updates only
- Updated student-centered examples, such as LeBron James
- New Economic Naturalist, “Why would Jeff Bezos live in a smaller house in Manhattan than in Medina, Washington?”

Chapter 4

- Minor updates only

Chapter 5

- Minor updates only

Chapter 6

- Updated student-centered examples, such as Instagram, electric scooter rentals, iTunes, HBO, Netflix, and cable Internet
- Updated end-of-chapter problems

Chapter 7

- Updated student-centered examples, such as the Ford Mustang and Chevrolet Camaro

Chapter 8

- New to this edition, this chapter serves as an introduction to behavioral economics for those who wish to incorporate this thought-provoking material

Chapter 9

- Updated student-centered examples, such as roommate conflicts
- Updated end-of-chapter questions

Chapter 10

- Updated information on carbon taxes, including mention of the Paris Agreement
- Updated material on welfare payments, in-kind transfers, and the negative income tax

Chapter 11

- Updated student-centered examples, such as interior designer Kelly Wearstler
- Revised Economic Naturalist that discusses the U.S.-China trade war that started in 2018, highlighting that there is more to trade than the exchange of

goods and services and its supply and demand analysis in this chapter; also covers issues such as intellectual property and national security

ORGANIZED LEARNING IN THE FOURTH EDITION

Chapter Learning Objectives

Students and professors can be confident that the organization of each chapter surrounds common themes outlined by four to seven learning objectives listed on the first page of each chapter. These objectives, along with AACSB and Bloom’s Taxonomy Learning Categories, are connected to all test bank questions and end-of-chapter material to offer a comprehensive, thorough teaching and learning experience. Reports available within Connect allow instructors to easily output data related to student performance across chapter learning objectives, AACSB criteria, and Bloom’s Taxonomy Learning Categories.

Assurance of Learning Ready

Many educational institutions today are focused on the notion of assurance of learning, an important element of some accreditation standards. *Principles of Microeconomics, A Streamlined Approach, 4/e*, is designed specifically to support your assurance of learning initiatives with a simple, yet powerful, solution. Instructors can use Connect to easily query for learning objectives that directly relate to the objectives of the course and then use the reporting features of Connect to aggregate student results in a similar fashion, making the collection and presentation of assurance of learning data simple and easy.

AACSB Statement

The McGraw Hill Companies is a proud corporate member of AACSB International. Recognizing the importance and value of AACSB accreditation, the authors of *Principles of Microeconomics, A Streamlined Approach, 4/e*, have sought to recognize the curricula guidelines detailed in AACSB standards for business accreditation by connecting questions in the test bank and end-of-chapter material to the general knowledge and skill guidelines found in AACSB standards.

It is important to note that the statements contained in *Principles of Microeconomics, A Streamlined Approach, 4/e*, are provided only as a guide for the users of this text.



ACKNOWLEDGMENTS

Our thanks first and foremost go to our portfolio director, Anke Weekes, and our senior product developer, Christina Kouvelis. Anke encouraged us to think deeply about how to improve the book and helped us transform our ideas into concrete changes. Christina shepherded us through the revision process with intelligence, sound advice, and good humor. We are grateful as well to the production team, whose professionalism (and patience) was outstanding: Christine Vaughan, lead content project manager; Emily Windelborn, content project manager; Matt Diamond, senior designer; and all of those who worked on the production team to turn our manuscript into the book you see now. Finally, we also thank Bobby Pearson, marketing manager, for getting our message into the wider world.

Special thanks to Peggy Dalton Verner, Frostburg State University, for her energy, creativity, and help in refining the assessment material in both the text and Connect; Sukanya Kemp, University of Akron, for her detailed accuracy check of the learning glass and economic naturalist videos; Alvin Angeles and team at the University of California, San Diego, for their efforts in the production and editing of the learning glass videos; and Kevin Bertotti and the team at ITVK for their creativity in transforming Economic Naturalist examples into dynamic and engaging video vignettes.

Finally, our sincere thanks to the following teachers and colleagues, whose thorough reviews and thoughtful suggestions led to innumerable substantive improvements to *Principles of Microeconomics, A Streamlined Approach*, 4/e.

Mark Abajian, *San Diego Mesa College*
 Richard Agesa, *Marshall University*
 Seemi Ahmad, *Dutchess Community College*
 Donald L. Alexander, *Western Michigan University*
 Chris Azevedo, *University of Central Missouri*
 Narine Badasyan, *Murray State University*
 Sigridur Benediktsdottir, *Yale University*
 Robert Blewett, *St. Lawrence University*
 Brian C. Brush, *Marquette University*
 Colleen Callahan, *American University*
 Giuliana Campanelli Andreopoulos, *William Paterson University*
 J. Lon Carlson, *Illinois State University*
 David Chaplin, *Northwest Nazarene University*
 Joni Charles, *Texas State University*

Anoshua Chaudhuri, *San Francisco State University*
 Nan-Ting Chou, *University of Louisville*
 Maria Luisa Corton, *University of South Florida–St. Petersburg*
 Manabendra Dasgupta, *University of Alabama at Birmingham*
 Craig Dorsey, *College of DuPage*
 Dennis Edwards, *Coastal Carolina University*
 Tracie Edwards, *University of Missouri–St. Louis*
 Roger Frantz, *San Diego State University*
 Mark Frascatore, *Clarkson University*
 Amanda Freeman, *Kansas State University*
 Greg George, *Macon State College*
 Seth Gershenson, *Michigan State University*
 Amy D. Gibson, *Christopher Newport University*
 Rajeev Goel, *Illinois State University*
 Mehdi Haririan, *Bloomsburg University of Pennsylvania*
 Susan He, *Washington State University*
 John Hejkal, *University of Iowa*
 Kuang-Chung Hsu, *Kishwaukee College*
 Greg Hunter, *California State University–Pomona*
 David W. Johnson, *University of Wisconsin–Madison*
 Derek Johnson, *University of Connecticut*
 Sukanya Kemp, *University of Akron*
 Brian Kench, *University of Tampa*
 Fredric R. Kolb, *University of Wisconsin–Eau Claire*
 Daniel D. Kuester, *Kansas State University*
 Valerie Lacarte, *American University*
 Donald J. Liu, *University of Minnesota–Twin Cities*
 Ida Mirzaie, *The Ohio State University*
 Jelena Nikolic, *Northeastern University*
 Anthony A. Noce, *State University of New York (SUNY)–Plattsburgh*
 Diego Nocetti, *Clarkson University*
 Stephanie Owings, *Fort Lewis College*
 Dishant Pandya, *Spalding University*
 Martin Pereyra, *University of Missouri*
 Tony Pizelo, *Northwest University*
 Ratha Ramoo, *Diablo Valley College*
 Thomas Rhoads, *Towson University*
 Bill Robinson, *University of Nevada–Las Vegas*
 Brian Rosario, *University of California–Davis*
 Elyce Rotella, *Indiana University*

Jeffrey Rubin, *Rutgers University*

Naveen Sarna, *Northern Virginia Community College*

Sumati Srinivas, *Radford University*

Thomas Stevens, *University of Massachusetts*

Carolyn Fabian Stumph, *Indiana University and Purdue University–Fort Wayne*

Markland Tuttle, *Sam Houston State University*

David Vera, *California State University–Fresno*

Nancy Virts, *California State University–Northridge*

Gilbert J. Werema, *Texas Woman’s University*

Elizabeth Wheaton, *Southern Methodist University*

William C. Wood, *James Madison University*

DISTINGUISHING FEATURES

Economic Naturalist Examples and Videos

Each Economic Naturalist example starts with a question to spark curiosity and interest in learning an answer. These examples fuel interest while teaching students to see economics in the world around them. Videos of select and new Economic Naturalist examples are denoted in the margin of the material to which they pertain. They are housed within Connect with accompanying questions. A full list of Economic Naturalist examples and videos can be found in the following pages.

Numbered Examples

Throughout the text, numbered and titled examples are referenced and called out to further illustrate concepts. Our engaging questions and examples from everyday life highlight how each human action is the result of an implicit or explicit cost-benefit calculation.

SELF-TESTS

These self-test questions in the body of the chapter enable students to determine whether the preceding material has been understood and reinforce understanding before reading further. Detailed answers to the self-test questions are found at the end of each chapter.

The Economic Naturalist 11.1

What is the China trade shock?

The China trade shock, a term most commonly associated with economists David Autor, David Dorn, and Gordon Hanson, is used to describe the dramatic change in international trade patterns that resulted from China's rise as a major player in the global economy over the past few decades.

In a series of influential studies, these economists and their collaborators investigated the costs of the shock to U.S. workers. They found that employment has fallen

ECONOMIC NATURALISM

With the rudiments of the cost-benefit framework under your belt, you are now in a position to become an "economic naturalist," someone who uses insights from economics to help make sense of observations from everyday life. People who have studied biology are able to observe and marvel at many details of nature that would otherwise have escaped their notice. For example, on a walk in the woods in early April, the novice may see only trees. In contrast, the biology student notices many different species of trees and understands why some are already in leaf while others still lie dormant. Likewise, the novice may notice that in some animal species males are much larger than females, but the biology student knows that pattern occurs only in species in which males take several mates. Natural selection favors larger males in those species because their greater size helps them prevail in the often bloody contests among males for access to females. In contrast, males tend to be roughly the same size as females in monogamous species, in which there is much less fighting for mates.



EXAMPLE 1.3

Implicit Cost

Should you use your frequent-flyer coupon to fly to Cancun for spring break?

With spring break only a week away, you are still undecided about whether to go to Cancun with a group of classmates at the University of Iowa. The round-trip airfare from Cedar Rapids is \$500, but you have a frequent-flyer coupon you could use for the trip. All other relevant costs for the vacation week at the beach total exactly \$1,000. The most you would be willing to pay for the Cancun vacation is \$1,350. That amount is your benefit of taking the vacation. Your only alternative use for your frequent-flyer coupon is for a trip to Boston the weekend after spring break to attend your brother's wedding. (Your coupon expires shortly thereafter.) If the Cedar Rapids–Boston round-trip airfare is \$400, should you use your frequent-flyer coupon to fly to Cancun for spring break?

The Cost-Benefit Principle tells us that you should go to Cancun if the benefits of the trip exceed its costs. If not for the complication of the frequent-flyer coupon, solving this problem would be a straightforward matter of comparing your benefit from the week at the beach to the sum of all relevant costs. And because your airfare and other costs would add up to \$1,500, or \$150 more than your benefit from the trip, you would not go to Cancun.

But what about the possibility of using your frequent-flyer coupon to make the trip? Using it for that purpose might make the flight to Cancun seem free, suggesting you'd reap an economic surplus of \$350 by making the trip. But doing so also would mean you'd have to fork over \$400 for your airfare to Boston. So the implicit cost of using your coupon to go to Cancun is really \$400. If you use it for that purpose, the trip still ends up being a loser because the cost of the vacation, \$1,400, exceeds the benefit by \$50. In cases like these, you're much more likely to decide sensibly if you ask yourself, "Should I use my frequent-flyer coupon for this trip or save it for an upcoming trip?"



Is your flight to Cancun "free" if you travel on a frequent-flyer coupon?

SELF-TEST 1.5

Should a basketball team's best player take all the team's shots?

A professional basketball team has a new assistant coach. The assistant notices that one player scores on a higher percentage of her shots than other players. Based on this information, the assistant suggests to the head coach that the star player take *all* the shots. That way, the assistant reasons, the team will score more points and win more games.

On hearing this suggestion, the head coach fires her assistant for incompetence. What was wrong with the assistant's idea?

COST-BENEFIT ANALYSIS**RECAP** ↑

Scarcity is a basic fact of economic life. Because of it, having more of one good thing almost always means having less of another. The *Cost-Benefit Principle* holds that an individual (or a firm or a society) should take an action if, and only if, the extra benefit from taking the action is at least as great as the extra cost. The benefit of taking any action minus the cost of taking the action is called the *economic surplus* from that action. Hence, the Cost-Benefit Principle suggests that we take only those actions that create additional economic surplus.

Recap

Sprinkled throughout each chapter are Recap boxes that underscore and summarize the importance of the preceding material and key concept takeaways.

Worked Problem Videos

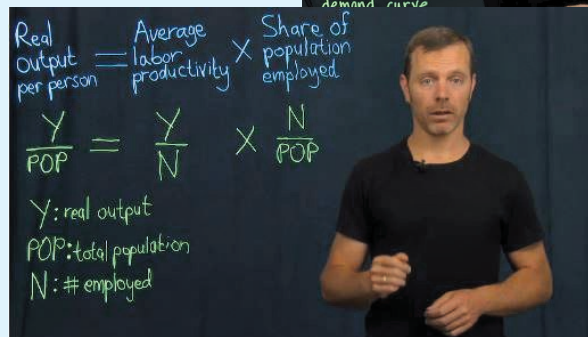
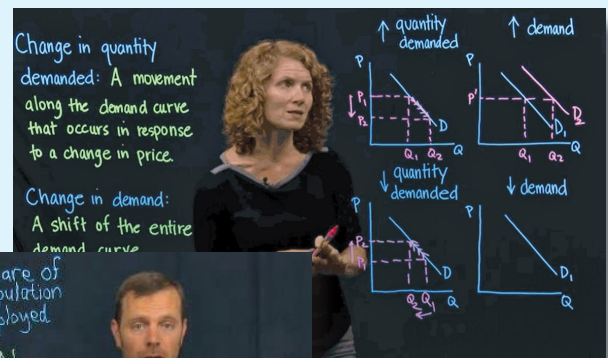
Brief videos work through end-of-chapter problems to aid in student understanding of core economic concepts and offer assistance with more challenging material. The videos are available as hints within Connect.

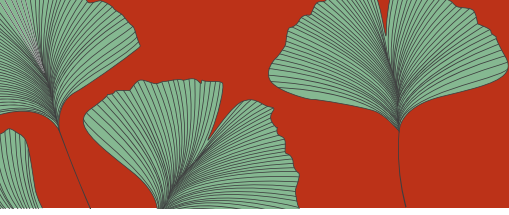
To earn extra money in the summer, you grow tomatoes and sell them at a local farmers' market for 30 cents per pound. By adding compost to your garden, you can increase your yield as shown in the table below. If compost costs 50 cents per pound and your goal is to make as much profit as possible, how many pounds of compost should you add?

Pounds of compost	Pounds of tomatoes	Additional pounds of tomatoes	Additional revenue (or marginal benefit)
	100		
1	120	20	\$6.00 ✓
2	125	5	\$1.50 ✓
3	128		
4	130		
5	131		
6	131.5		

LEARNING GLASS VIDEOS

Dozens of lecture videos featuring authors Kate Antonovics and Ori Heffetz utilize learning glass technology to provide you with an overview of important concepts. These videos can be accessed as resources within SmartBook® or as assignable content with questions via McGraw Hill Connect®.





ECONOMIC NATURALIST VIDEO SERIES

Behavioral Economics: Why do real estate agents often show clients two nearly identical houses, even though one is both cheaper and in better condition than the other?

Comparative Advantage: Why are many products designed in one place yet assembled in another?

Cost Benefit 1: Why does the light come on when you open the refrigerator door but not when you open the freezer?

Cost Benefit 2: Why are child safety seats required in automobiles but not in airplanes?

Discount Pricing: Why might an appliance retailer hammer dents into the sides of its stoves and refrigerators?

Elasticity: Why do people buy the same amount of salt as before even when the price of salt doubles?

Menu Costs: Will new technologies eliminate menu costs?

Monopolistic Competition: Why do we often see convenience stores located on adjacent street corners?

Prisoner's Dilemma: Why do people shout at parties?

Production Costs: Why are brown eggs more expensive than white ones?

Supply and Demand: Why are rotisserie chickens less expensive than fresh chickens?

Tariffs: Why do consumers in the United States often pay more than double the world price for sugar?

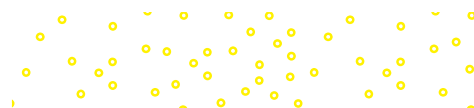
The Demand for Money: Why does the average Argentine citizen hold more U.S. dollars than the average U.S. citizen?

The Invisible Hand: Why do supermarket checkout lines all tend to be roughly the same length?

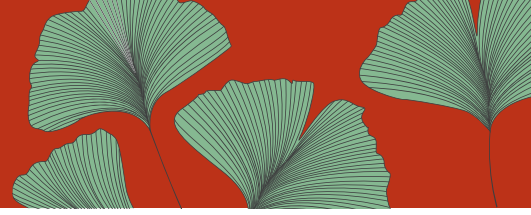
The Law of Demand: Why are smaller automobile engines more common in Europe than in the United States?

The Optimal Amount of Information: Why might a patient be more likely to receive an expensive magnetic resonance imaging (MRI) exam for a sore knee if covered under a conventional health insurance rather than a health maintenance organization (HMO) plan?

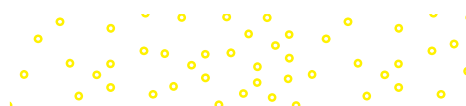
The Tragedy of the Commons and Property Rights: Why do blackberries in public parks get picked before they're completely ripe?



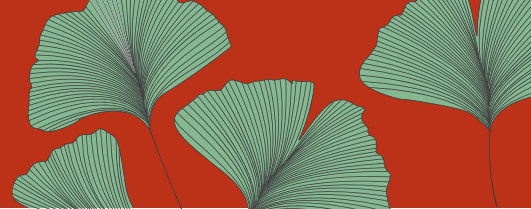
ECONOMIC NATURALIST EXAMPLES



- 1.1 Why do many hardware manufacturers include more than \$1,000 worth of “free” software with a computer selling for only slightly more than that?
- 1.2 Why don’t auto manufacturers make cars without heaters?
- 1.3 Why do the keypad buttons on drive-up automated teller machines have Braille dots?
- 2.1 When the federal government implements a large pay increase for its employees, why do rents for apartments located near Washington Metro stations go up relative to rents for apartments located far away from Metro stations?
- 2.2 Why do major term papers go through so many more revisions today than in the 1970s?
- 2.3 Why do the prices of some goods, like airline tickets to Europe, go up during the months of heaviest consumption, while others, like sweet corn, go down?
- 2.4 Why was there a shortage of toilet paper during the COVID-19 pandemic?
- 3.1 Why does California experience chronic water shortages?
- 3.2 Why would Jeff Bezos live in a smaller house in Manhattan than in Medina, Washington?
- 3.3 Why did people turn to four-cylinder cars in the 1970s, only to shift back to six- and eight-cylinder cars in the 1990s?
- 3.4 Why are automobile engines smaller in England than in the United States?
- 3.5 Why are waiting lines longer in poorer neighborhoods?
- 3.6 Will a higher tax on cigarettes curb teenage smoking?
- 3.7 Why was the luxury tax on yachts such a disaster?
- 4.1 When recycling is left to private market forces, why are many more aluminum beverage containers recycled than glass ones?
- 4.2 Why are gasoline prices so much more volatile than car prices?
- 5.1 Why do supermarket checkout lines all tend to be roughly the same length?
- 5.2 Are there “too many” smart people working as corporate earnings forecasters?
- 6.1 Why does Intel sell the overwhelming majority of all microprocessors used in personal computers?
- 6.2 Why do many movie theaters offer discount tickets to students?
- 6.3 Why might an appliance retailer instruct its clerks to hammer dents into the sides of its stoves and refrigerators?
- 7.1 Why are cartel agreements notoriously unstable?
- 7.2 How did Congress unwittingly solve the television advertising dilemma confronting cigarette producers?
- 7.3 Why do people shout at parties?
- 7.4 Why do we often see convenience stores located on adjacent street corners?
- 8.1 Why did the American Olympic swimmer Shirley Babashoff, who set one world record and six national records at the 1976 Olympics, refuse to appear on the cover of *Sports Illustrated*?
- 8.2 Why would people pay thousands of dollars to attend a weight-loss camp that will feed them only 1,500 calories per day?
- 8.3 Why was Obamacare difficult to enact and is harder still to repeal?
- 8.4 Why have attempts to privatize Social Security proved so politically unpopular in the United States?
- 8.5 If prosperous voters would be happier if they spent less on positional goods and lived in environments with more generously funded public sectors, why haven’t they elected politicians who would deliver what they want?
- 9.1 What is the purpose of free speech laws?
- 9.2 Why does the government subsidize private property owners to plant trees on their hillsides?
- 9.3 Why do blackberries in public parks get picked too soon?
- 9.4 Why are shared milkshakes consumed too quickly?
- 9.5 Why do football players take anabolic steroids?
- 10.1 Why is a patient with a sore knee more likely to receive an MRI exam if he has conventional health insurance than if he belongs to a health maintenance organization?
- 11.1 What is the China trade shock?
- 11.2 Why did the U.S. start a trade war with China?
- 11.3 What is fast track authority?



SUPPLEMENTS



The following ancillaries are available for quick download and convenient access via the Instructor Resource material available through McGraw Hill Connect®.

Solutions Manual

Prepared by the authors with assistance from Peggy Dalton Verner, Frostburg State University, this manual provides detailed answers to the end-of-chapter review questions and problems.

Test Bank

The test bank has been carefully revised and reviewed for accuracy. Hundreds of questions have been categorized by chapter learning objectives, AACSB learning categories, Bloom's Taxonomy objectives, and level of difficulty. Select problems are now offered as an algorithmic alternative, providing even more variation.

Test Builder in Connect

Available within Connect, Test Builder is a cloud-based tool that enables instructors to format tests that can be printed or administered within an LMS. Test Builder offers a modern, streamlined interface for easy content configuration that matches course needs, without requiring a download.

Test Builder allows you to:

- access all test bank content from a particular title.
- easily pinpoint the most relevant content through robust filtering options.
- manipulate the order of questions or scramble questions and/or answers.
- pin questions to a specific location within a test.
- determine your preferred treatment of algorithmic questions.
- choose the layout and spacing.
- add instructions and configure default settings.

Test Builder provides a secure interface for better protection of content and allows for just-in-time updates to flow directly into assessments.

PowerPoints

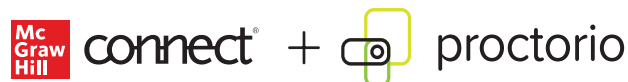
Presentation slides contain a detailed, chapter-by-chapter review of the important ideas presented in the

textbook, accompanied by animated graphs and slide notes. You can edit, print, or rearrange the slides to fit the needs of your course.

Customizable Micro Lecture Notes and PowerPoints

One of the biggest hurdles to an instructor considering changing textbooks is the prospect of having to prepare new lecture notes and slides. For the microeconomics chapters, this hurdle no longer exists. A full set of lecture notes for principles of microeconomics, prepared by Bob Frank for his award-winning introductory microeconomics course at Cornell University, is available as Microsoft Word files that instructors are welcome to customize as they see fit. The challenge for any instructor is to reinforce the lessons of the text in lectures without generating student unrest by merely repeating what's in the book. These lecture notes address that challenge by constructing examples that run parallel to those presented in the book, yet are different from them in interesting contextual ways. Also available is a complete set of richly illustrated PowerPoint files to accompany these lecture notes. Instructors are also welcome to customize these files as they wish.

Remote Proctoring and Browser-Locking Capabilities



New remote proctoring and browser-locking capabilities, hosted by Proctorio within Connect, provide control of the assessment environment by enabling security options and verifying the identity of the student.

Seamlessly integrated within Connect, these services allow instructors to control students' assessment experience by restricting browser activity, recording students' activity, and verifying students are doing their own work.

Instant and detailed reporting gives instructors an at-a-glance view of potential academic integrity concerns, thereby avoiding personal bias and supporting evidence-based claims.

FOR MORE INFORMATION ABOUT CONNECT AND ITS AVAILABLE RESOURCES, REFER TO THE PAGES THAT FOLLOW.



Connect Economics Asset Alignment with Bloom's Taxonomy

Principles of Microeconomics: A Streamlined Approach, 4e

We Take Students Higher

As a learning science company we create content that supports higher order thinking skills. Within Connect®, we tag assessments accordingly so you can filter your search, assign it, and receive reporting on it. These content asset types can be associated with one or more levels of Bloom's Taxonomy.

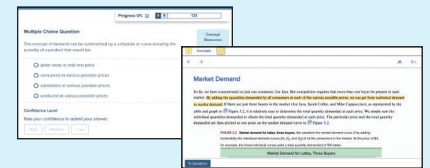
The chart below shows a few of the key assignable economics assets with *McGraw Hill Connect* aligned with Bloom's Taxonomy. Take your students higher by assigning a variety of applications, moving them from simple memorization to concept application.

		SmartBook 2.0	ECON Adaptive Math Prep	Videos	Exercises	Interactive Graphs	Application-Based Activities	ECON Everyday Current Events Blog*	Writing Assignment Plus
Higher Order Thinking Skills	CREATE								✓
	EVALUATE						✓	✓	✓
	ANALYZE				✓	✓	✓	✓	✓
Lower Order Thinking Skills	APPLY		✓	✓	✓	✓	✓	✓	✓
	UNDERSTAND	✓	✓	✓	✓	✓	✓	✓	✓
	REMEMBER	✓	✓	✓	✓	✓	✓	✓	✓

* Outside of Connect.

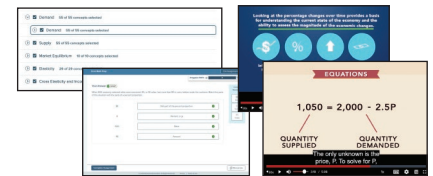
SmartBook 2.0

Adaptively aids students to study more efficiently by highlighting where in the chapter to focus, asking review questions and pointing them to passages in the text until they understand. Assignable and assessable.



Adaptive Econ Prep

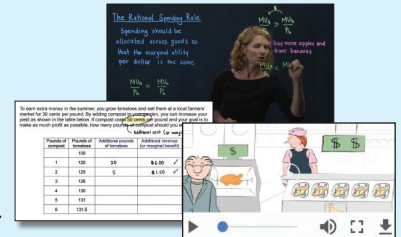
Math and graphing preparedness assignments help students refresh important prerequisite topics necessary to be successful in economics. New Adaptive Econ Prep Tool provides students just-in-time math remediation that is prerequisite to success in Principles of Economics courses and adapts to each student.



Videos

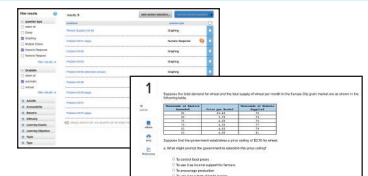
Worked examples and real-world application videos help students learn economics.

Learning Glass videos reinforce challenging topics featuring the authors and innovative learning glass technology. **Economic Naturalist videos** bring examples to life showing interesting applications of economic concepts. **Worked Problem videos** work through select end-of-chapter questions for extra help and guidance through challenging material.



Exercises

Exercises with algorithmic variations provide ample opportunities for students to practice and hone quantitative skills. Graphing Exercises provide opportunities for students to draw, interact with, manipulate, and analyze graphs.



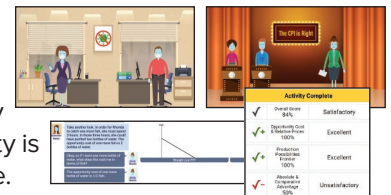
Interactive Graphs

Interactive Graphs provide visual displays of real data and economic concepts for students to manipulate. All graphs are accompanied by assignable assessment questions and feedback to guide students through the experience of learning to read and interpret graphs and data.



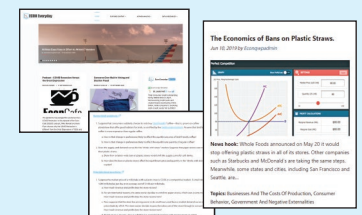
Application-Based Activities

Immersive real-life scenarios engage students and put them in the role of everyday economists. Students practice their economic thinking and problem-solving skills as they apply course concepts and see the implications of their decisions as they go. Each activity is designed as a 15-minute experience, unless students eagerly replay for a better outcome.



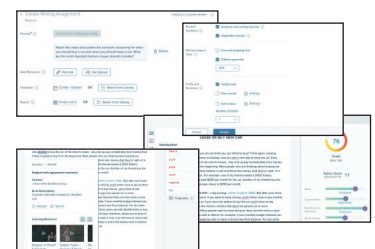
ECON Everyday Blog*

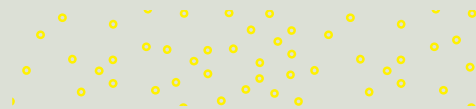
Our ECON Everyday blog saves instructors time bringing current, student-centered content into their course all semester long. Short articles, written for principles-level students, are tagged by topic to bring currency into your course. We also provide discussion questions to help you drive the conversation forward. Visit www.econeveryday.com and subscribe for updates. (*Outside of Connect.)



Writing Assignment Plus

Writing Assignment Plus delivers a learning experience that helps students improve their written communication skills and conceptual understanding. Faculty can assign, monitor, grade, and provide feedback on writing projects efficiently. Built-in grammar and writing review helps students improve writing quality while an originality check helps students correct central plagiarism before submission. End result? Improved workplace skills of writing in critical thinking.



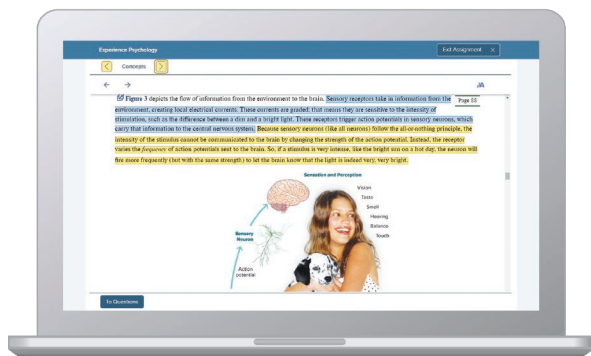


Instructors: Student Success Starts with You

Tools to enhance your unique voice

Want to build your own course? No problem. Prefer to use our turnkey, prebuilt course? Easy. Want to make changes throughout the semester? Sure. And you'll save time with Connect's auto-grading too.

65%
Less Time
Grading



Laptop: McGraw Hill; Woman/dog: George Doyle/Getty Images

Study made personal

Incorporate adaptive study resources like SmartBook® 2.0 into your course and help your students be better prepared in less time. Learn more about the powerful personalized learning experience available in SmartBook 2.0 at www.mheducation.com/highered/connect/smartbook

Affordable solutions, added value



Make technology work for you with LMS integration for single sign-on access, mobile access to the digital textbook, and reports to quickly show you how each of your students is doing. And with our Inclusive Access program you can provide all these tools at a discount to your students. Ask your McGraw Hill representative for more information.

Padlock: Jobalou/Getty Images

Solutions for your challenges



A product isn't a solution. Real solutions are affordable, reliable, and come with training and ongoing support when you need it and how you want it. Visit **www.supportateverystep.com** for videos and resources both you and your students can use throughout the semester.

Checkmark: Jobalou/Getty Images



Students: Get Learning That Fits You

Effective tools for efficient studying

Connect is designed to make you more productive with simple, flexible, intuitive tools that maximize your study time and meet your individual learning needs. Get learning that works for you with Connect.

Study anytime, anywhere

Download the free ReadAnywhere app and access your online eBook or SmartBook 2.0 assignments when it's convenient, even if you're offline. And since the app automatically syncs with your eBook and SmartBook 2.0 assignments in Connect, all of your work is available every time you open it. Find out more at www.mheducation.com/readanywhere

"I really liked this app—it made it easy to study when you don't have your textbook in front of you."

- Jordan Cunningham,
Eastern Washington University



Calendar: owattaphotos/Getty Images

Everything you need in one place

Your Connect course has everything you need—whether reading on your digital eBook or completing assignments for class, Connect makes it easy to get your work done.

Learning for everyone

McGraw Hill works directly with Accessibility Services Departments and faculty to meet the learning needs of all students. Please contact your Accessibility Services Office and ask them to email accessibility@mheducation.com, or visit www.mheducation.com/about/accessibility for more information.

Top: Jenner Images/Getty Images, Left: Hero Images/Getty Images, Right: Hero Images/Getty Images



COMPARISON GUIDE FOR FRANK, BERNANKE, ANTONOVICS, AND HEFFETZ PRODUCTS

Principles of Economics provides enhanced coverage and offers more topics and mathematical rigor. *Principles of Economics: A Streamlined Approach* is a stripped-down version of the big book featuring core content with a less-is-more approach. See which product is right for you!

Comparison Guide							
Principles of Economics, 8th edition				Principles of Economics: A Streamlined Approach, 4th edition			
Chapter Title	Econ 8e	Micro 8e	Macro 8e	Chapter Title	Streamlined 4e Econ	Streamlined 4e Micro	Streamlined 4e Macro
Thinking Like an Economist	1	1	1	Thinking Like an Economist	1	1	1
Comparative Advantage	2	2	2				
Supply and Demand	3	3	3	Supply and Demand	2	2	2
Elasticity	4	4		Demand and Elasticity	3	3	
Demand	5	5					
Perfectly Competitive Supply	6	6		Perfectly Competitive Supply	4	4	
Efficiency, Exchange, and the Invisible Hand in Action	7	7		Efficiency, Exchange, and the Invisible Hand in Action	5	5	
Monopoly, Oligopoly, and Monopolistic Competition	8	8		Monopoly, Oligopoly, and Monopolistic Competition	6	6	
Games and Strategic Behavior	9	9		Games and Strategic Behavior	7	7	
An Introduction to Behavioral Economics	10	10		An Introduction to Behavioral Economics (NEW)	8	8	
Externalities, Property Rights, and the Environment	11	11		Externalities and Property Rights	9	9	
The Economics of Information	12	12		Using Economics to Make Better Policy Choices	10	10	
Labor Markets, Poverty, and Income Distribution	13	13					
Public Goods and Tax Policy	14	14					
International Trade and Trade Policy	15	15	16	International Trade and Trade Policy	11	11	12
Macroeconomics: The Bird's-Eye View of the Economy	16		4	Macroeconomics: The Bird's Eye View of the Economy	12		3
Measuring Economic Activity: GDP and Unemployment	17		5	Measuring Economic Activity: GDP, Unemployment, and Inflation	13		4
Measuring the Price Level and Inflation	18		6				
Economic Growth, Productivity, and Living Standards	19		7	Economic Growth, Productivity, and Living Standards	14		5
The Labor Market: Workers, Wages, and Unemployment	20		8	The Labor Market: Workers, Wages, and Unemployment	15		6
Saving and Capital Formation	21		9	Saving and Capital Formation	16		7
Money, Prices, and the Federal Reserve	22		10	Money, the Federal Reserve, and Global Financial Markets	17		8
Financial Markets and International Capital Flows	23		11				
Short-Term Economic Fluctuations: An Introduction	24		12	Short-Term Economic Fluctuations and Fiscal Policy	18		9
Spending and Output in the Short Run	25		13				
Stabilizing the Economy: The Role of the Fed	26		14	Stabilizing the Economy: The Role of the Fed	19		10
Aggregate Demand, Aggregate Supply, and Inflation	27		15	Aggregate Demand, Aggregate Supply, and Inflation	20		11
Exchange Rates and the Open Economy	28		17	Exchange Rates and the Open Economy	21		13

PART 1 Introduction

- 1** Thinking Like an Economist 1
- 2** Supply and Demand 31

PART 2 Competition and the Invisible Hand

- 3** Demand and Elasticity 65
- 4** Perfectly Competitive Supply 91
- 5** Efficiency, Exchange, and the Invisible Hand in Action 119

PART 3 Market Imperfections

- 6** Monopoly, Oligopoly, and Monopolistic Competition 141
- 7** Games and Strategic Behavior 169
- 8** An Introduction to Behavioral Economics 195
- 9** Externalities and Property Rights 217

PART 4 Economics of Public Policy

- 10** Using Economics to Make Better Policy Choices 241

PART 5 International Trade

- 11** International Trade and Trade Policy 261

CONTENTS

PART 1 Introduction

Chapter 1 Thinking Like an Economist 1

Economics: Studying Choice in a World of Scarcity 2

Applying the Cost-Benefit Principle 3

Economic Surplus 4

Opportunity Cost 4

The Role of Economic Models 5

Three Important Decision Pitfalls 6

Pitfall 1: Measuring Costs and Benefits as Proportions

Rather Than Absolute Dollar Amounts 6

Pitfall 2: Ignoring Implicit Costs 7

Pitfall 3: Failing to Think at the Margin 8

Normative Economics versus Positive Economics 13

Economics: Micro and Macro 13

The Approach of This Text 14

Economic Naturalism 14

THE ECONOMIC NATURALIST 1.1 15

THE ECONOMIC NATURALIST 1.2 15

THE ECONOMIC NATURALIST 1.3 16

Summary 17 • *Key Terms* 17 • *Review Questions* 17 •
Problems 18 • *Answers to Self-Tests* 19 • *Appendix: Working
with Equations, Graphs, and Tables* 21

Chapter 2 Supply and Demand 31

What, How, and for Whom? Central Planning
versus the Market 33

Buyers and Sellers in Markets 34

The Demand Curve 35

The Supply Curve 36

Market Equilibrium 38

Rent Controls Reconsidered 41

Pizza Price Controls? 43

Predicting and Explaining Changes in Prices
and Quantities 44

Shifts in Demand 45

THE ECONOMIC NATURALIST 2.1 47

Shifts in the Supply Curve 48

THE ECONOMIC NATURALIST 2.2 51

Four Simple Rules 51

THE ECONOMIC NATURALIST 2.3 54

Efficiency and Equilibrium 54

Cash on the Table 55

Smart for One, Dumb for All 56

THE ECONOMIC NATURALIST 2.4 57

Summary 58 • *Key Terms* 59 • *Review Questions* 59 •
Problems 59 • *Answers to Self-Tests* 60 • *Appendix: The
Algebra of Supply and Demand* 63

PART 2 Competition and the Invisible Hand

Chapter 3 Demand and Elasticity 65

The Law of Demand 66

The Origins of Demand 66

Needs versus Wants 67

THE ECONOMIC NATURALIST 3.1 67

Applying the Law of Demand 68

Substitution at Work 68

THE ECONOMIC NATURALIST 3.2 68

THE ECONOMIC NATURALIST 3.3 69

THE ECONOMIC NATURALIST 3.4 70

The Importance of Income Differences 70

THE ECONOMIC NATURALIST 3.5 71

Individual and Market Demand Curves 71

Horizontal Addition 71

Elasticity 73

Price Elasticity of Demand 73

Price Elasticity Defined 73

Determinants of Price Elasticity of Demand 75

Some Representative Elasticity Estimates 76

Using Price Elasticity of Demand 77

THE ECONOMIC NATURALIST 3.6 77

THE ECONOMIC NATURALIST 3.7 77

A Graphical Interpretation of Price Elasticity 78

Price Elasticity Changes along a Straight-Line

Demand Curve 80

Two Special Cases 81

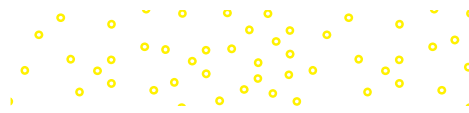
Elasticity and Total Expenditure 82

Income Elasticity and Cross-Price Elasticity
of Demand 86

Summary 87 • *Key Terms* 87 • *Review Questions* 88 •
Problems 88 • *Answers to Self-Tests* 89

Chapter 4 Perfectly Competitive Supply 91

Thinking about Supply: The Importance of
Opportunity Cost 92

**Individual and Market Supply Curves 94****Profit-Maximizing Firms in Perfectly****Competitive Markets 95**

Profit Maximization 95

The Demand Curve Facing a Perfectly

Competitive Firm 96

Production in the Short Run 97

Choosing Output to Maximize Profit 98

Price Equals Marginal Cost: The Seller's Supply
Rule 101

Graphing Marginal Cost 101

The "Law" of Supply 103

Applying the Theory of Supply 104

THE ECONOMIC NATURALIST 4.1 104**Determinants of Supply Revisited 107**

Technology 107

Input Prices 107

The Number of Suppliers 107

Expectations 108

Changes in Prices of Other Products 108

The Price Elasticity of Supply 108

Determinants of Supply Elasticity 111

THE ECONOMIC NATURALIST 4.2 112Unique and Essential Inputs: The Ultimate Supply
Bottleneck 114*Summary 115 • Key Terms 116 • Review Questions 116 •
Problems 116 • Answers to Self-Tests 117***Chapter 5 Efficiency, Exchange, and the
Invisible Hand in Action 119****The Central Role of Economic Profit 120**

Three Types of Profit 120

The Invisible Hand Theory 123

Two Functions of Price 123

Responses to Profits and Losses 123

The Effect of Market Forces on Economic
Profit 125The Importance of Free Entry
and Exit 126

The Invisible Hand in Action 127

THE ECONOMIC NATURALIST 5.1 127**Economic Rent versus Economic Profit 129****The Distinction between an Equilibrium and a
Social Optimum 130**

Smart for One, Dumb for All 131

THE ECONOMIC NATURALIST 5.2 131

Market Equilibrium and Efficiency 132

The Cost of Preventing Price Adjustments 136*Summary 138 • Key Terms 139 •**Review Questions 139 • Problems 139 •**Answers to Self-Tests 140***PART 3 Market Imperfections****Chapter 6 Monopoly, Oligopoly, and
Monopolistic Competition 141****Perfect and Imperfect Competition 142**

Different Forms of Imperfect Competition 142

The Essential Difference between Perfectly and
Imperfectly Competitive Firms 144**Five Sources of Market Power 145**

Exclusive Control over Important Inputs 145

Patents and Copyrights 145

Government Licenses or Franchises 145

Economies of Scale and Natural Monopolies 146

Network Economies 146

Economies of Scale and the Importance of Start-Up Costs 147**THE ECONOMIC NATURALIST 6.1 150****Profit Maximization for the Monopolist 150**

Marginal Revenue for the Monopolist 150

The Monopolist's Profit-Maximizing Decision Rule 152

Being a Monopolist Doesn't Guarantee
an Economic Profit 154**Why the Invisible Hand Breaks Down under
Monopoly 155****Using Discounts to Expand the Market 156**

Price Discrimination Defined 157

THE ECONOMIC NATURALIST 6.2 157

How Price Discrimination Affects Output 157

The Hurdle Method of Price Discrimination 160

Is Price Discrimination a Bad Thing? 163

Examples of Price Discrimination 163

THE ECONOMIC NATURALIST 6.3 164*Summary 165 • Key Terms 166 •**Review Questions 166 • Problems 166 •**Answers to Self-Tests 168***Chapter 7 Games and Strategic Behavior 169****Using Game Theory to Analyze Strategic Decisions 170**

The Three Elements of a Game 170

Nash Equilibrium 172

The Prisoner's Dilemma 174

The Original Prisoner's Dilemma 174

The Economics of Cartels 175

THE ECONOMIC NATURALIST 7.1 175

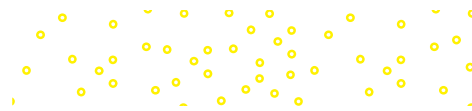
Tit-for-Tat and the Repeated Prisoner's Dilemma 177

THE ECONOMIC NATURALIST 7.2 178**THE ECONOMIC NATURALIST 7.3 179****Games in Which Timing Matters 180**

Credible Threats and Promises 182

Monopolistic Competition When Location Matters 184

THE ECONOMIC NATURALIST 7.4 184

**Commitment Problems** 186

Solving Commitment Problems with Psychological
Incentives 188

Summary 190 • *Key Terms* 190 • *Review Questions* 191 •
Problems 191 • *Answers to Self-Tests* 194

**Chapter 8 An Introduction to Behavioral
Economics** 195**Judgmental Heuristics or Rules of Thumb** 196

Availability 196
Representativeness 197
Regression to the Mean 198

THE ECONOMIC NATURALIST 8.1 198

Anchoring and Adjustment 199

Misinterpretation of Contextual Clues 200

The Psychophysics of Perception 200
The Difficulty of Actually Deciding 200

Impulse-Control Problems 202**THE ECONOMIC NATURALIST 8.2** 203**Loss Aversion and Status Quo Bias** 205**THE ECONOMIC NATURALIST 8.3** 206**Concerns about Relative Position** 207**THE ECONOMIC NATURALIST 8.4** 210**THE ECONOMIC NATURALIST 8.5** 211

Summary 214 • *Key Terms* 215 • *Review Questions* 215 •
Problems 215 • *Answers to Self-Tests* 216

Chapter 9 Externalities and Property Rights 217**External Costs and Benefits** 217

How Externalities Affect Resource Allocation 218

The Coase Theorem 219**Remedies for Externalities** 223

Laws and Regulations 223

THE ECONOMIC NATURALIST 9.1 225

The Optimal Amount of Negative
Externalities Is Not Zero 225
Compensatory Taxes and Subsidies 226

THE ECONOMIC NATURALIST 9.2 226**Property Rights and the Tragedy of the Commons** 227

The Problem of Unpriced Resources 227
The Effect of Private Ownership 229
When Private Ownership Is Impractical 230

THE ECONOMIC NATURALIST 9.3 230**THE ECONOMIC NATURALIST 9.4** 231**Positional Externalities** 232

Payoffs That Depend on Relative Performance 232

THE ECONOMIC NATURALIST 9.5 232

Positional Arms Races and Positional
Arms Control Agreements 234
Social Norms as Positional Arms Control Agreements 235

Summary 237 • *Key Terms* 237 • *Review Questions* 237 •
Problems 238 • *Answers to Self-Tests* 239

PART 4 Economics of Public Policy**Chapter 10 Using Economics to Make Better
Policy Choices** 241**The Economics of Health Care** 242

The Case for Mandatory Immunization Laws 242

Explaining Rising Health Care Costs 243

Designing a Solution 244
The HMO Revolution 245

THE ECONOMIC NATURALIST 10.1 245

The Problem with Health Care Provision
through Private Insurance 246
The Affordable Care Act of 2010 247

Using Price Incentives in Environmental Regulation 248

Taxing Pollution 248
Auctioning Pollution Permits 250
Climate Change and Carbon Taxes 252

Methods of Income Redistribution 253

Welfare Payments and In-Kind Transfers 254
Means-Tested Benefit Programs 254
The Negative Income Tax 255
Minimum Wages 255
The Earned-Income Tax Credit 256
Public Employment for the Poor 256
A Combination of Methods 256

Summary 258 • *Key Terms* 258 • *Review Questions* 258 •
Problems 258 • *Answers to Self-Tests* 259

PART 5 International Trade**Chapter 11 International Trade and Trade
Policy** 261**Comparative Advantage as a Basis for Trade** 262**A Supply and Demand Perspective on Trade** 266

Winners and Losers from Trade 268

THE ECONOMIC NATURALIST 11.1 269**Protectionist Policies: Tariffs and Quotas** 270

Tariffs 270
Quotas 272

THE ECONOMIC NATURALIST 11.2 275

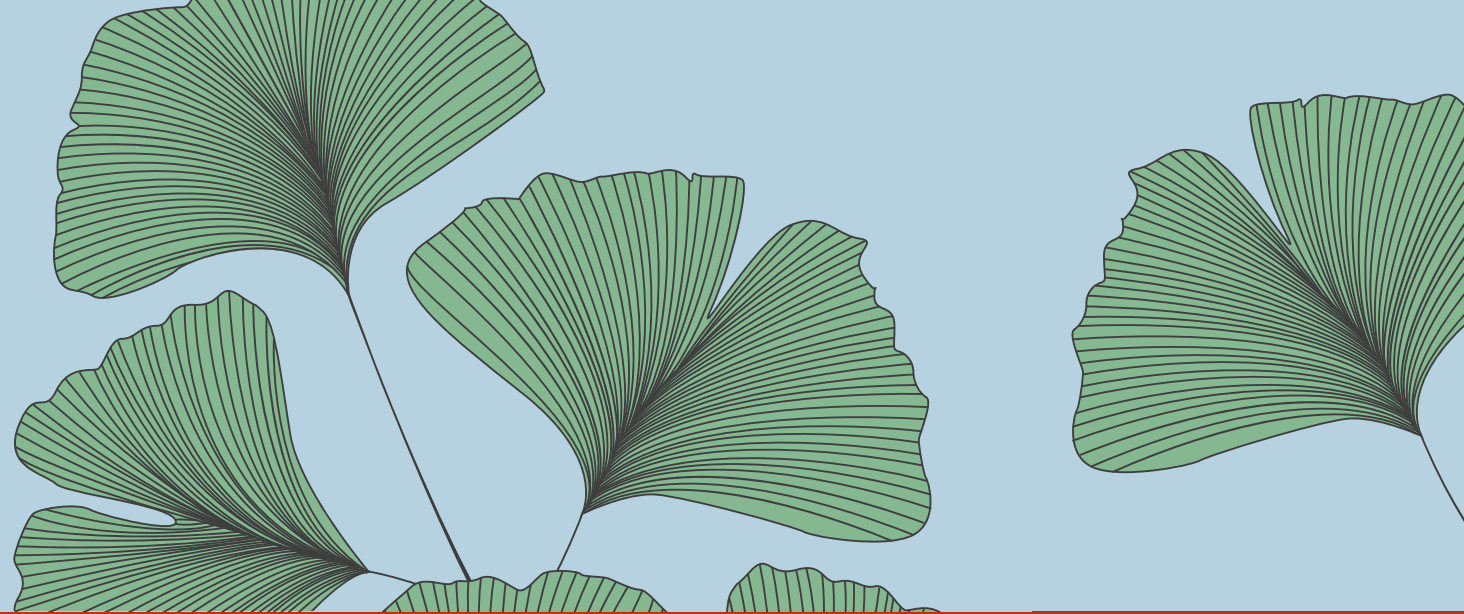
The Inefficiency of Protectionism 276

THE ECONOMIC NATURALIST 11.3 276

Summary 277 • *Key Terms* 278 • *Review Questions* 278 •
Problems 278 • *Answers to Self-Tests* 280

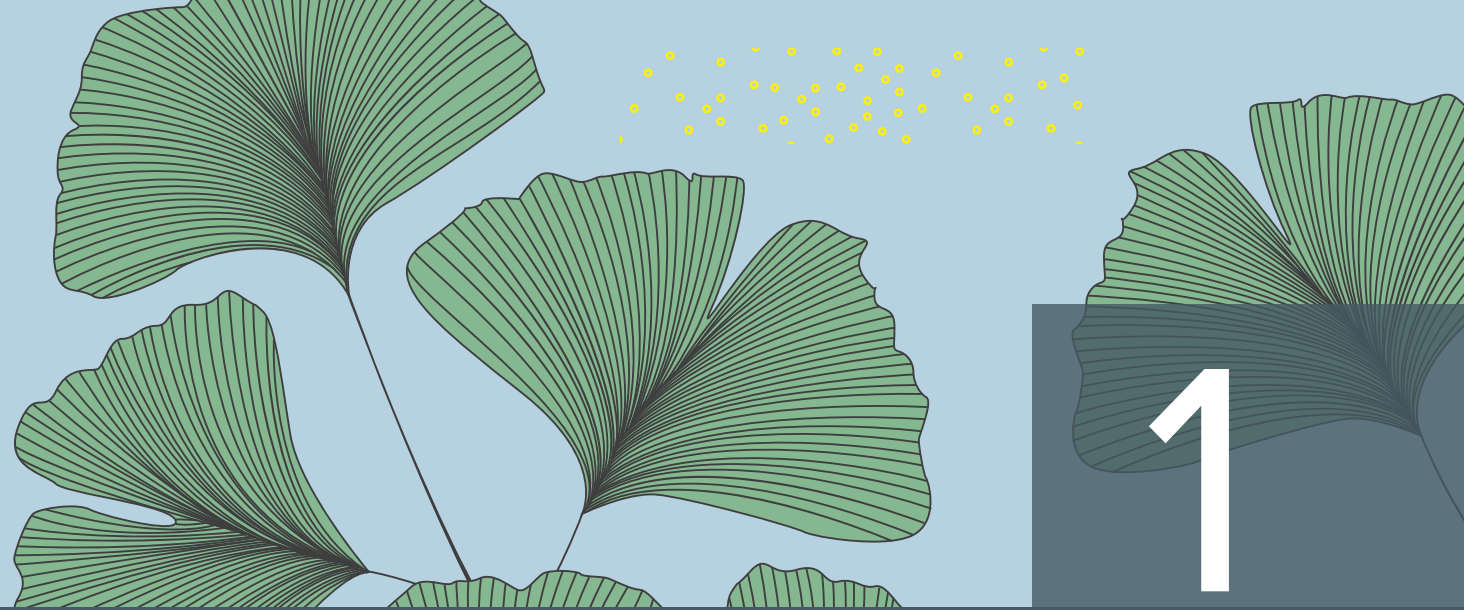
Glossary G-1

Index I-1



FOURTH EDITION

Principles of
MICROECONOMICS
A STREAMLINED APPROACH



1

LEARNING OBJECTIVES

After reading this chapter, you should be able to:

- LO1** Explain why having more of any good thing necessarily requires having less of something else.
- LO2** Explain and apply the Cost-Benefit Principle, which says that an action should be taken if, but only if, its benefit is at least as great as its cost.
- LO3** Discuss three important pitfalls that occur when applying the Cost-Benefit Principle inconsistently.
- LO4** Explain why if you want to predict people's behavior, a good place to start is by examining their incentives.

Thinking Like an Economist

How many students are in your introductory economics class? Some classes have just 20 or so. Others average 35, 100, or 200 students. At some schools, introductory economics classes may have as many as 2,000 students. What size is best?

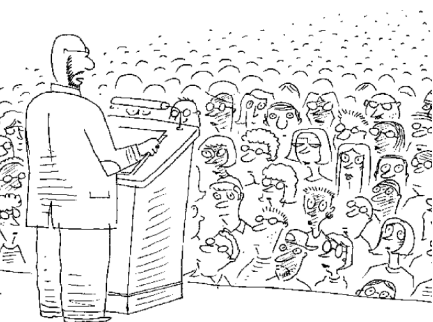
If cost were no object, the best size might be a single student. Think about it: the whole course, all term long, with just you and your professor! Everything could be custom-tailored to your own background and ability. You could cover the material at just the right pace. The tutorial format also would promote close communication and personal trust between you and your professor. And your grade would depend more heavily on what you actually learned than on your luck when taking multiple-choice exams. Let's suppose, for the sake of discussion, that students have been shown to learn best in the tutorial format.

Why, then, do so many introductory classes still have hundreds of students? The simple reason is that costs *do* matter. They matter not just to the university administrators who must build classrooms and pay faculty salaries, but also to *you*. The direct cost of providing you with your own personal introductory economics course might easily top \$50,000. *Someone* has to pay these costs. In private universities, a large share of the cost would be recovered directly from higher tuition payments. In state universities, the burden would be split between higher tuition payments and higher tax payments. But, in either case, the course would be unaffordable for most students.

With larger classes, of course, the cost per student goes down. For example, an introductory economics course with 300 students might cost as little as \$200 per student. But a class that large could easily compromise the quality of the learning environment. Compared to the custom tutorial format, however, it would be dramatically more affordable.

In choosing what size introductory economics course to offer, then, university administrators confront a classic economic trade-off. In making the class larger, they risk lowering the quality of instruction—a bad thing. At the same time, they reduce costs and hence the tuition students must pay—a good thing.





Are small classes “better” than large ones?

economics the study of how people make choices under conditions of scarcity and of the results of those choices for society

In this chapter, we’ll introduce some simple ideas that will help you understand and explain patterns of behavior you observe in the world around you. These principles also will help you avoid three pitfalls that plague decision makers in everyday life.

ECONOMICS: STUDYING CHOICE IN A WORLD OF SCARCITY

Even in rich societies like the United States, *scarcity* is a fundamental fact of life. There is never enough time, money, or energy to do everything we want to do or have everything we’d like to have. **Economics** is the study of how people make choices under conditions of scarcity and of the results of those choices for society.

In the class-size example just discussed, a motivated economics student might definitely prefer to be in a class of 20 rather than a class of 100, everything else being equal. But other things, of course, are not equal. Students can enjoy the benefits of having smaller classes, but only at the price of having less money for other activities. The student’s choice inevitably will come down to the relative importance of competing activities.

That such trade-offs are widespread and important is one of the core principles of economics. Although we have boundless needs and wants, the resources available to us are limited. So having more of one good thing usually means having less of another.

Inherent in the idea of a trade-off is the fact that choice involves compromise between competing interests. Economists resolve such trade-offs by using cost-benefit analysis, which is based on the disarmingly simple principle that an action should be taken if, and only if, its benefits exceed its costs. We call this statement the *Cost-Benefit Principle*, a core principle of economics.

With this principle in mind, let’s think about our class-size question again. Imagine that classrooms come in only two sizes—100-seat lecture halls and 20-seat classrooms—and that your university currently offers introductory economics courses to classes of 100 students. Question: Should administrators reduce the class size to 20 students? Answer: Reduce if, and only if, the value of the improvement in instruction outweighs its additional cost.

This rule sounds simple. But to apply it we need some way to measure the relevant costs and benefits, a task that’s often difficult in practice. If we make a few simplifying assumptions, however, we can see how the analysis might work. On the cost side, the primary expense of reducing class size from 100 to 20 is that we’ll now need five professors instead of just one. We’ll also need five smaller classrooms rather than a single big one, and this too may add slightly to the expense of the move. Let’s suppose that classes with 20 cost \$1,000 per student more than those with 100. Should administrators switch to the smaller class size? If they apply the Cost-Benefit Principle, they will realize that *doing so makes sense only if the value of attending the smaller class is at least \$1,000 per student greater than the value of attending the larger class*.

Would you (or your family) be willing to pay an extra \$1,000 for a smaller class? If not, and if other students feel the same way, then sticking with the larger class size makes sense. But if you and others would be willing to pay the extra tuition, then reducing the class size makes good economic sense.

Notice that the “best” class size, from an economic point of view, will generally not be the same as the “best” size from the point of view of an educational psychologist. That’s because the economic definition of “best” takes into account both the benefits *and* the costs of different class sizes. The psychologist ignores costs and looks only at the learning benefits of different class sizes.

In practice, of course, different people feel differently about the value of smaller classes. People with high incomes, for example, tend to be willing to pay more for the advantage. That helps explain why average class size is smaller, and

tuition higher, at private schools whose students come predominantly from high-income families.

The cost-benefit framework for thinking about the class-size problem also suggests a possible reason for the gradual increase in average class size that has been taking place in American colleges and universities. During the last 30 years, professors' salaries have risen sharply, making smaller classes more costly. During the same period, median family income—and hence the willingness to pay for smaller classes—has remained roughly constant. When the cost of offering smaller classes goes up but willingness to pay for smaller classes does not, universities shift to larger class sizes.

Scarcity and the trade-offs that result also apply to resources other than money. Jeff Bezos is one of the richest people on Earth. His wealth is estimated at more than \$180 billion. That's more than the combined wealth of the poorest 54 percent of Americans. Bezos could buy more houses, cars, vacations, and other consumer goods than he could possibly use. Yet he, like the rest of us, has only 24 hours each day and a limited amount of energy. So even he confronts trade-offs. Any activity he pursues—whether it be building his business empire or redecorating his mansion—uses up time and energy that he could otherwise spend on other things. Indeed, someone once calculated that the value of Bezos's time is so great that pausing to pick up a \$100 bill from the sidewalk simply wouldn't be worth his while.



Chip Somodevilla/Getty Images

If Jeff Bezos saw a \$100 bill lying on the sidewalk, would it be worth his time to pick it up?

APPLYING THE COST-BENEFIT PRINCIPLE

In studying choice under scarcity, we'll usually begin with the premise that people are **rational**, which means they have well-defined goals and try to fulfill them as best they can. The Cost-Benefit Principle is a fundamental tool for the study of how rational people make choices.

As in the class-size example, often the only real difficulty in applying the cost-benefit rule is to come up with reasonable measures of the relevant benefits and costs. Only in rare instances will exact dollar measures be conveniently available. But the cost-benefit framework can lend structure to your thinking even when no relevant market data are available.

To illustrate how we proceed in such cases, the following example asks you to decide whether to perform an action whose cost is described only in vague, qualitative terms.

rational person someone with well-defined goals who tries to fulfill those goals as best he or she can

EXAMPLE 1.1

Comparing Costs and Benefits

Should you walk downtown to save \$10 on a \$25 wireless keyboard?

Imagine you are about to buy a \$25 wireless keyboard at the nearby campus store when a friend tells you that the same keyboard is on sale at a downtown store for only \$15. If the downtown store is a 30-minute walk away, where should you buy the keyboard?

The Cost-Benefit Principle tells us that you should buy it downtown if the benefit of doing so exceeds the cost. The benefit of taking any action is the dollar value of everything you gain by taking it. Here, the benefit of buying downtown is exactly \$10, because that's the amount you'll save on the price of the keyboard. The cost of taking any action is the dollar value of everything you give up by taking it. Here, the cost of buying downtown is the dollar value you assign to the time and trouble it takes to make the trip. But how do we estimate that value?

One way is to perform the following hypothetical auction. Imagine that a stranger has offered to pay you to do an errand that involves the same walk downtown (perhaps to drop off a package for her at the post office). If she offered you a payment of, say, \$1,000, would you accept? If so, we know that your cost of walking downtown and back must be less than \$1,000. Now imagine her offer being reduced in small increments until you finally refuse the last offer. For example, if you'd agree to walk downtown and back for \$9 but not for \$8.99, then your cost of making the trip is \$9. In this case, you should buy the keyboard downtown because the \$10 you'll save (your benefit) is greater than your \$9 cost of making the trip.

But suppose your cost of making the trip had been greater than \$10. In that case, your best bet would have been to buy the keyboard from the nearby campus store. Confronted with this choice, different people may choose differently, depending on how costly they think it is to make the trip downtown. But although there is no uniquely correct choice, most people who are asked what they would do in this situation say they would buy the keyboard downtown.

economic surplus the benefit of taking an action minus its cost

Economic Surplus

Suppose that in Example 1.1 your "cost" of making the trip downtown was \$9. Compared to the alternative of buying the keyboard at the campus store, buying it downtown resulted in an **economic surplus** of \$1, the difference between the benefit of making the trip and its cost. In general, your goal as an economic decision maker is to choose those actions that generate the largest possible economic surplus. This means taking all actions that yield a positive total economic surplus, which is just another way of restating the Cost-Benefit Principle.

Note that the fact that your best choice was to buy the keyboard downtown doesn't imply that you *enjoy* making the trip, any more than choosing a large class means that you prefer large classes to small ones. It simply means that the trip is less unpleasant than the prospect of paying \$10 extra for the keyboard. Once again, you've faced a trade-off. In this case, the choice was between a cheaper keyboard and the free time gained by avoiding the trip.

Opportunity Cost

Of course, your mental auction could have produced a different outcome. Suppose, for example, that the time required for the trip is the only time you have left to study for a difficult test the next day. Or suppose you are watching one of your favorite shows on Netflix, or that you are tired and would love a short nap. In such cases, we say that the **opportunity cost** of making the trip—that is, the value of what you must sacrifice to walk downtown and back—is high and you are more likely to decide against making the trip.

opportunity cost the value of what must be forgone to undertake an activity

Strictly speaking, your opportunity cost of engaging in an activity is the value of everything you must sacrifice to engage in it. For instance, if seeing a movie requires not only that you buy a \$10 ticket, but also that you give up a \$20 dogwalking job that you would have been willing to do for free, then the opportunity cost of seeing the film is \$30.

Under this definition, *all* costs—both implicit and explicit—are opportunity costs. Unless otherwise stated, we will adhere to this strict definition.

We must warn you, however, that some economists use the term *opportunity cost* to refer only to the implicit value of opportunities forgone. Thus, in the example just discussed, these economists wouldn't include the \$10 ticket price when calculating the opportunity cost of seeing the film. But virtually all economists would agree that your opportunity cost of not doing the dogwalking job is \$20.

In the previous example, if watching another hour of your favorite show on Netflix is the most valuable opportunity that conflicts with the trip downtown, the opportunity cost of making the trip is the dollar value you place on pursuing that opportunity. It is the largest amount you'd be willing to pay to avoid watching your show at another time. Note that the opportunity cost of making the trip is not the combined value of *all* possible activities you could have pursued, but only the value of your *best* alternative—the one you would have chosen had you not made the trip.

Throughout the text we'll pose self-tests like the one that follows. You'll find that pausing to answer them will help you master key concepts in economics. Because doing these self-tests isn't very costly (indeed, many students report that they're actually fun), the Cost-Benefit Principle indicates that it's well worth your while to do them.

SELF-TEST 1.1

You would again save \$10 by buying the wireless keyboard downtown rather than at the campus store, but your cost of making the trip is now \$12, not \$9. By how much would your economic surplus be smaller if you bought the keyboard downtown rather than at the campus store?

The Role of Economic Models

Economists use the Cost-Benefit Principle as an abstract model of how an idealized rational individual would choose among competing alternatives. (By “abstract model” we mean a simplified description that captures the essential elements of a situation and allows us to analyze them in a logical way.) A computer model of a complex phenomenon like climate change, which must ignore many details and includes only the major forces at work, is an example of an abstract model.

Noneconomists are sometimes harshly critical of the economist's cost-benefit model on the grounds that people in the real world never conduct hypothetical mental auctions before deciding whether to make trips downtown. But this criticism betrays a fundamental misunderstanding of how abstract models can help explain and predict human behavior. Economists know perfectly well that people don't conduct hypothetical mental auctions when they make simple decisions. All the Cost-Benefit Principle really says is that a rational decision is one that is explicitly or implicitly based on a weighing of costs and benefits.

Most of us make sensible decisions most of the time, without being consciously aware that we are weighing costs and benefits, just as most people ride a bike without being consciously aware of what keeps them from falling. Through trial and error, we gradually learn what kinds of choices tend to work best in different contexts, just as bicycle riders internalize the relevant laws of physics, usually without being conscious of them.

Even so, learning the explicit principles of cost-benefit analysis can help us make better decisions, just as knowing about physics can help in learning to ride a bicycle. For instance, when a young economist was teaching his oldest son to ride a bike, he followed the time-honored tradition of running alongside the bike and holding onto his son, then giving him a push and hoping for the best. After several hours and painfully skinned elbows and knees, his son finally got it. A year later, someone pointed out that the trick to riding a bike is to turn slightly in whichever direction the bike is leaning. Of course! The economist passed this information along to his second son, who learned to ride almost instantly. Just as knowing a little physics can help you learn to ride a bike, knowing a little economics can help you make better decisions.

RECAP ↑**COST-BENEFIT ANALYSIS**

Scarcity is a basic fact of economic life. Because of it, having more of one good thing almost always means having less of another. The *Cost-Benefit Principle* holds that an individual (or a firm or a society) should take an action if, and only if, the extra benefit from taking the action is at least as great as the extra cost. The benefit of taking any action minus the cost of taking the action is called the *economic surplus* from that action. Hence, the Cost-Benefit Principle suggests that we take only those actions that create additional economic surplus.

THREE IMPORTANT DECISION PITFALLS¹

Rational people will apply the Cost-Benefit Principle most of the time, although probably in an intuitive and approximate way, rather than through explicit and precise calculation. Knowing that rational people tend to compare costs and benefits enables economists to predict their likely behavior. As noted earlier, for example, we can predict that students from wealthy families are more likely than others to attend colleges that offer small classes. (Again, while the cost of small classes is the same for all families, their benefit, as measured by what people are willing to pay for them, tends to be higher for wealthier families.)

Yet researchers have identified situations in which people tend to apply the Cost-Benefit Principle inconsistently. In these situations, the Cost-Benefit Principle may not predict behavior accurately. But it proves helpful in another way, by identifying specific strategies for avoiding bad decisions.

Pitfall 1: Measuring Costs and Benefits as Proportions Rather Than Absolute Dollar Amounts

As the next example makes clear, even people who seem to know they should weigh the pros and cons of the actions they are contemplating sometimes don't have a clear sense of how to measure the relevant costs and benefits.

EXAMPLE 1.2**Comparing Costs and Benefits****Should you walk downtown to save \$10 on a \$2,020 laptop computer?**

You are about to buy a \$2,020 laptop computer at the nearby campus store when a friend tells you that the same computer is on sale at a downtown store for only \$2,010. If the downtown store is half an hour's walk away, where should you buy the computer?

Assuming that the laptop is light enough to carry without effort, the structure of this example is exactly the same as that of Example 1.1. The only difference is that the price of the laptop is dramatically higher than the price of the wireless keyboard. As before, the benefit of buying downtown is the dollar amount you'll save, namely, \$10. And because it's exactly the same trip, its cost also must be the

¹The examples in this section are inspired by the pioneering research of Daniel Kahneman and the late Amos Tversky. Kahneman was awarded the 2002 Nobel Prize in Economics for his efforts to integrate insights from psychology into economics. You can read more about this work in Kahneman's brilliant 2011 book, *Thinking Fast and Slow* (New York: Macmillan).

same as before. So if you are perfectly rational, you should make the same decision in both cases. Yet when people are asked what they would do in these situations, the overwhelming majority say they'd walk downtown to buy the keyboard but would buy the laptop at the campus store. When asked to explain, most of them say something like, "The trip was worth it for the keyboard because you save 40 percent, but not worth it for the laptop because you save only \$10 out of \$2,020."

This is faulty reasoning. The benefit of the trip downtown is not the *proportion* you save on the original price. Rather, it is the *absolute dollar amount* you save. The benefit of walking downtown to buy the laptop is \$10, exactly the same as for the wireless keyboard. And because the cost of the trip must also be the same in both cases, the economic surplus from making both trips must be exactly the same. That means that a rational decision maker would make the same decision in both cases. Yet, as noted, most people choose differently.

The pattern of faulty reasoning in the decision just discussed is one of several decision pitfalls to which people are often prone. In the discussion that follows, we will identify two additional decision pitfalls. In some cases, people ignore costs or benefits that they ought to take into account. On other occasions they are influenced by costs or benefits that are irrelevant.

SELF-TEST 1.2

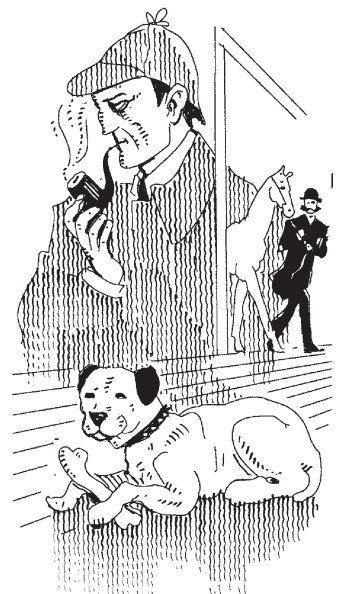
Which is more valuable: saving \$100 on a \$2,000 plane ticket to Tokyo or saving \$90 on a \$200 plane ticket to Chicago?

Pitfall 2: Ignoring Implicit Costs

Sherlock Holmes, Arthur Conan Doyle's legendary detective, was successful because he saw details that most others overlooked. In *Silver Blaze*, Holmes is called on to investigate the theft of an expensive racehorse from its stable. A Scotland Yard inspector assigned to the case asks Holmes whether some particular aspect of the crime requires further study. "Yes," Holmes replies, and describes "the curious incident of the dog in the nighttime."² "The dog did nothing in the nighttime,"² responds the puzzled inspector. But, as Holmes realized, that was precisely the problem! The watchdog's failure to bark when Silver Blaze was stolen meant that the watchdog knew the thief. This clue ultimately proved the key to unraveling the mystery.

Just as we often don't notice when a dog fails to bark, many of us tend to overlook the implicit value of activities that fail to happen. As discussed earlier, however, intelligent decisions require taking the value of forgone opportunities properly into account.

The opportunity cost of an activity, once again, is the value of all that must be forgone in order to engage in that activity. If buying a wireless keyboard downtown means not watching another hour of your favorite show on Netflix, then the value to you of watching the show is an implicit cost of the trip. Many people make bad decisions because they tend to ignore the value of such forgone opportunities. To avoid overlooking implicit costs, economists often translate questions like "Should I walk downtown?" into ones like "Should I walk downtown or watch another hour of my favorite show?"



Implicit costs are like dogs that fail to bark in the night. Many of us tend to overlook activities that fail to happen.

²Arthur Conan Doyle, "The Adventure of Silver Blaze," *The Memoirs of Sherlock Holmes* (London: George Newnes Ltd., 1893).

EXAMPLE 1.3**Implicit Cost****Should you use your frequent-flyer coupon to fly to Cancun for spring break?**

With spring break only a week away, you are still undecided about whether to go to Cancun with a group of classmates at the University of Iowa. The round-trip airfare from Cedar Rapids is \$500, but you have a frequent-flyer coupon you could use for the trip. All other relevant costs for the vacation week at the beach total exactly \$1,000. The most you would be willing to pay for the Cancun vacation is \$1,350. That amount is your benefit of taking the vacation. Your only alternative use for your frequent-flyer coupon is for a trip to Boston the weekend after spring break to attend your brother's wedding. (Your coupon expires shortly thereafter.) If the Cedar Rapids–Boston round-trip airfare is \$400, should you use your frequent-flyer coupon to fly to Cancun for spring break?

The Cost-Benefit Principle tells us that you should go to Cancun if the benefits of the trip exceed its costs. If not for the complication of the frequent-flyer coupon, solving this problem would be a straightforward matter of comparing your benefit from the week at the beach to the sum of all relevant costs. And because your airfare and other costs would add up to \$1,500, or \$150 more than your benefit from the trip, you would not go to Cancun.

But what about the possibility of using your frequent-flyer coupon to make the trip? Using it for that purpose might make the flight to Cancun seem free, suggesting you'd reap an economic surplus of \$350 by making the trip. But doing so also would mean you'd have to fork over \$400 for your airfare to Boston. So the implicit cost of using your coupon to go to Cancun is really \$400. If you use it for that purpose, the trip still ends up being a loser because the cost of the vacation, \$1,400, exceeds the benefit by \$50. In cases like these, you're much more likely to decide sensibly if you ask yourself, "Should I use my frequent-flyer coupon for this trip or save it for an upcoming trip?"



Bkinder/E+/Getty Images

Is your flight to Cancun “free” if you travel on a frequent-flyer coupon?

We cannot emphasize strongly enough that the key to using the Cost-Benefit Principle correctly lies in recognizing precisely what taking a given action prevents us from doing. Self-Test 1.3 illustrates this point by modifying the details of Example 1.3 slightly.

SELF-TEST 1.3

Refer to given information in Example 1.3, but this time your frequent-flyer coupon expires in a week, so your only chance to use it will be for the Cancun trip. Should you use your coupon?

Pitfall 3: Failing to Think at the Margin

When deciding whether to take an action, the only relevant costs and benefits are those that would occur as a result of taking the action. Sometimes people are influenced by costs they ought to ignore. Other times they compare the wrong costs and benefits. *The only costs that should influence a decision about whether to take an action are those we can avoid by not taking the action. Similarly, the only benefits we should consider are those that would not occur unless the action were taken.* As a practical matter, however, many decision makers appear to be influenced by costs or benefits that would

have occurred no matter what. Thus, people are often influenced by **sunk costs**—costs that are beyond recovery at the moment a decision is made. For example, money spent on a nontransferable, nonrefundable airline ticket is a sunk cost.

As the following example illustrates, sunk costs must be borne *whether or not an action is taken*, so they are irrelevant to the decision of whether to take the action.

sunk cost a cost that is beyond recovery at the moment a decision must be made

EXAMPLE 1.4 Sunk Cost

How much should you eat at an all-you-can-eat restaurant?

Sangam, an Indian restaurant in Philadelphia, offers an all-you-can-eat lunch buffet for \$10. Customers pay \$10 at the door, and no matter how many times they refill their plates, there is no additional charge. One day, as a goodwill gesture, the owner of the restaurant tells 20 randomly selected guests that they can eat at the all-you-can-eat buffet for free. The remaining guests pay the usual price. If all diners are rational, will those who are able to eat at the buffet for free consume a different amount of food, on average, than those who have to pay \$10 for the buffet?

Having eaten their first helping, diners in each group confront the following question: “Should I go back for another helping?” For rational diners, if the benefit of doing so exceeds the cost, the answer is yes; otherwise it is no. Note that at the moment of decision, the \$10 charge for the lunch is a sunk cost. Those who paid it have no way to recover it. Thus, for both groups, the (extra) cost of another helping is exactly zero. And because the people who received the free lunch were chosen at random, there’s no reason their appetites or incomes should be any different from those of other diners. The benefit of another helping thus should be the same, on average, for people in both groups. And because their respective costs and benefits are the same, the two groups should eat the same number of helpings, on average.

Psychologists and economists have experimental evidence, however, that people in such groups do *not* eat similar amounts.³ In particular, those who have to pay for the all-you-can-eat buffet tend to eat substantially more than those for whom the buffet is free. People in the former group somehow seem determined to “get their money’s worth.” Their implicit goal is apparently to minimize the average cost per bite of the food they eat. Yet minimizing average cost is not a particularly sensible objective. The irony is that diners who are determined to get their money’s worth usually end up eating too much.

The fact that the cost-benefit criterion failed the test of prediction in Example 1.4 does nothing to invalidate its advice about what people *should* do. If you are letting sunk costs influence your decisions, you can do better by changing your behavior.

In addition to paying attention to costs and benefits that should be ignored, people often use incorrect measures of the relevant costs and benefits. This error often occurs when we must choose the *extent* to which an activity should be pursued (as opposed to choosing whether to pursue it at all). We can apply the Cost-Benefit Principle in such situations by repeatedly asking the question, “Should I increase the level at which I am currently pursuing the activity?”

In attempting to answer this question, the focus should always be on the benefit and cost of an *additional* unit of activity. To emphasize this focus, economists refer to the cost of an additional unit of activity as its **marginal cost**. Similarly, the benefit of an additional unit of the activity is its **marginal benefit**.

When the problem is to discover the proper level for an activity, the cost-benefit rule is to keep increasing the level as long as the marginal benefit of the activity

marginal cost the increase in total cost that results from carrying out one additional unit of an activity

marginal benefit the increase in total benefit that results from carrying out one additional unit of an activity

³See, for example, Richard Thaler, “Toward a Positive Theory of Consumer Choice,” *Journal of Economic Behavior and Organization* 1, no. 1 (1980).

exceeds its marginal cost. As the following example illustrates, however, people often fail to apply this rule correctly.

EXAMPLE 1.5 Focusing on Marginal Costs and Benefits

Should SpaceX expand its launch program from four launches per year to five?

SpaceX accountants have estimated that the gains from the company’s jumbo rocket launch program are currently \$24 billion a year (an average of \$6 billion per launch) and that its costs are currently \$20 billion a year (an average \$5 billion per launch). On the basis of these estimates, they have recommended that the company increase its number of launches. Should SpaceX CEO Elon Musk follow their advice?

average cost the total cost of undertaking n units of an activity divided by n

average benefit the total benefit of undertaking n units of an activity divided by n

To discover whether the advice makes economic sense, we must compare the marginal cost of a launch to its marginal benefit. The accountants’ estimates, however, tell us only the **average cost** and **average benefit** of the program. These are, respectively, the total cost of the program divided by the number of launches and the total benefit divided by the number of launches.

Knowing the average benefit and average cost per launch for all rockets launched thus far is simply not useful for deciding whether to expand the program. Of course, the average cost of the launches undertaken so far *might* be the same as the cost of adding another launch. But it also might be either higher or lower than the marginal cost of a launch. The same holds true regarding average and marginal benefits.

Suppose, for the sake of discussion, that the benefit of an additional launch is in fact the same as the average benefit per launch thus far, \$6 billion. Should SpaceX add another launch? Not if the cost of adding the fifth launch would be more than \$6 billion. And the fact that the average cost per launch is only \$5 billion simply does not tell us anything about the marginal cost of the fifth launch.

Suppose, for example, that the relationship between the number of rockets launched and the total cost of the program is as described in Table 1.1. The average cost per launch (third column) when there are four launches would then be \$20 billion/4 = \$5 billion per launch, just as the accountants reported. But note in the second column of the table that adding a fifth launch would raise costs from \$20 billion to \$32 billion, making the marginal cost of the fifth launch \$12 billion. So if the benefit of an additional launch is \$6 billion, increasing the number of launches from four to five would make absolutely no economic sense.

TABLE 1.1
How Total Cost Varies with the Number of Launches

Number of launches	Total cost (\$ billions)	Average cost (\$ billion/launch)
0	0	0
1	3	3
2	7	3.5
3	12	4
4	20	5
5	32	6.4

The following example illustrates how to apply the *Cost-Benefit Principle* correctly in this case.

EXAMPLE 1.6 Focusing on Marginal Costs and Benefits

How many rockets should SpaceX launch?

SpaceX must decide how many rockets to launch. The benefit of each launch is estimated to be \$6 billion, and the total cost of the program again depends on the number of launches as shown in Table 1.1. How many rockets should SpaceX launch?

SpaceX should continue to launch its jumbo rockets as long as the marginal benefit of the program exceeds its marginal cost. In this example, the marginal benefit is constant at \$6 billion per launch, regardless of the number of rockets launched. SpaceX should thus keep launching rockets as long as the marginal cost per launch is less than or equal to \$6 billion.

Applying the definition of marginal cost to the total cost entries in the second column of Table 1.1 yields the marginal cost values in the third column of Table 1.2. (Because marginal cost is the change in total cost that results when we change the number of launches by one, we place each marginal cost entry midway between the rows showing the corresponding total cost entries.) Thus, for example, the marginal cost of increasing the number of launches from one to two is \$4 billion, the difference between the \$7 billion total cost of two launches and the \$3 billion total cost of one launch.

TABLE 1.2
How Marginal Cost Varies with the Number of Launches

Number of launches	Total cost (\$ billions)	Marginal cost (\$ billion/launch)
0	0	
1	3	3
2	7	4
3	12	5
4	20	8
5	32	12

As we see from a comparison of the \$6 billion marginal benefit per launch with the marginal cost entries in the third column of Table 1.2, the first three launches satisfy the cost-benefit test, but the fourth and fifth launches do not. SpaceX should thus launch three rockets.

SELF-TEST 1.4

If the marginal benefit of each launch had been not \$6 billion but \$9 billion, how many rockets should SpaceX have launched?

The cost-benefit framework emphasizes that the only relevant costs and benefits in deciding whether to pursue an activity further are *marginal* costs and benefits—measures that correspond to the *increment* of activity under consideration. In many

contexts, however, people seem more inclined to compare the *average* cost and benefit of the activity. As Example 1.5 made clear, increasing the level of an activity may not be justified, even though its average benefit at the current level is significantly greater than its average cost.

SELF-TEST 1.5

Should a basketball team's best player take all the team's shots?

A professional basketball team has a new assistant coach. The assistant notices that one player scores on a higher percentage of her shots than other players. Based on this information, the assistant suggests to the head coach that the star player take *all* the shots. That way, the assistant reasons, the team will score more points and win more games.

On hearing this suggestion, the head coach fires her assistant for incompetence. What was wrong with the assistant's idea?

RECAP

THREE IMPORTANT DECISION PITFALLS

1. **The pitfall of measuring costs or benefits proportionally.** Many decision makers treat a change in cost or benefit as insignificant if it constitutes only a small proportion of the original amount. Absolute dollar amounts, not proportions, should be employed to measure costs and benefits.
2. **The pitfall of ignoring implicit costs.** When performing a cost-benefit analysis of an action, it is important to account for all relevant costs, including the implicit value of alternatives that must be forgone in order to carry out the action. A resource (such as a frequent-flyer coupon) may have a high implicit cost, even if you originally got it “for free,” if its best alternative use has high value. The identical resource may have a low implicit cost, however, if it has no good alternative uses.
3. **The pitfall of failing to think at the margin.** When deciding whether to perform an action, the only costs and benefits that are relevant are those that would result from taking the action. It is important to ignore sunk costs—those costs that cannot be avoided even if the action isn't taken. Even though a ticket to a concert may have cost you \$100, if you've already bought it and cannot sell it to anyone else, the \$100 is a sunk cost and shouldn't influence your decision about whether to go to the concert. It's also important not to confuse average costs and benefits with marginal costs and benefits. Decision makers often have ready information about the total cost and benefit of an activity, and from these it's simple to compute the activity's average cost and benefit. A common mistake is to conclude that an activity should be increased if its average benefit exceeds its average cost. The Cost-Benefit Principle tells us that the level of an activity should be increased if, and only if, its *marginal* benefit exceeds its *marginal* cost.

Some costs and benefits, especially marginal costs and benefits and implicit costs, are important for decision making, while others, like sunk costs and average costs and benefits, are essentially irrelevant. This conclusion is implicit in our original statement of the Cost-Benefit Principle (an action should be taken if, and only if, the extra benefits of taking it exceed the extra costs).

NORMATIVE ECONOMICS VERSUS POSITIVE ECONOMICS

The examples discussed in the preceding section make the point that people *sometimes* choose irrationally. We must stress that our purpose in discussing these examples was not to suggest that people *generally* make irrational choices. On the contrary, most people appear to choose sensibly most of the time, especially when their decisions are important or familiar ones. The economist's focus on rational choice thus offers not only useful advice about making better decisions, but also a basis for predicting and explaining human behavior. We used the cost-benefit approach in this way when discussing how rising faculty salaries have led to larger class sizes. And as we will see, similar reasoning helps explain human behavior in virtually every other domain.

The Cost-Benefit Principle is an example of a **normative economic principle**, one that provides guidance about how we *should* behave. For example, according to the Cost-Benefit Principle, we should ignore sunk costs when making decisions about the future. As our discussion of the various decision pitfalls makes clear, however, the Cost-Benefit Principle is not always a **positive**, or **descriptive**, **economic principle**, one that describes how we actually *will* behave. As we saw, the Cost-Benefit Principle can be tricky to implement, and people sometimes fail to heed its prescriptions.

That said, we stress that knowing the relevant costs and benefits surely does enable us to predict how people will behave much of the time. If the benefit of an action goes up, it is generally reasonable to predict that people will be more likely to take that action. And conversely, if the cost of an action goes up, the safest prediction will be that people will be less likely to take that action.

When the Cost-Benefit Principle helps us predict people's behavior, it also acts as a positive economic principle. The principle stresses that the relevant costs and benefits usually help us predict behavior, but at the same time does not insist that people behave rationally in each instance. For example, if the price of heating oil were to rise sharply, we would invoke the Cost-Benefit Principle to say that people *should* turn down their thermostats. And although some may not follow that advice, the Cost-Benefit Principle would also predict that average thermostat settings *will* in fact go down.

normative economic principle one that says how people should behave

positive (or descriptive) economic principle one that predicts how people will behave

ECONOMICS: MICRO AND MACRO

By convention, we use the term **microeconomics** to describe the study of individual choices and of group behavior in individual markets. **Macroeconomics**, by contrast, is the study of the performance of national economies and of the policies that governments use to try to improve that performance. Macroeconomics tries to understand the determinants of such things as the national unemployment rate, the overall price level, and the total value of national output.

Our focus in this chapter is on issues that confront the individual decision maker, whether that individual confronts a personal decision, a family decision, a business decision, a government policy decision, or indeed any other type of decision. Further on, we'll consider economic models of groups of individuals such as all buyers or all sellers in a specific market. Later still we'll turn to broader economic issues and measures.

No matter which of these levels is our focus, however, our thinking will be shaped by the fact that, although economic needs and wants are effectively unlimited, the material and human resources that can be used to satisfy them are finite. Clear thinking about economic problems must therefore always take into account the idea of trade-offs—the idea that having more of one good thing usually means having less of another. Our economy and our society are shaped to a substantial degree by the choices people have made when faced with trade-offs.

microeconomics the study of individual choice under scarcity and its implications for the behavior of prices and quantities in individual markets

macroeconomics the study of the performance of national economies and the policies that governments use to try to improve that performance

THE APPROACH OF THIS TEXT

Choosing the number of students to register in each class is just one of many important decisions in planning an introductory economics course. Another concerns which topics to include on the course syllabus. There's a virtually inexhaustible set of issues that might be covered in an introductory course, but only limited time in which to cover them. There's no free lunch. Covering some inevitably means omitting others.

All textbook authors are forced to pick and choose. A textbook that covered *all* the issues would take up more than a whole floor of your campus library. It is our firm view that most introductory textbooks try to cover far too much. One reason that each of us was drawn to the study of economics is that a relatively short list of the discipline's core ideas can explain a great deal of the behavior and events we see in the world around us. So rather than cover a large number of ideas at a superficial level, our strategy is to focus on this short list of core ideas, returning to each entry again and again, in many different contexts. This strategy will enable you to internalize these ideas remarkably well in the brief span of a single course. And the benefit of learning a small number of important ideas well will far outweigh the cost of having to ignore a host of other, less important ones.

A second important element in our philosophy is a belief in the importance of active learning. In the same way that you can learn Spanish only by speaking and writing it, or tennis only by playing the game, you can learn economics only by *doing* economics. And because we want you to learn how to do economics, rather than just to read or listen passively as the authors or your instructor does economics, we'll make every effort to encourage you to stay actively involved.

For example, instead of just telling you about an idea, we'll usually first motivate the idea by showing you how it works in the context of a specific example. Often, these examples will be followed by self-tests for you to try, as well as applications that show the relevance of the idea to real life. Try working the self-tests *before* looking up the answers (which are at the back of the corresponding chapter).

Think critically about the applications: Do you see how they illustrate the point being made? Do they give you new insight into the issue? Work the problems at the end of the chapters and take extra care with those relating to points that you don't fully understand. Apply economic principles to the world around you. (We'll say more about this when we discuss economic naturalism below.) Finally, when you come across an idea or example that you find interesting, tell a friend about it. You'll be surprised to discover how much the mere act of explaining it helps you understand and remember the underlying principle. The more actively you can become engaged in the learning process, the more effective your learning will be.

ECONOMIC NATURALISM

With the rudiments of the cost-benefit framework under your belt, you are now in a position to become an "economic naturalist," someone who uses insights from economics to help make sense of observations from everyday life. People who have studied biology are able to observe and marvel at many details of nature that would otherwise have escaped their notice. For example, on a walk in the woods in early April, the novice may see only trees. In contrast, the biology student notices many different species of trees and understands why some are already in leaf while others still lie dormant. Likewise, the novice may notice that in some animal species males are much larger than females, but the biology student knows that pattern occurs only in species in which males take several mates. Natural selection favors larger males in those species because their greater size helps them prevail in the often bloody contests among males for access to females. In contrast, males tend to be roughly the same size as females in monogamous species, in which there is much less fighting for mates.



► Visit your instructor's Connect course and access your eBook to view this video.



Why does the light come on when you open the refrigerator door but not when you open the freezer?

Learning a few simple economic principles broadens our vision in a similar way. It enables us to see the mundane details of ordinary human existence in a new light. Whereas the uninitiated often fail even to notice these details, the economic naturalist not only sees them, but also becomes actively engaged in the attempt to understand them. Let's consider a few examples of questions economic naturalists might pose for themselves.

The Economic Naturalist 1.1



Why do many hardware manufacturers include more than \$1,000 worth of “free” software with a computer selling for only slightly more than that?

The software industry is different from many others in the sense that its customers care a great deal about product compatibility. When you and your classmates are working on a project together, for example, your task will be much simpler if you all use the same word-processing program. Likewise, an executive's life will be easier at tax time if her financial software is the same as her accountant's.

The implication is that the benefit of owning and using any given software program increases with the number of other people who use that same product. This unusual relationship gives the producers of the most popular programs an enormous advantage and often makes it hard for new programs to break into the market.

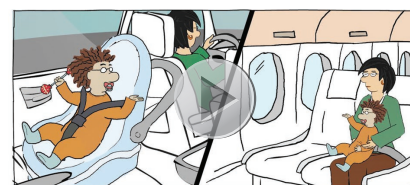
Recognizing this pattern, Intuit Corp. offered computer makers free copies of *Quicken*, its personal financial-management software. Computer makers, for their part, were only too happy to include the program because it made their new computers more attractive to buyers. *Quicken* soon became the standard for personal financial-management programs. By giving away free copies of the program, Intuit “primed the pump,” creating an enormous demand for upgrades of *Quicken* and for more advanced versions of related software. Thus, *TurboTax*, Intuit's personal income-tax software, has become the standard for tax-preparation programs.

Inspired by this success story, other software developers have jumped onto the bandwagon. Most hardware now comes bundled with a host of free software programs. Some software developers are even rumored to *pay* computer makers to include their programs!



connect

► Visit your instructor's Connect course and access your eBook to view this video.



Why are child safety seats required in automobiles but not in airplanes?

The Economic Naturalist 1.1 illustrates a case in which the *benefit* of a product depends on the number of other people who own that product. As the next Economic Naturalist demonstrates, the *cost* of a product may also depend on the number of others who own it.

The Economic Naturalist 1.2



Why don't auto manufacturers make cars without heaters?

Virtually every new car sold in the United States today has a heater. But not every car has a satellite navigation system. Why this difference?

One might be tempted to answer that, although everyone *needs* a heater, people can get along without navigation systems. Yet heaters are of limited use in places like Hawaii and Southern California.

Although heaters cost extra money to manufacture and are not useful in all parts of the country, they do not cost *much* money and are useful on at least a few days each year in most parts of the country. As time passed and people's incomes

grew, manufacturers found that people were ordering fewer and fewer cars without heaters. At some point it actually became cheaper to put heaters in *all* cars, rather than bear the administrative expense of making some cars with heaters and others without. No doubt a few buyers would still order a car without a heater if they could save some money in the process, but catering to these customers is just no longer worth it.

Similar reasoning explains why certain cars today cannot be purchased without a satellite navigation system. Buyers of the 2020 BMW 750i, for example, got one whether they wanted it or not. Most buyers of this car, which sells for more than \$85,000, have high incomes, so the overwhelming majority of them would have chosen to order a navigation system had it been sold as an option. Because of the savings made possible when all cars are produced with the same equipment, it would have actually cost BMW more to supply cars for the few who would want them without navigation systems.

Buyers of the least-expensive makes of car have much lower incomes on average than BMW 750i buyers. Accordingly, most of them have more pressing alternative uses for their money than to buy navigation systems for their cars, and this explains why some inexpensive makes continue to offer navigation systems only as options. But as incomes continue to grow, new cars without navigation systems will eventually disappear.

The insights afforded by The Economic Naturalist 1.2 suggest an answer to the following strange question:

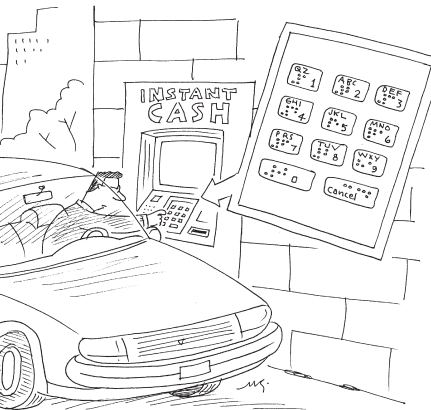


The Economic Naturalist 1.3

Why do the keypad buttons on drive-up automated teller machines have Braille dots?

Braille dots on elevator buttons and on the keypads of walk-up automated teller machines enable blind people to participate more fully in the normal flow of daily activity. But even though blind people can do many remarkable things, they cannot drive automobiles on public roads. Why, then, do the manufacturers of automated teller machines install Braille dots on the machines at drive-up locations?

The answer to this riddle is that once the keypad molds have been manufactured, the cost of producing buttons with Braille dots is no higher than the cost of producing smooth buttons. Making both would require separate sets of molds and two different types of inventory. If the patrons of drive-up machines found buttons with Braille dots harder to use, there might be a reason to incur these extra costs. But since the dots pose no difficulty for sighted users, the best and cheapest solution is to produce only keypads with dots.



Why do the keypad buttons on drive-up automated teller machines have Braille dots?

The preceding example was suggested by Cornell student Bill Tjoa, in response to the following assignment:

SELF-TEST 1.6

In 500 words or less, use cost-benefit analysis to explain some pattern of events or behavior you have observed in your own environment.

There is probably no more useful step you can take in your study of economics than to perform several versions of the assignment in Self-Test 1.6. Students who do so almost invariably become lifelong economic naturalists. Their mastery of economic concepts not only does not decay with the passage of time, but it actually grows stronger. We urge you, in the strongest possible terms, to make this investment!

SUMMARY

- Economics is the study of how people make choices under conditions of scarcity and of the results of those choices for society. Economic analysis of human behavior begins with the assumption that people are rational—that they have well-defined goals and try to achieve them as best they can. In trying to achieve their goals, people normally face trade-offs: Because material and human resources are limited, having more of one good thing means making do with less of some other good thing. (LO1)
- Our focus in this chapter has been on how rational people make choices among alternative courses of action. Our basic tool for analyzing these decisions is cost-benefit analysis. The Cost-Benefit Principle says that a person should take an action if, and only if, the benefit of that action is at least as great as its cost. The benefit of an action is defined as the largest dollar amount the person would be willing to pay in order to take the action. The cost of an action is defined as the dollar value of everything the person must give up in order to take the action. (LO2)
- In using the cost-benefit framework, we need not presume that people choose rationally all the time. Indeed, we identified three common pitfalls that plague decision makers in all walks of life: a tendency to treat small proportional changes as insignificant, a tendency to ignore implicit costs, and a tendency to fail to think at the margin—for example, by failing to ignore sunk costs or by failing to compare marginal costs and benefits. (LO3)
- Often the question is not whether to pursue an activity but rather how many units of it to pursue. In these cases, the rational person pursues additional units as long as the marginal benefit of the activity (the benefit from pursuing an additional unit of it) exceeds its marginal cost (the cost of pursuing an additional unit of it). (LO4)
- Microeconomics is the study of individual choices and of group behavior in individual markets, while macroeconomics is the study of the performance of national economics and of the policies that governments use to try to improve economic performance.

KEY TERMS

average benefit
average cost
economic surplus
economics
macroeconomics

marginal benefit
marginal cost
microeconomics
normative economic principle
opportunity cost

positive (or descriptive) economic principle
rational person
sunk cost

REVIEW QUESTIONS

1. A friend of yours on the tennis team says, “Private tennis lessons are definitely better than group lessons.” Explain what you think she means by this statement. Then use the Cost-Benefit Principle to explain why private lessons are not necessarily the best choice for everyone. (LO2)
2. True or false: Your willingness to drive downtown to save \$30 on a new appliance should depend on what fraction of the total selling price \$30 is. Explain. (LO3)
3. Why might someone who is trying to decide whether to see a movie be more likely to focus on the \$10 ticket price than on the \$20 he or she would fail to earn by not dogwalking? (LO3)
4. Many people think of their air travel as being free when they use frequent-flyer coupons. Explain why

these people are likely to make wasteful travel decisions. (LO3)

5. Is the nonrefundable tuition payment you made to your university this semester a sunk cost? How

would your answer differ if your university were to offer a full tuition refund to any student who dropped out of school during the first two months of the semester? (LO3)

PROBLEMS

McGraw Hill connect

- Suppose your school is considering whether to spend \$20 million building a new state-of-the-art recreation facility. All of the students agree that the existing facility is in disrepair and that a new facility would be much nicer. Despite this, however, when students are asked to vote on whether they would like the school to build the new recreation facility, over 78 percent vote no. Why might such a large fraction of students vote no even though they all agree that a new recreation facility would be much nicer than the existing one? (LO1)
- Suppose the most you would be willing to pay to have a freshly washed car before going out on a date is \$6. The smallest amount for which you would be willing to wash someone else's car is \$3.50. You are going out this evening and your car is dirty. How much economic surplus would you receive from washing it? (LO2)
- To earn extra money in the summer, you grow tomatoes and sell them at a local farmers' market for 30 cents per pound. By adding compost to your garden, you can increase your yield as shown in the accompanying table. If compost costs 50 cents per pound and your goal is to make as much profit as possible, how many pounds of compost should you add? (LO2)

Pounds of compost	Pounds of tomatoes
	100
1	120
2	125
3	128
4	130
5	131
6	131.5

- 4* You and your friend Jamal have identical tastes. At 2 p.m., you go to the local Ticketmaster outlet and buy a nonrefundable \$30 ticket to a basketball game to be played that night in Syracuse, 50 miles north of your home in Ithaca. Jamal plans to attend the same game, but because he cannot get to the Ticketmaster outlet, he plans to buy his ticket at the game. Tickets sold at the game cost only \$25 because they carry no Ticketmaster surcharge. (Many people nonetheless

pay the higher price at Ticketmaster, to be sure of getting good seats.) At 4 p.m., an unexpected snowstorm begins, making the prospect of the drive to Syracuse much less attractive than before (but ensuring the availability of good seats). If both you and Jamal are rational, is one of you more likely to attend the game than the other? (LO2)

- Kenya is a mushroom farmer. She invests all her spare cash in additional mushrooms, which grow on otherwise useless land behind her barn. The mushrooms double in weight during their first year, after which time they are harvested and sold at a constant price per pound. Kenya's friend Fatima asks Kenya for a loan of \$200, which she promises to repay after one year. How much interest will Fatima have to pay Kenya in order for Kenya to recover her opportunity cost of making the loan? Explain briefly. (LO3)
- Suppose that in the last few seconds you devoted to question 1 on your physics exam you earned 4 extra points, while in the last few seconds you devoted to question 2 you earned 10 extra points. You earned a total of 48 and 12 points, respectively, on the two questions, and the total time you spent on each was the same. If you could take the exam again, how—if at all—should you reallocate your time between these questions? (LO3)
- Monica and Rachel have the same preferences and incomes. Just as Monica arrived at the theater to see a play, she discovered that she had lost the \$10 ticket she had purchased earlier. Rachel also just arrived at the theater planning to buy a ticket to see the same play when she discovered that she had lost a \$10 bill from her wallet. If both Monica and Rachel are rational and both still have enough money to pay for a ticket, is one of them more likely than the other to go ahead and see the play anyway? (LO3)
- Residents of your city are charged a fixed weekly fee of \$6 for garbage collection. They are allowed to put out as many cans as they wish. The average household disposes of three cans of garbage per week under this plan. Now suppose that your city changes to a "tag" system. Each can of garbage to be collected must have a tag affixed to it. The tags cost \$2 each and are not reusable. What effect do you think the introduction of

*Denotes more difficult problem.

the tag system will have on the total quantity of garbage collected in your city? Explain briefly. (LO4)

9. Once a week, Hector purchases a six-pack of cola and puts it in his refrigerator for his two children. He invariably discovers that all six cans are gone on the first day. Jin also purchases a six-pack of cola once a week for his two children, but unlike Hector, he tells them that each may drink no more than three cans per week. If the children use cost-benefit analysis each time they decide whether to drink a can of cola, explain why the cola lasts much longer at Jin's house than at Hector's. (LO4)
- 10.* Suppose there is only one electric scooter company in Adriana's hometown. Currently, the company charges 20 cents per minute, and there is no fee to unlock a scooter. The scooter company is considering changing its pricing plan so that it

would charge \$1 to unlock a scooter and 10 cents per minute. If Adriana never takes a ride that lasts less than 10 minutes, then what will happen to the average length of her rides if the scooter company switches to the new pricing plan? Explain. (LO4)

- 11.* The meal plan at University A lets students eat as much as they like for a fixed fee of \$500 per semester. The average student there eats 250 pounds of food per semester. University B charges \$500 for a book of meal tickets that entitles the student to eat 250 pounds of food per semester. If the student eats more than 250 pounds, he or she pays \$2 for each additional pound; if the student eats less, he or she gets a \$2 per pound refund. If students are rational, at which university will average food consumption be higher? Explain briefly. (LO4)

ANSWERS TO SELF-TESTS

- 1.1 The benefit of buying the wireless keyboard downtown is again \$10 but the cost is now \$12, so your economic surplus would be \$2 smaller than if you'd bought it at the campus store. (LO2)
- 1.2 Saving \$100 is \$10 more valuable than saving \$90, even though the percentage saved is much greater in the case of the Chicago ticket. (LO3)
- 1.3 Since you now have no alternative use for your coupon, the opportunity cost of using it to pay for the Cancun trip is zero. That means your economic surplus from the trip will be $\$1,350 - \$1,000 = \$350 > 0$, so you should use your coupon and go to Cancun. (LO3)
- 1.4 The marginal benefit of the fourth launch is \$9 billion, which exceeds its marginal cost of \$8 billion, so

the fourth launch should be added. But the fifth launch should not, because its marginal cost (\$12 billion) exceeds its marginal benefit (\$9 billion). (LO3)

- 1.5 If the star player takes one more shot, some other player must take one less. The fact that the star player's *average* success rate is higher than the other players' does not mean that the probability of making her *next* shot (the marginal benefit of having her shoot once more) is higher than the probability of another player making her next shot. Indeed, if the best player took all her team's shots, the other team would focus its defensive effort entirely on her, in which case letting others shoot would definitely pay. (LO3)
- 1.6 Answers will vary.

Working with Equations, Graphs, and Tables

Although many of the examples and most of the end-of-chapter problems in this book are quantitative, none requires mathematical skills beyond rudimentary high school algebra and geometry. In this brief appendix, we review some of the skills you'll need for dealing with these examples and problems.

One important skill is to be able to read simple verbal descriptions and translate the information they provide into the relevant equations or graphs. You'll also need to be able to translate information given in tabular form into an equation or graph, and sometimes you'll need to translate graphical information into a table or equation. Finally, you'll need to be able to solve simple systems with two equations and two unknowns. The following examples illustrate all the tools you'll need.

USING A VERBAL DESCRIPTION TO CONSTRUCT AN EQUATION

We begin with an example that shows how to construct a billing equation for a ride on an electric scooter from a verbal description of the billing plan.

EXAMPLE 1A.1 A Verbal Description

An electric scooter company charges \$1 to unlock a scooter plus 20 cents per minute. Write an equation that describes your bill for riding this scooter.

An **equation** is a simple mathematical expression that describes the relationship between two or more **variables**, or quantities, that are free to assume different values in some range. The most common type of equation we'll work with contains two types of variables: **dependent variables** and **independent variables**. In this example, the dependent variable is the dollar amount of your scooter bill and the independent variable is the variable on which your bill depends, namely, the number of minutes you ride the scooter. Your bill also depends on the \$1 fee to unlock the scooter and the 20 cents per minute charge. But, in this example, those amounts are **constants**, not variables. A constant, also called a **parameter**, is a quantity in an equation that is fixed in value, not free to vary. As the terms suggest, the dependent variable describes an outcome that depends on the value taken by the independent variable.

Once you've identified the dependent variable and the independent variable, choose simple symbols to represent them. In algebra courses, X is typically used to represent the independent variable and Y the dependent variable. Many people find it easier to remember what the variables stand for, however, if they choose symbols that are linked in some straightforward way to the quantities that

equation a mathematical expression that describes the relationship between two or more variables

variable a quantity that is free to take a range of different values

dependent variable a variable in an equation whose value is determined by the value taken by another variable in the equation

independent variable a variable in an equation whose value determines the value taken by another variable in the equation

constant (or parameter) a quantity that is fixed in value

the variables represent. Thus, in this example, we might use B to represent your scooter *bill* in dollars and T to represent the total *time* in minutes of your ride.

Having identified the relevant variables and chosen symbols to represent them, you are now in a position to write the equation that links them

$$B = 1 + 0.20T, \quad (1A.1)$$

where B is your scooter bill in dollars and T is the number of minutes you spend riding the scooter. The fee to unlock the scooter (1) and the charge per minute (0.20) are parameters in this equation. Note the importance of being clear about the units of measure. Because B represents the scooter bill in dollars, we must also express the fee to unlock the scooter and the per-minute charge in dollars, which is why the latter number appears in Equation 1A.1 as 0.20 rather than 20. Equation 1A.1 follows the normal convention in which the dependent variable appears by itself on the left-hand side while the independent variable or variables and constants appear on the right-hand side.

Once we have the equation for the scooter bill, we can use it to calculate how much you'll owe as a function of the number of minutes you spend riding. For example, if you ride the scooter for 16 minutes, you can calculate your scooter bill by simply substituting 16 minutes for T in Equation 1A.1

$$B = 1 + 0.20(16) = 4.20. \quad (1A.2)$$

Your scooter bill when you ride for 16 minutes is thus equal to \$4.20.

SELF-TEST 1A.1

Under the scooter billing plan described in Example 1A.1, how much would you owe for a 22-minute ride?

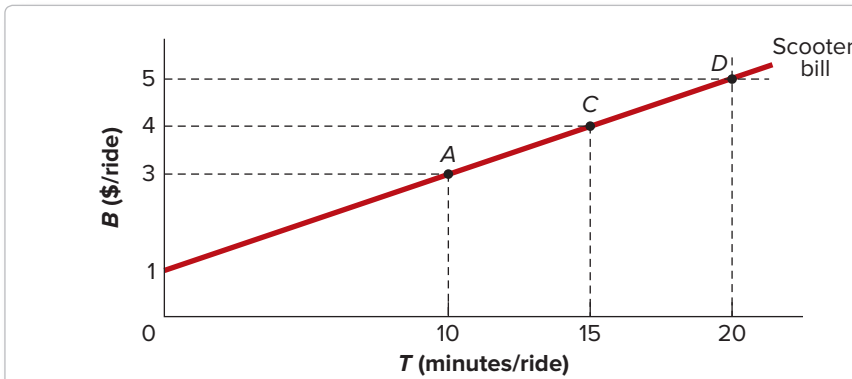
GRAPHING THE EQUATION OF A STRAIGHT LINE

The next example shows how to portray the billing plan described in Example 1A.1 as a graph.

EXAMPLE 1A.2 Graphing an Equation

Construct a graph that portrays the scooter billing plan described in Example 1A.1, putting your total bill, in dollars, on the vertical axis and the length of your ride, in minutes, on the horizontal axis.

The first step in responding to this instruction is the one we just took, namely, to translate the verbal description of the billing plan into an equation. When graphing an equation, the normal convention is to use the vertical axis to represent the dependent variable and the horizontal axis to represent the independent variable. In Figure 1A.1, we therefore put B on the vertical axis and T on the horizontal axis. One way to construct the graph shown in the figure is to begin by plotting the bill values that correspond to several different ride lengths. For example, someone who rides for 10 minutes would have a bill of $B = 1 + 0.20(10) = \$3$. Thus, in Figure 1A.1 the value of 10 minutes on the horizontal axis corresponds to a bill of \$3 on the vertical axis (point A). Someone who rides for 15 minutes would have a bill of

**FIGURE 1A.1****The Scooter Billing Plan in Example 1A.1.**

The graph of the equation $B = 1 + 0.20T$ is the straight line shown. Its vertical intercept is 1 and its slope is 0.20.

$B = 1 + 0.20(15) = \$4$, so the value of 15 minutes on the horizontal axis corresponds to \$4 on the vertical axis (point C). Similarly, someone who rides for 20 minutes would have a bill of $B = 1 + 0.20(20) = \$5$, so the value of 20 minutes on the horizontal axis corresponds to \$5 on the vertical axis (point D). The line joining these points is the graph of the billing Equation 1A.1.

As shown in Figure 1A.1, the graph of the equation $B = 1 + 0.20T$ is a straight line. The parameter 1 is the **vertical intercept** of the line—the value of B when $T = 0$, or the point at which the line intersects the vertical axis. The parameter 0.20 is the **slope** of the line, which is the ratio of the **rise** of the line to the corresponding **run**. The ratio rise/run is simply the vertical distance between any two points on the line divided by the horizontal distance between those points. For example, if we choose points A and C in Figure 1A.1, the rise is $4 - 3 = 1$ and the corresponding run is $15 - 10 = 5$, so $\text{rise/run} = 1/5 = 0.20$. More generally, for the graph of any equation $Y = a + bX$, the parameter a is the vertical intercept and the parameter b is the slope.

vertical intercept in a straight line, the value taken by the dependent variable when the independent variable equals zero

slope in a straight line, the ratio of the vertical distance the straight line travels between any two points (**rise**) to the corresponding horizontal distance (**run**)

DERIVING THE EQUATION OF A STRAIGHT LINE FROM ITS GRAPH

The next example shows how to derive the equation for a straight line from a graph of the line.

EXAMPLE 1A.3 Deriving an Equation from a Graph

Figure 1A.2 shows the graph of the billing plan for an electric scooter company. What is the equation for this graph? How much is the fee to unlock a scooter under this plan? How much is the charge per minute?

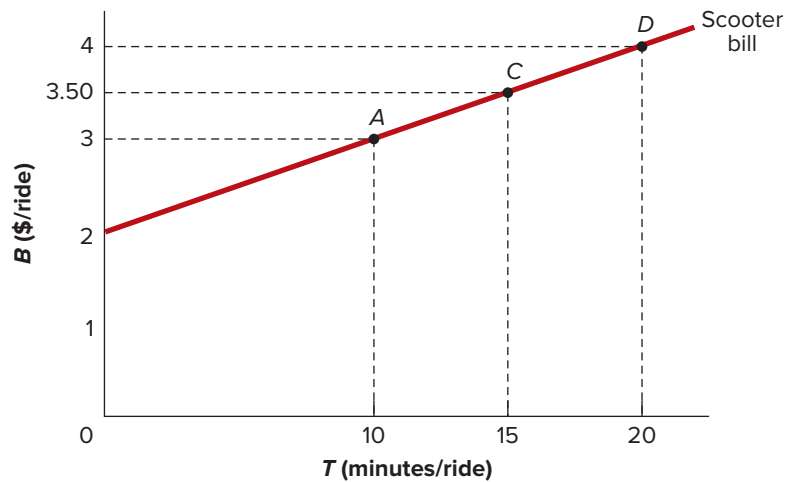
The slope of the line shown is the rise between any two points divided by the corresponding run. For points A and C, $\text{rise} = 3.50 - 3 = 0.50$ and $\text{run} = 15 - 10 = 5$, so the slope equals $\text{rise/run} = 0.50/5 = 0.10$. And since the horizontal intercept of the line is 2, its equation must be given by

$$B = 2 + 0.10T. \quad (1A.3)$$

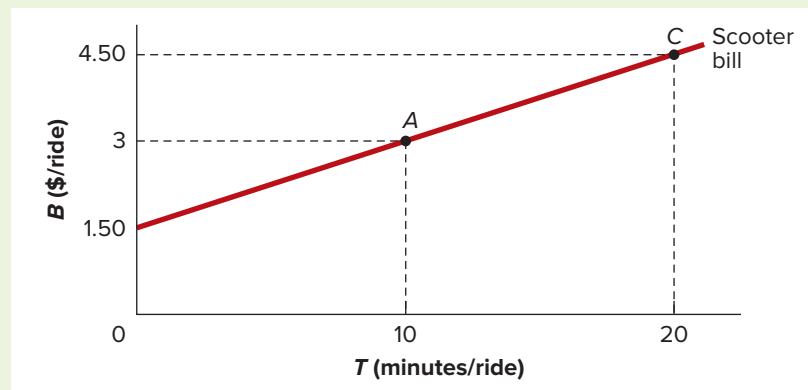
Under this plan, the fee to unlock a scooter is the value of the bill when $T = 0$, which is \$2. The charge per minute is the slope of the billing line, 0.10, or 10 cents per minute.

FIGURE 1A.2**Another Scooter Billing Plan.**

The vertical distance between points A and C is $3.50 - 3 = 0.50$ units, and the horizontal distance between points A and C is $15 - 10 = 5$, so the slope of the line is $0.50/5 = 0.10$. The vertical intercept (the value of B when $T = 0$) is 2. So the equation for the billing plan shown is $B = 2 + 0.10T$.

**SELF-TEST 1A.2**

Write the equation for the billing plan shown in the accompanying graph. How much does it cost to unlock a scooter? What is the charge per minute?



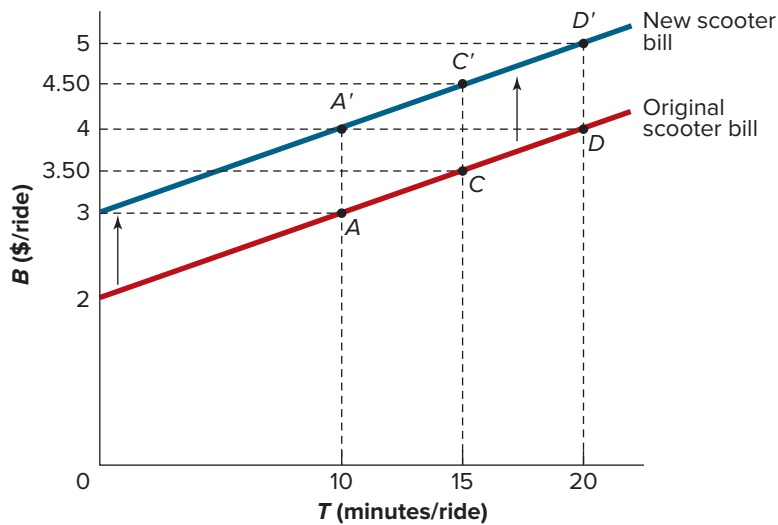
CHANGES IN THE VERTICAL INTERCEPT AND SLOPE

The next two examples and self-tests provide practice in seeing how a line shifts with a change in its vertical intercept or slope.

EXAMPLE 1A.4**Change in Vertical Intercept**

Show how the billing plan whose graph is in Figure 1A.2 would change if the fee to unlock a scooter were increased from \$2 to \$3.

An increase in the fee to unlock a scooter from \$2 to \$3 would increase the vertical intercept of the billing plan by \$1 but would leave its slope unchanged. An increase in the fee to unlock a scooter thus leads to a parallel upward shift in the billing plan by \$1, as shown in Figure 1A.3. For any given number of minutes the new bill will be \$1 higher than the old bill. Thus, a 10-minute ride cost \$3 under the original plan

**FIGURE 1A.3****The Effect of an Increase in the Vertical Intercept.**

An increase in the vertical intercept of a straight line produces an upward parallel shift in the line.

(point A) but \$4 under the new plan (point A'). A 15-minute ride cost \$3.50 under the original plan (point C), but \$4.50 under the new plan (point C'). And a 20-minute ride cost \$4 under the original plan (point D), but \$5 under the new plan (point D').

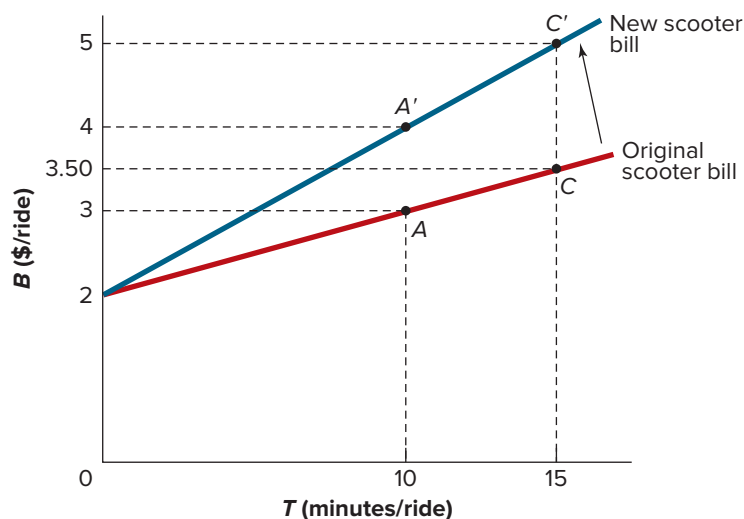
SELF-TEST 1A.3

Show how the billing plan whose graph is in Figure 1A.2 would change if the fee to unlock a scooter were reduced from \$2 to \$1.

EXAMPLE 1A.5 Change in Slope

Show how the billing plan whose graph is in Figure 1A.2 would change if the charge per minute were increased from 10 cents to 20 cents.

Because the fee to unlock a scooter is unchanged, the vertical intercept of the new billing plan continues to be 2. But the slope of the new plan, shown in

**FIGURE 1A.4****The Effect of an Increase in the Charge per Minute.**

Because the fixed monthly fee continues to be \$2, the vertical intercept of the new plan is the same as that of the original plan. With the new charge per minute of 20 cents, the slope of the billing plan rises from 0.10 to 0.20.

Figure 1A.4, is 0.20, or twice the slope of the original plan. More generally, in the equation $Y = a + bX$, an increase in b makes the slope of the graph of the equation steeper.

SELF-TEST 1A.4

Show how the billing plan whose graph is in Figure 1A.2 would change if the charge per minute were reduced from 10 cents to 5 cents.

Self-Test 1A.4 illustrates the general rule that in an equation $Y = a + bX$, a reduction in b makes the slope of the graph of the equation less steep.

CONSTRUCTING EQUATIONS AND GRAPHS FROM TABLES

The next example and self-test show how to transform tabular information into an equation or graph.

EXAMPLE 1A.6 Transforming a Table to a Graph

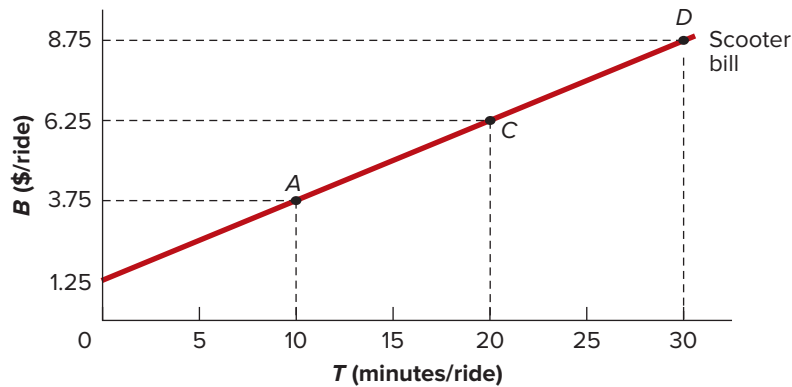
Table 1A.1 shows four points from a scooter billing equation. If all points on this billing equation lie on a straight line, find the vertical intercept of the equation and graph it. What is the fee to unlock a scooter? What is the charge per minute? Calculate the total bill for a 30-minute ride.

TABLE 1A.1
Points on a Scooter Billing Plan

Total bill (\$/ride)	Length of ride (minutes/ride)
2.50	5
3.75	10
5.00	15
6.25	20

One approach to this problem is simply to plot any two points from the table on a graph. Because we are told that the billing equation is a straight line, that line must be the one that passes through any two of its points. Thus, in Figure 1A.5 we use A to denote the point from Table 1A.1 for which a bill of \$3.75 corresponds to a 10-minute ride (second row) and C to denote the point for which a bill of \$6.25 corresponds to a 20-minute ride (fourth row). The straight line passing through these points is the graph of the billing equation.

Unless you have a steady hand, however, or use extremely large graph paper, the method of extending a line between two points on the billing plan is unlikely to be very accurate. An alternative approach is to calculate the equation for the billing plan directly. Because the equation is a straight line, we know that it takes the general form $B = f + sT$, where f is the fixed monthly fee and s is the slope.

**FIGURE 1A.5**

Plotting the Scooter Billing Plan from a Sample of Points.

Point A is taken from row 2, Table 1A.1, and point C from row 4. The billing plan is the straight line that passes through these points.

Our goal is to calculate the vertical intercept f and the slope s . From the same two points we plotted earlier, A and C, we can calculate the slope of the billing plan as $s = \text{rise/run} = 2.50/10 = 0.25$.

So all that remains is to calculate f , the fee to unlock a scooter. At point C on the billing plan, the total bill is \$6.25 for 20 minutes, so we can substitute $B = 6.25$, $s = 0.25$, and $T = 20$ into the general equation $B = f + sT$ to obtain

$$6.25 = f + 0.25(20), \quad (1A.4)$$

or

$$6.25 = f + 5, \quad (1A.5)$$

which solves for $f = 1.25$. So the billing equation must be

$$B = 1.25 + 0.25T. \quad (1A.6)$$

For this billing equation, the fee to unlock a scooter is \$1.25, the charge per minute is 25 cents (\$0.25/minute), and the total bill for a 30-minute ride is $B = 1.25 + 0.25(30) = \$8.75$, just as shown in Figure 1A.5.

SELF-TEST 1A.5

The following table shows four points from a scooter billing plan.

Total bill (\$/ride)	Length of ride (minutes/ride)
2.50	5
4.25	10
6.00	15
7.75	20

If all points on this billing plan lie on a straight line, find the vertical intercept of the corresponding equation without graphing it. What is the fee to unlock a scooter? What is the charge per minute? How much would the charges be for a 30-minute ride?

SOLVING SIMULTANEOUS EQUATIONS

The next example and self-test demonstrate how to proceed when you need to solve two equations with two unknowns.

EXAMPLE 1A.7 Solving Simultaneous Equations

Suppose you are trying to choose between two electric scooter companies, each with a different pricing plan. If you choose Company 1, your charges will be computed according to the equation

$$B = 0.50 + 0.30T, \quad (1A.7)$$

where B is again your bill in dollars and T is the length of your ride in minutes. If you choose Company 2, your bill will be computed according to the equation

$$B = 2 + 0.15T. \quad (1A.8)$$

How many minutes long would your ride have to be to make Company 2 cheaper?

Company 1 has the attractive feature of a relatively low fee to unlock a scooter, but also the unattractive feature of a relatively high rate per minute. In contrast, Company 2 has a relatively high fee to unlock a scooter but a relatively low rate per minute. Someone who wanted to make a short ride (for example, 4 minutes) would do better under Company 1 (bill = \$1.70) than under Company 2 (bill = \$2.60) because the low fee to unlock a scooter for Company 1 would more than compensate for its higher rate per minute. Conversely, someone who wanted to make a long ride (say, 15 minutes) would do better under Company 2 (bill = \$4.25) than under Company 1 (bill = \$5.00) because Company 2's lower rate per minute would more than compensate for its higher fee to unlock a scooter.

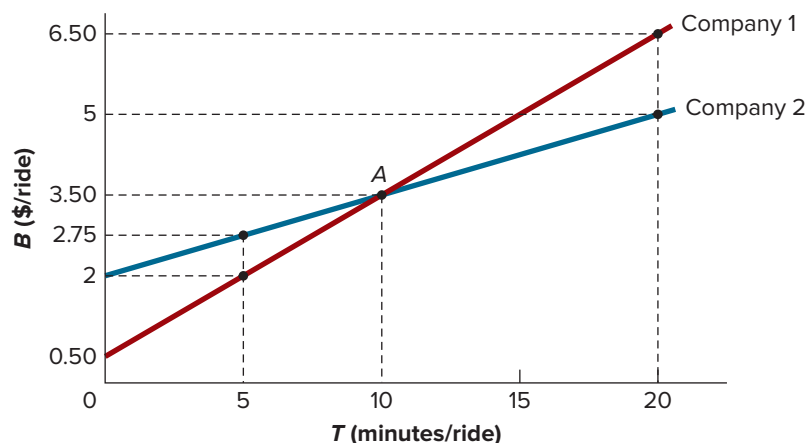
Our task here is to find the break-even ride length, which is ride length for which the bill is the same under the two plans. One way to answer this question is to graph the two billing plans and see where they cross. At that crossing point, the two equations are satisfied simultaneously, which means that the length of the rides will be the same, as will the bills.

In Figure 1A.6, we see that the graphs of the two plans cross at A , where both yield a bill of \$3.50 for a 10-minute ride. The break-even ride length for these two

FIGURE 1A.6

The Break-Even Volume of Ride Length.

For 10-minute rides, your bill will be the same under both plans. For longer rides, Company 2 is cheaper; Company 1 is cheaper for shorter rides.



companies is thus 10 minutes. If the length of your ride is longer than that, you will save money by choosing Company 2. For example, if you take a 20-minute ride, your bill under Company 2 (\$5) will be \$1.50 cheaper than under Company 1 (\$6.50). Conversely, if you ride for less than 10 minutes, you will do better under Company 1. For example, if you take a 5-minute ride, your bill under Company 1 (\$2) will be 75 cents cheaper than under Company 2 (\$2.75). For 10-minute rides, the two companies cost exactly the same (\$3.50).

The question posed here also may be answered algebraically. As in the graphical approach just discussed, our goal is to find the point (T, B) that satisfies both billing equations simultaneously. As a first step, we rewrite the two billing equations, one on top of the other, as

$$\begin{array}{ll} B = 0.50 + 0.30T & \text{(Company 1).} \\ B = 2 + 0.15T & \text{(Company 2).} \end{array}$$

As you'll recall from high school algebra, if we subtract the terms from each side of one equation from the corresponding terms of the other equation, the resulting differences must be equal. So if we subtract the terms on each side of the Company 2 equation from the corresponding terms in the Company 1 equation, we get

$$\begin{array}{ll} B = 0.50 + 0.30T & \text{(Company 1).} \\ -B = -2 - 0.15T & \text{(-Company 2).} \\ \hline 0 = -1.5 + 0.15T & \text{(Company 1 - Company 2).} \end{array}$$

Finally, we solve the last equation (Company 1 - Company 2) to get $T = 10$.

Plugging $T = 10$ into either plan's equation, we then find $B = 3.50$. For example, Company 1's equation yields $0.50 + 0.30(10) = 3.50$, as does Company 2's $2 + 0.15(10) = 3.50$.

Because the point $(T, B) = (10, 3.50)$ lies on the equations for both plans simultaneously, the algebraic approach just described is often called the *method of simultaneous equations*.

SELF-TEST 1A.6

Suppose you are trying to choose between two electric scooter companies. If you choose Company 1, your bill will be computed according to the equation

$$B = 0.20 + 0.40T \quad \text{(Company 1),}$$

where B is again your bill in dollars and T is the length of your ride in minutes. If you choose Company 2, your bill will be computed according to the equation

$$B = 5 + 0.10T \quad \text{(Company 2).}$$

Use the algebraic approach described in the preceding example to find the break-even ride length for these plans.

KEY TERMS

constant (or parameter)
dependent variable
equation

independent variable
rise
run

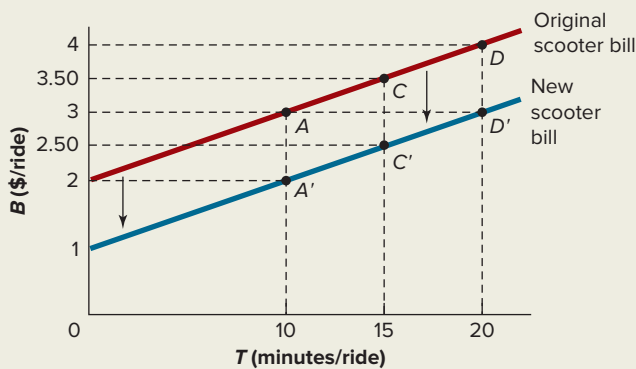
slope
variable
vertical intercept

ANSWERS TO APPENDIX SELF-TESTS

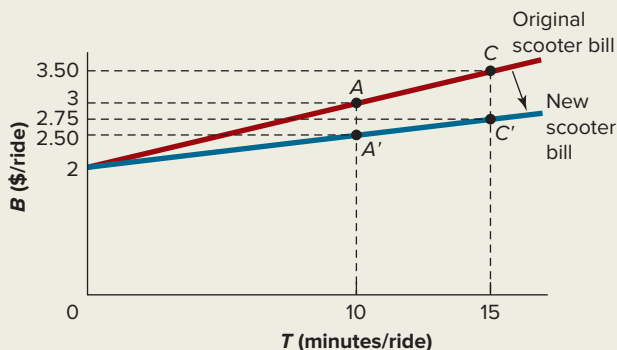
1A.1 To calculate your bill for a 22-minute ride, substitute 22 minutes for T in Equation 1A.1 to get $B = 1 + 0.20(22) = \$5.40$.

1A.2 Calculating the slope using points A and C , we have $\text{rise} = 4.50 - 3 = 1.50$ and $\text{run} = 20 - 10 = 10$, so $\text{rise/run} = 1.50/10 = 0.15$. And since the horizontal intercept of the line is 1.5, its equation is $B = 1.5 + 0.15T$. Under this plan, the fee to unlock a scooter is \$1.50 and the charge per minute is the slope of the billing line, 0.15, or 15 cents per minute.

1A.3 A \$1 reduction in the fee to unlock a scooter would produce a downward parallel shift in the billing plan by \$1.



1A.4 With the fee to unlock a scooter unchanged, the vertical intercept of the new billing plan continues to be 2. The slope of the new plan is 0.05, half the slope of the original plan.

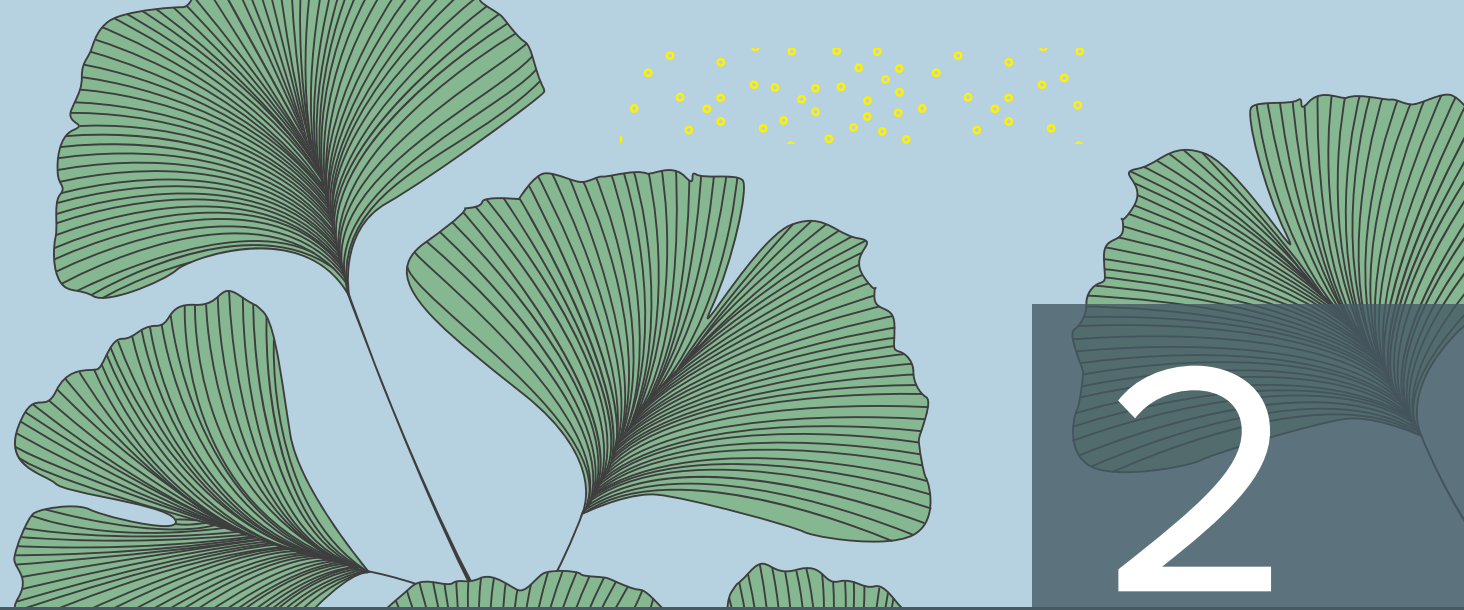


1A.5 Let the billing equation be $B = f + sT$, where f is the fee to unlock a scooter and s is the slope. From the first two points in the table, calculate the slope $s = \text{rise/run} = 1.75/5 = 0.35$. To calculate f , we can use the information in the first row of the table to write the billing equation as $2.50 = f + 0.35(5)$ and solve for $f = 0.75$. So the billing equation must be $B = 0.75 + 0.35T$. For this billing equation, the fee to unlock a scooter is 75 cents, the charge per minute is 35 cents, and the total bill for a 30-minute ride is $B = 0.75 + 0.35(30) = \$11.25$.

1A.6 Subtracting the Company 2 equation from the Company 1 equation yields the equation

$$0 = -4.8 + 0.30T \quad (\text{Company 1} - \text{Company 2}),$$

which solves for $T = 16$. So if you ride for more than 16 minutes, you'll do better with Company 2.



Supply and Demand

The stock of foodstuffs on hand at any moment in New York City's grocery stores, restaurants, and private kitchens is sufficient to feed the area's 10 million residents for at most a week or so. Since most of these residents have nutritionally adequate and highly varied diets, and since almost no food is produced within the city proper, provisioning New York requires that millions of pounds of food and drink be delivered to locations throughout the city each day.

No doubt many New Yorkers, buying groceries at their favorite local markets or eating at their favorite Italian restaurants, give little or no thought to the nearly miraculous coordination of people and resources required to feed city residents on a daily basis. But near-miraculous it is, nevertheless. Even if the supplying of New York City consisted only of transporting a fixed collection of foods to a given list of destinations each day, it would be quite an impressive operation, requiring at least a small (and well-managed) army to carry out.

Yet the entire process is astonishingly more complex than that. For example, the system must somehow ensure that not only *enough* food is delivered to satisfy New Yorkers' discriminating palates, but also the *right kinds* of food. There can't be too much pheasant and not enough smoked eel, or too much bacon and not enough eggs, or too much caviar and not enough canned tuna, and so on. Similar judgments must be made *within* each category of food and drink: there must be the right amount of Swiss cheese and the right amounts of provolone, gorgonzola, and feta.

But even this doesn't begin to describe the complexity of the decisions and actions required to provide our nation's largest city with its daily bread. Someone has to decide where each particular type of food gets produced, and how, and by whom. Someone must decide how much of each type of food gets delivered to *each* of the tens of thousands of restaurants and grocery stores in the city. Someone must determine whether the deliveries should be made in big trucks or small ones, arrange that the trucks be in the right place at the right time, and ensure that gasoline and qualified drivers be available.

LEARNING OBJECTIVES

After reading this chapter, you should be able to:

- LO1** Describe how the demand and supply curves summarize the behavior of buyers and sellers in the marketplace.
- LO2** Discuss how the supply and demand curves interact to determine equilibrium price and quantity.
- LO3** Illustrate how shifts in supply and demand curves cause prices and quantities to change.
- LO4** Explain why markets in equilibrium tend to leave no unexploited opportunities available to individuals.

Thousands of individuals must decide what role, if any, they will play in this collective effort. Some people—just the right number—must choose to drive food delivery trucks rather than trucks that deliver lumber. Others—again, just the right number—must become the mechanics who fix these trucks rather than carpenters who build houses. Others must become farmers rather than architects or bricklayers. Still others must become chefs in upscale restaurants, or flip burgers at McDonald's, instead of becoming plumbers or electricians.

Yet despite the almost incomprehensible number and complexity of the tasks involved, somehow the supplying of New York City manages to get done remarkably smoothly. Oh, a grocery store will occasionally run out of flank steak or a diner will sometimes be told that someone else has just ordered the last serving of roast duck. But if episodes like these stick in memory, it is only because they are rare. For the most part, New York's food delivery system—like that of every other city in the country—functions so seamlessly that it attracts virtually no notice.

The situation is strikingly different in New York City's rental housing market. According to one estimate, growth in the city's population has long outstripped growth in the city's supply of housing units (with the exception of a recent surge in construction of ultra-luxury condominiums).¹ As a result, America's most densely populated city has been experiencing a protracted housing shortage. Yet, paradoxically, in the midst of this shortage, apartment houses are being demolished; and in the vacant lots left behind, people from the neighborhoods are planting flower gardens!

New York City is experiencing not only a growing shortage of rental housing, but also chronically strained relations between landlords and tenants. In one all-too-typical case, for example, a photographer living in a loft on the Lower East Side waged an eight-year court battle with his landlord that generated literally thousands of pages of legal documents. "Once we put up a doorbell for ourselves," the photographer recalled, "and [the landlord] pulled it out, so we pulled out the wires to his doorbell."² The landlord, for his part, accused the photographer of obstructing his efforts to renovate the apartment. According to the landlord, the tenant preferred that the apartment remain in substandard condition since that gave him an excuse to withhold rent payments.

Same city, two strikingly different patterns: In the food industry, goods and services are available in wide variety and people (at least those with adequate income) are generally satisfied with what they receive and the choices available to them. In contrast, in the rental housing industry, chronic shortages and chronic dissatisfaction are rife among both buyers and sellers. Why this difference?

The brief answer is that New York City relies on a complex system of administrative rent regulations to allocate housing units but leaves the allocation of food essentially in the hands of market forces—the forces of supply and demand. Although intuition might suggest otherwise, both theory and experience suggest that the seemingly chaotic and unplanned outcomes of market forces, in most cases, can do a better job of allocating economic resources than can (for example) a government agency, even if the agency has the best of intentions.

In this chapter we'll explore how markets allocate food, housing, and other goods and services, usually with remarkable efficiency despite the complexity of the tasks. To be sure, markets are by no means perfect, and our stress on their virtues is to some extent an attempt to counteract what most economists view as an



Melissam/Shutterstock



Thinkstock/Jupiter Images

Why does New York City's food distribution system work so much better than its housing market?

¹For additional information see www.citylab.com/equity/2017/05/is-housing-catching-up/528246/.

²John Tierney, "The Rentocracy: At the Intersection of Supply and Demand," *New York Times Magazine*, May 4, 1997, p. 39.

underappreciation by the general public of their remarkable strengths. But, in the course of our discussion, we'll see why markets function so smoothly most of the time and why bureaucratic rules and regulations rarely work as well in solving complex economic problems.

To convey an understanding of how markets work is a major goal of this course, and in this chapter we provide only a brief introduction and overview. As the course proceeds, we'll discuss the economic role of markets in considerably more detail, paying attention to some of the problems of markets as well as their strengths.

WHAT, HOW, AND FOR WHOM? CENTRAL PLANNING VERSUS THE MARKET

No city, state, or society—regardless of how it is organized—can escape the need to answer certain basic economic questions. For example, how much of our limited time and other resources should we devote to building housing, how much to the production of food, and how much to providing other goods and services? What techniques should we use to produce each good? Who should be assigned to each specific task? And how should the resulting goods and services be distributed among people?

In the thousands of different societies for which records are available, issues like these have been decided in essentially one of two ways. One approach is for all economic decisions to be made centrally, by an individual or small number of individuals on behalf of a larger group. For example, in many agrarian societies throughout history, families or other small groups consumed only those goods and services that they produced for themselves, and a single clan or family leader made most important production and distribution decisions. On an immensely larger scale, the economic organization of the former Soviet Union (and other communist countries) was also largely centralized. In so-called centrally planned communist nations, a central bureaucratic committee established production targets for the country's farms and factories, developed a master plan for how to achieve the targets (including detailed instructions concerning who was to produce what), and set up guidelines for the distribution and use of the goods and services produced.

Neither form of centralized economic organization is much in evidence today. When implemented on a small scale, as in a self-sufficient family enterprise, centralized decision making is certainly feasible. For the reasons discussed in the preceding chapter, however, the jack-of-all-trades approach was doomed once it became clear how dramatically people could improve their living standards by specialization—that is, by having each individual focus his or her efforts on a relatively narrow range of tasks. And with the fall of the Soviet Union and its satellite nations in the late 1980s, there are now only three communist economies left in the world: Cuba, North Korea, and China. The first two of these appear to be on their last legs, economically speaking, and China has largely abandoned any attempt to control production and distribution decisions from the center. The major remaining examples of centralized allocation and control now reside in the bureaucratic agencies that administer programs like New York City's rent controls—programs that are themselves becoming increasingly rare.

At the beginning of the twenty-first century, we are therefore left, for the most part, with the second major form of economic system, one in which production and distribution decisions are left to individuals interacting in private markets. In the so-called capitalist, or free-market, economies, people decide for themselves which careers to pursue and which products to produce or buy. In fact, there are no *pure* free-market economies today. Modern industrial countries are more properly described as "mixed economies." Their goods and services are allocated by a combination of free markets, regulation, and other forms of collective control. Still, it makes sense to refer to such systems as free-market economies because people are for the most part free to start businesses, shut them down, or sell them. And within

broad limits, the distribution of goods and services is determined by individual preferences backed by individual purchasing power, which in most cases comes from the income people earn in the labor market.

In country after country, markets have replaced centralized control for the simple reason that they tend to assign production tasks and consumption benefits much more effectively. The popular press and conventional wisdom often assert that economists disagree about important issues. (As someone once quipped, “If you lay all the economists in the world end to end, they still wouldn’t reach a conclusion.”) The fact is, however, that there is overwhelming agreement among economists about a broad range of issues. A substantial majority believes that markets are the most effective means for allocating society’s scarce resources. For example, a recent survey found that more than 90 percent of American professional economists believe that rent regulations like the ones implemented by New York City do more harm than good. That the stated aim of these regulations—to make rental housing more affordable for middle- and low-income families—is clearly benign was not enough to prevent them from wreaking havoc on New York City’s housing market. To see why, we must explore how goods and services are allocated in private markets and why non-market means of allocating goods and services often do not produce the expected results.

BUYERS AND SELLERS IN MARKETS

market the market for any good consists of all buyers or sellers of that good

Beginning with some simple concepts and definitions, we will explore how the interactions among buyers and sellers in markets determine the prices and quantities of the various goods and services traded. We begin by defining a market: the **market** for any good consists of all the buyers and sellers of that good. So, for example, the market for pizza on a given day in a given place is just the set of people (or other economic actors such as firms) potentially able to buy or sell pizza at that time and location.

In the market for pizza, sellers comprise the individuals and companies that either do sell—or might, under the right circumstances, sell—pizza. Similarly, buyers in this market include all individuals who buy—or might buy—pizza.

In most parts of the country, a decent pizza can still be had for less than \$10. Where does the market price of pizza come from? Looking beyond pizza to the vast array of other goods that are bought and sold every day, we may ask, “Why are some goods cheap and others expensive?” Aristotle had no idea. Nor did Plato, or Copernicus, or Newton. On reflection, it is astonishing that, for almost the entire span of human history, not even the most intelligent and creative minds on Earth had any real inkling of how to answer that seemingly simple question. Even Adam Smith, the Scottish moral philosopher whose *Wealth of Nations* launched the discipline of economics in 1776, suffered confusion on this issue.

Smith and other early economists (including Karl Marx) thought that the market price of a good was determined by its cost of production. But although costs surely do affect prices, they cannot explain why one of Claude Monet’s paintings sells for so much more than one of Pierre-Auguste Renoir’s.

Stanley Jevons and other nineteenth-century economists tried to explain price by focusing on the value people derived from consuming different goods and services. It certainly seems plausible that people will pay a lot for a good they value highly. Yet willingness to pay cannot be the whole story, either. Deprive a person in the desert of water, for example, and he will be dead in a matter of hours, and yet water sells for less than a penny a gallon. By contrast, human



Why do Claude Monet’s paintings sell for so much more than Pierre-Auguste Renoir’s?

The Metropolitan Museum of Art, New York, H. O. Havemeyer Collection, Bequest of Mrs. H. O. Havemeyer, 1929

beings can get along perfectly well without gold, and yet gold sells for more than \$1,800 an ounce.

Cost of production? Value to the user? Which is it? The answer, which seems obvious to today's economists, is that both matter. Writing in the late nineteenth century, the British economist Alfred Marshall was among the first to show clearly how costs and value interact to determine both the prevailing market price for a good and the amount of it that is bought and sold. Our task in the pages ahead will be to explore Marshall's insights and gain some practice in applying them. As a first step, we introduce the two main components of Marshall's pathbreaking analysis: the demand curve and the supply curve.

The Demand Curve

In the market for pizza, the **demand curve** for pizza is a simple schedule or graph that tells us how many slices people would be willing to buy at different prices. By convention, economists usually put price on the vertical axis of the demand curve and quantity on the horizontal axis.

A fundamental property of the demand curve is that it is downward-sloping with respect to price. For example, the demand curve for pizza tells us that as the price of pizza falls, buyers will buy more slices. Thus, the daily demand curve for pizza in Chicago on a given day might look like the curve seen in Figure 2.1. (Although economists usually refer to demand and supply "curves," we often draw them as straight lines in examples.)

The demand curve in Figure 2.1 tells us that when the price of pizza is low—say, \$2 per slice—buyers will want to buy 16,000 slices per day, whereas they will want to buy only 12,000 slices at a price of \$3 and only 8,000 at a price of \$4. The demand curve for pizza—as for any other good—slopes downward for multiple reasons. Some have to do with the individual consumer's reactions to price changes. Thus, as pizza becomes more expensive, a consumer may switch to chicken sandwiches, hamburgers, or other foods that substitute for pizza. This is called the **substitution effect** of a price change. In addition, a price increase reduces the quantity demanded because it reduces purchasing power: a consumer simply can't afford to buy as many slices of pizza at higher prices as at lower prices. This is called the **income effect** of a price change.

Another reason the demand curve slopes downward is that consumers differ in terms of how much they're willing to pay for the good. The Cost-Benefit Principle tells us that a given person will buy the good if the benefit she expects to receive from



A Pierre-Auguste Renoir painting.

Source: Courtesy of National Gallery of Art, Washington

demand curve a schedule or graph showing the quantity of a good that buyers wish to buy at each price

substitution effect the change in the quantity demanded of a good that results because buyers switch to or from substitutes when the price of the good changes

income effect the change in the quantity demanded of a good that results because a change in the price of a good changes the buyer's purchasing power

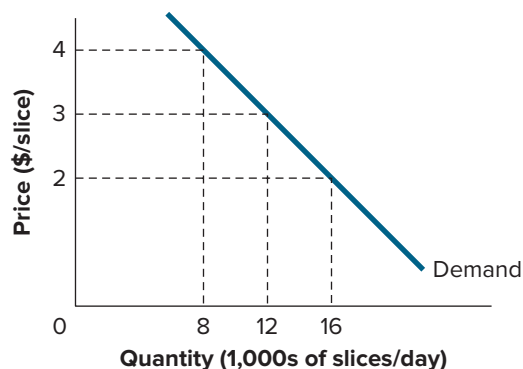


FIGURE 2.1

The Daily Demand Curve for Pizza in Chicago.

The demand curve for any good is a downward-sloping function of its price.

buyer's reservation price

the largest dollar amount the buyer would be willing to pay for a good

it exceeds its cost. The benefit is the **buyer's reservation price**, the highest dollar amount she'd be willing to pay for the good. The cost of the good is the actual amount that the buyer actually must pay for it, which is the market price of the good. In most markets, different buyers have different reservation prices. So, when the good sells for a high price, it will satisfy the cost-benefit test for fewer buyers than when it sells for a lower price.

To put this same point another way, the fact that the demand curve for a good is downward-sloping reflects the fact that the reservation price of the marginal buyer declines as the quantity of the good bought increases. Here the marginal buyer is the person who purchases the last unit of the good sold. If buyers are currently purchasing 12,000 slices of pizza a day in Figure 2.1, for example, the reservation price for the buyer of the 12,000th slice must be \$3. (If someone had been willing to pay more than that, the quantity demanded at a price of \$3 would have been more than 12,000 to begin with.) By similar reasoning, when the quantity sold is 16,000 slices per day, the marginal buyer's reservation price must be only \$2.

We defined the demand curve for any good as a schedule telling how much of it consumers wish to purchase at various prices. This is called the *horizontal interpretation* of the demand curve. Using the horizontal interpretation, we start with price on the vertical axis and read the corresponding quantity demanded on the horizontal axis. Thus, at a price of \$4 per slice, the demand curve in Figure 2.1 tells us that the quantity of pizza demanded will be 8,000 slices per day.

The demand curve also can be interpreted in a second way, which is to start with quantity on the horizontal axis and then read the marginal buyer's reservation price on the vertical axis. Thus, when the quantity of pizza sold is 8,000 slices per day, the demand curve in Figure 2.1 tells us that the marginal buyer's reservation price is \$4 per slice. This second way of reading the demand curve is called the *vertical interpretation*.

SELF-TEST 2.1

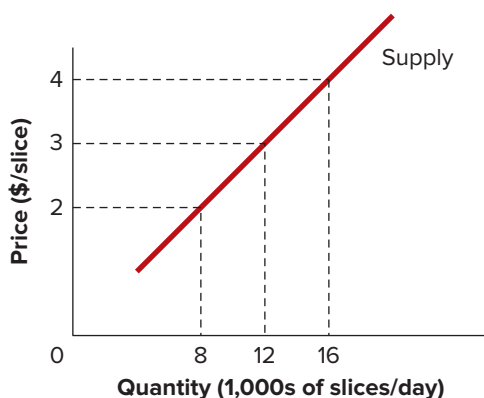
In Figure 2.1, what is the marginal buyer's reservation price when the quantity of pizza sold is 10,000 slices per day? For the same demand curve, what will be the quantity of pizza demanded at a price of \$2.50 per slice?

supply curve a graph or schedule showing the quantity of a good that sellers wish to sell at each price

The Supply Curve

In the market for pizza, the **supply curve** is a simple schedule or graph that tells us, for each possible price, the total number of slices that all pizza vendors would be willing to sell at that price. What does the supply curve of pizza look like? The answer to this question is based on the logical assumption that suppliers should be willing to sell additional slices as long as the price they receive is sufficient to cover their opportunity cost of supplying them. Thus, if what someone could earn by selling a slice of pizza is insufficient to compensate her for what she could have earned if she had spent her time and invested her money in some other way, she will not sell that slice. Otherwise, she will.

Just as buyers differ with respect to the amounts they are willing to pay for pizza, sellers also differ with respect to their opportunity cost of supplying pizza. For those with limited education and work experience, the opportunity cost of selling pizza is relatively low (because such individuals typically do not have a lot of high-paying alternatives). For others, the opportunity cost of selling pizza is of moderate value, and for still others—like rock stars and professional athletes—it is prohibitively high. In part because of these differences in opportunity cost among people, the daily supply curve of pizza will be *upward-sloping* with respect to price. As an illustration, see Figure 2.2, which shows a hypothetical supply curve for pizza in the Chicago market on a given day.

**FIGURE 2.2****The Daily Supply Curve of Pizza in Chicago.**

At higher prices, sellers generally offer more units for sale.

The reason the supply curve slopes upward may be seen as a consequence of the fact that as we expand the production of pizza, we turn first to those whose opportunity cost of producing pizza is lowest, and only then to others with a higher opportunity cost.

Like the demand curve, the supply curve can be interpreted either horizontally or vertically. Under the horizontal interpretation, we begin with a price and then go over to the supply curve to read the quantity that sellers wish to sell at that price on the horizontal axis. For instance, at a price of \$2 per slice, sellers in Figure 2.2 wish to sell 8,000 slices per day.

Under the vertical interpretation, we begin with a quantity and then go up to the supply curve to read the corresponding marginal cost on the vertical axis. Thus, if sellers in Figure 2.2 are currently supplying 12,000 slices per day, the opportunity cost of the marginal seller is \$3 per slice. In other words, the supply curve tells us that the marginal cost of producing the 12,000th slice of pizza is \$3. (If someone could produce a 12,001st slice for less than \$3, she would have an incentive to supply it, so the quantity of pizza supplied at \$3 per slice would not have been 12,000 slices per day to begin with.) By similar reasoning, when the quantity of pizza supplied is 16,000 slices per day, the marginal cost of producing another slice must be \$4. The **seller's reservation price** for selling an additional unit of a good is her marginal cost of producing that good. It is the smallest dollar amount for which she would not be worse off if she sold an additional unit.

seller's reservation price

the smallest dollar amount for which a seller would be willing to sell an additional unit, generally equal to marginal cost

SELF-TEST 2.2

In Figure 2.2, what is the marginal cost of a slice of pizza when the quantity of pizza sold is 10,000 slices per day? For the same supply curve, what will be the quantity of pizza supplied at a price of \$3.50 per slice?

RECAP ↑**DEMAND AND SUPPLY CURVES**

The *market* for a good consists of the actual and potential buyers and sellers of that good. For any given price, the *demand curve* shows the quantity that demanders would be willing to buy and the *supply curve* shows the quantity that suppliers of the good would be willing to sell. Suppliers are willing to sell more at higher prices (supply curves slope upward) and demanders are willing to buy less at higher prices (demand curves slope downward).

equilibrium a balanced or unchanging situation in which all forces at work within a system are canceled by others

equilibrium price and **equilibrium quantity** the price and quantity at the intersection of the supply and demand curves for the good

market equilibrium occurs in a market when all buyers and sellers are satisfied with their respective quantities at the market price

MARKET EQUILIBRIUM

The concept of **equilibrium** is employed in both the physical and social sciences, and it is of central importance in economic analysis. In general, a system is in equilibrium when all forces at work within the system are canceled by others, resulting in a balanced or unchanging situation. In physics, for example, a ball hanging from a spring is said to be in equilibrium when the spring has stretched sufficiently that the upward force it exerts on the ball is exactly counterbalanced by the downward force of gravity. In economics, a market is said to be in equilibrium when no participant in the market has any reason to alter his or her behavior, so that there is no tendency for production or prices in that market to change.

If we want to determine the final position of a ball hanging from a spring, we need to find the point at which the forces of gravity and spring tension are balanced and the system is in equilibrium. Similarly, if we want to find the price at which a good will sell (which we will call the **equilibrium price**) and the quantity of it that will be sold (the **equilibrium quantity**), we need to find the equilibrium in the market for that good. The basic tools for finding the equilibrium in a market for a good are the supply and demand curves for that good. For reasons we will explain, the equilibrium price and equilibrium quantity of a good are the price and quantity at which the supply and demand curves for the good intersect. For the hypothetical supply and demand curves shown earlier for the pizza market in Chicago, the equilibrium price will therefore be \$3 per slice, and the equilibrium quantity of pizza sold will be 12,000 slices per day, as shown in Figure 2.3.

Note that at the equilibrium price of \$3 per slice, both sellers and buyers are “satisfied” in the following sense: buyers are buying exactly the quantity of pizza they wish to buy at that price (12,000 slices per day) and sellers are selling exactly the quantity of pizza they wish to sell (also 12,000 slices per day). And since they are satisfied in this sense, neither buyers nor sellers face any incentives to change their behavior.

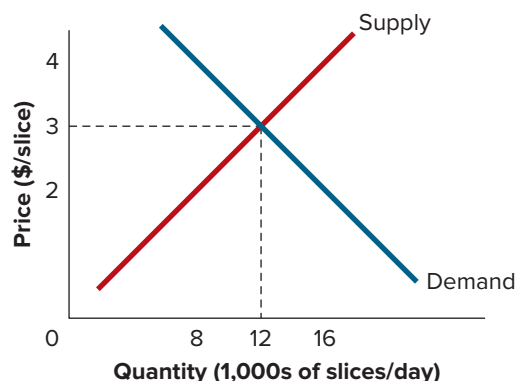
Note the limited sense of the term *satisfied* in the definition of **market equilibrium**. It doesn’t mean that sellers would be displeased to receive a price higher than the equilibrium price. Rather, it means only that they’re able to sell all they wish to sell at that price. Similarly, to say that buyers are satisfied at the equilibrium price doesn’t mean that they wouldn’t be happy to pay less than that price. Rather, it means only that they’re able to buy exactly as many units of the good as they wish to at the equilibrium price.

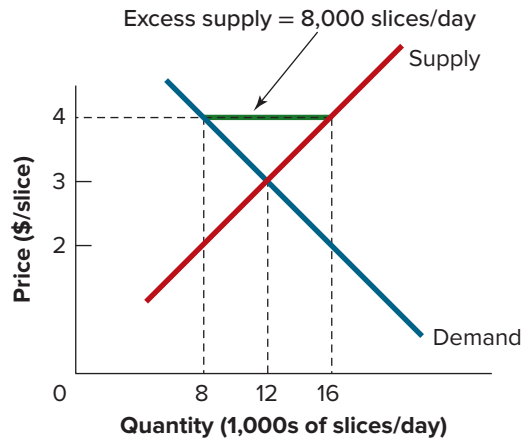
Note also that if the price of pizza in our Chicago market were anything other than \$3 per slice, either buyers or sellers would be frustrated. Suppose, for example, that the price of pizza were \$4 per slice, as shown in Figure 2.4. At that price, buyers wish to buy only 8,000 slices per day, but sellers wish to sell 16,000. And since no one can force someone to buy a slice of pizza against her wishes, this means that buyers will buy only the 8,000 slices they wish to buy. So when price exceeds the

FIGURE 2.3

The Equilibrium Price and Quantity of Pizza in Chicago.

The equilibrium price and quantity of a product are the values that correspond to the intersection of the supply and demand curves for that product.



**FIGURE 2.4****Excess Supply.**

When price exceeds equilibrium price, there is excess supply, or surplus, the difference between quantity supplied and quantity demanded.

equilibrium price, it is sellers who end up being frustrated. At a price of \$4 in this example, they are left with an **excess supply**, or **surplus**, of 8,000 slices per day.

Conversely, suppose that the price of pizza in our Chicago market were less than the equilibrium price—say, \$2 per slice. As shown in Figure 2.5, buyers want to buy 16,000 slices per day at that price, whereas sellers want to sell only 8,000. And since sellers cannot be forced to sell pizza against their wishes, this time it is the buyers who end up being frustrated. At a price of \$2 per slice in this example, they experience an **excess demand**, or **shortage**, of 8,000 slices per day.

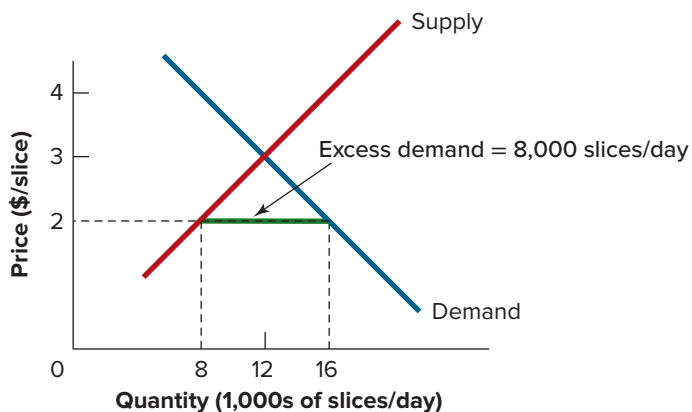
An extraordinary feature of private markets for goods and services is their automatic tendency to gravitate toward their respective equilibrium prices and quantities. The mechanisms by which the adjustment happens are implicit in our definitions of excess supply and excess demand. Suppose, for example, that the price of pizza in our hypothetical market was \$4 per slice, leading to excess supply as shown in Figure 2.4. Because sellers are frustrated in the sense of wanting to sell more pizza than buyers wish to buy, sellers have an incentive to take whatever steps they can to increase their sales. The simplest strategy available to them is to cut their price slightly. Thus, if one seller reduced her price from \$4 to, say, \$3.95 per slice, she would attract many of the buyers who had been paying \$4 per slice for pizza supplied by other sellers. Those sellers, in order to recover their lost business, would then have an incentive to match the price cut. But notice that if all sellers lowered their prices to \$3.95 per slice, there would still be considerable excess supply. So sellers would face continuing incentives to cut their prices. This pressure to cut prices won't go away until prices fall all the way to \$3 per slice.

excess supply (or surplus)

the amount by which quantity supplied exceeds quantity demanded when the price of the good exceeds the equilibrium price

excess demand (or

shortage) the amount by which quantity demanded exceeds quantity supplied when the price of a good lies below the equilibrium price

**FIGURE 2.5****Excess Demand.**

When price lies below equilibrium price, there is excess demand, the difference between quantity demanded and quantity supplied.

Conversely, suppose that price starts out less than the equilibrium price—say, \$2 per slice. This time it is buyers who are frustrated. A person who can’t get all the pizza she wants at a price of \$2 per slice has an incentive to offer a higher price, hoping to obtain pizza that would otherwise have been sold to other buyers. And sellers, for their part, will be only too happy to post higher prices as long as queues of frustrated buyers remain.

The upshot is that price has a tendency to gravitate to its equilibrium level under conditions of either excess supply or excess demand. And when price reaches its equilibrium level, both buyers and sellers are satisfied in the technical sense of being able to buy or sell precisely the amounts of their choosing.

EXAMPLE 2.1 Market Equilibrium

Samples of points on the demand and supply curves of a pizza market are provided in Table 2.1. Graph the demand and supply curves for this market and find its equilibrium price and quantity.

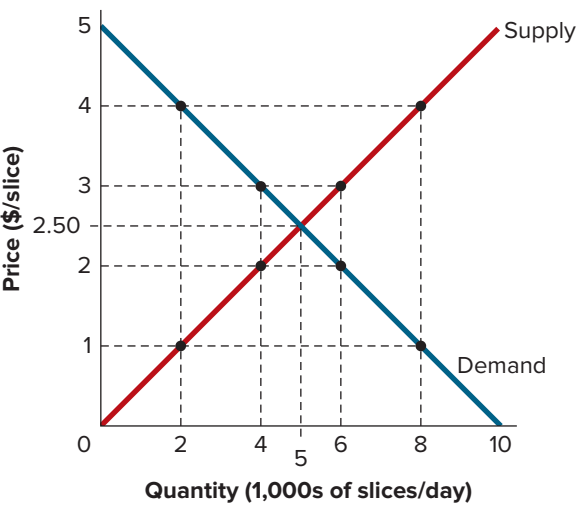
TABLE 2.1
Points along the Demand and Supply Curves of a Pizza Market

Demand for Pizza		Supply of Pizza	
Price (\$/slice)	Quantity demanded (1,000s of slices/day)	Price (\$/slice)	Quantity supplied (1,000s of slices/day)
1	8	1	2
2	6	2	4
3	4	3	6
4	2	4	8

The points in the table are plotted in Figure 2.6 and then joined to indicate the supply and demand curves for this market. These curves intersect to yield an equilibrium price of \$2.50 per slice and an equilibrium quantity of 5,000 slices per day.

FIGURE 2.6
Graphing Supply and Demand and Finding Equilibrium Price and Quantity.

To graph the demand and supply curves, plot the relevant points given in the table and then join them with a line. Equilibrium price and quantity occur at the intersection of these curves.



We emphasize that market equilibrium doesn't necessarily produce an ideal outcome for all market participants. Thus, in Example 2.1, market participants are satisfied with the amount of pizza they buy and sell at a price of \$2.50 per slice, but for a poor buyer this may signify little more than that he *can't* buy additional pizza without sacrificing other more highly valued purchases.

Indeed, buyers with extremely low incomes often have difficulty purchasing even basic goods and services, which has prompted governments in almost every society to attempt to ease the burdens of the poor. Yet the laws of supply and demand cannot simply be repealed by an act of the legislature. In the next section, we'll see that when legislators attempt to prevent markets from reaching their equilibrium prices and quantities, they often do more harm than good. Fortunately, there are other, more effective, ways of providing assistance to families in need.

Rent Controls Reconsidered

Consider again the market for rental housing units in New York City and suppose that the demand and supply curves for one-bedroom apartments are as shown in Figure 2.7. This market, left alone, would reach an equilibrium monthly rent of \$1,600, at which 2 million one-bedroom apartments would be rented. Both landlords and tenants would be satisfied, in the sense that they would not wish to rent either more or fewer units at that price.

This wouldn't necessarily mean, of course, that all is well and good. Many potential tenants, for example, might simply be unable to afford a rent of \$1,600 per month and thus be forced to remain homeless (or to move out of the city to a cheaper location). Suppose that, acting purely out of benign motives, legislators made it unlawful for landlords to charge more than \$800 per month for one-bedroom apartments. Their stated aim in enacting this law was that no person should have to remain homeless because decent housing was unaffordable.

But note in Figure 2.8 that when rents for one-bedroom apartments are prevented from rising above \$800 per month, landlords are willing to supply only 1 million apartments per month, 1 million fewer than at the equilibrium monthly rent of \$1,600. Note also that at the controlled rent of \$800 per month, tenants want to rent 3 million one-bedroom apartments per month. (For example, many people who would have decided to live in New Jersey rather than pay \$1,600 a month in New York will now choose to live in the city.) So when rents are prevented from rising above \$800 per month, we see an excess demand for one-bedroom apartments of 2 million units each month. Put another way, the rent controls result in a housing shortage of 2 million units each month. What is more, the number of apartments actually available *declines* by 1 million units per month.

If the housing market were completely unregulated, the immediate response to such a high level of excess demand would be for rents to rise sharply. But here the law

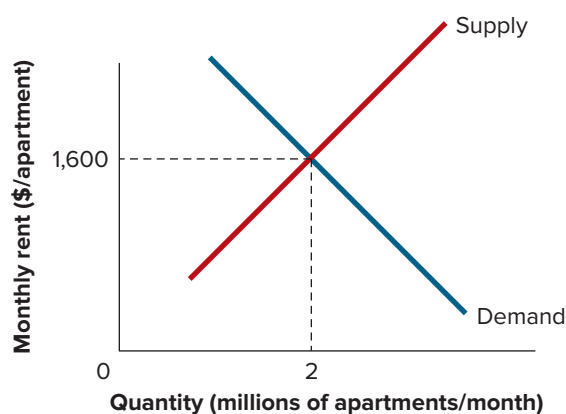


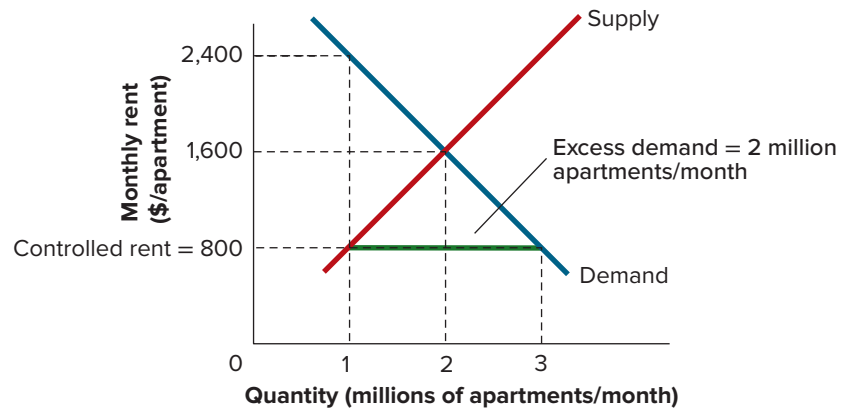
FIGURE 2.7

An Unregulated Housing Market.

For the supply and demand curves shown, the equilibrium monthly rent is \$1,600 and 2 million apartments will be rented at that price.

FIGURE 2.8**Rent Controls.**

When rents are prohibited from rising to the equilibrium level, the result is excess demand in the housing market.



prevents them from rising above \$800. Many other ways exist, however, in which market participants can respond to the pressures of excess demand. For instance, owners will quickly learn that they are free to spend less on maintaining their rental units. After all, if there are scores of renters knocking at the door of each vacant apartment, a landlord has considerable room to maneuver. Leaking pipes, peeling paint, broken furnaces, and other problems are less likely to receive prompt attention—or, indeed, any attention at all—when rents are set well below market-clearing levels.

Nor are reduced availability of apartments and poorer maintenance of existing apartments the only difficulties. With an offering of only 1 million apartments per month, we see in Figure 2.8 that there are renters who'd be willing to pay as much as \$2,400 per month for an apartment. This pressure will almost always find ways, legal or illegal, of expressing itself. In New York City, for example, it is not uncommon to see “finder’s fees” or “key deposits” as high as several thousand dollars. Owners who cannot charge a market-clearing rent for their apartments also have the option of converting them to condominiums or co-ops, which enables them to sell their assets for prices much closer to their true economic value.

Even when rent-controlled apartment owners don’t hike their prices in these various ways, serious misallocations result. For instance, ill-suited roommates often remain together despite their constant bickering because each is reluctant to reenter the housing market. Or a widow might steadfastly remain in her seven-room apartment even after her children have left home because it is much cheaper than alternative dwellings not covered by rent control. It would be much better for all concerned if she relinquished that space to a larger family that valued it more highly. But under rent controls, she has no economic incentive to do so.

There’s also another more insidious cost of rent controls. In markets without rent controls, landlords cannot discriminate against potential tenants on the basis of race, religion, sexual orientation, physical disability, or national origin without suffering an economic penalty. Refusal to rent to members of specific groups would reduce the demand for their apartments, which would mean having to accept lower rents. When rents are artificially pegged below their equilibrium level, however, the resulting excess demand for apartments enables landlords to engage in discrimination with no further economic penalty.

Rent controls are not the only instance in which governments have attempted to repeal the law of supply and demand in the interest of helping the poor. During the late 1970s, for example, the federal government tried to hold the price of gasoline below its equilibrium level out of concern that high gasoline prices imposed unacceptable hardships on low-income drivers. As with controls in the rental housing market, unintended consequences of price controls in the gasoline market made the policy an extremely costly way of trying to aid the poor. For example, gasoline shortages resulted in long lines at the pumps, a waste not only of valuable time, but also of gasoline as cars sat idling for extended periods.

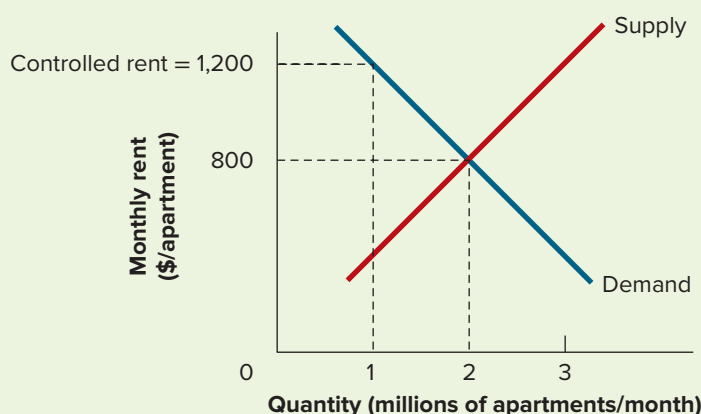
In their opposition to rent controls and similar measures, are economists revealing a total lack of concern for the poor? Although this claim is sometimes made by those who don't understand the issues, or who stand to benefit in some way from government regulations, there is little justification for it. *Economists simply realize that there are much more effective ways to help poor people than to try to give them apartments and other goods at artificially low prices.*

One straightforward approach would be to give the poor additional income and let them decide for themselves how to spend it. True, there are also practical difficulties involved in transferring additional purchasing power into the hands of the poor—most importantly, the difficulty of targeting cash to the genuinely needy without weakening others' incentives to fend for themselves. But there are practical ways to overcome this difficulty. For example, for far less than the waste caused by price controls, the government could afford generous subsidies to the wages of the working poor and could sponsor public-service employment for those who are unable to find jobs in the private sector.

Regulations that peg prices below equilibrium levels have far-reaching effects on market outcomes. The following self-test asks you to consider what happens when a price control is established at a level above the equilibrium price.

SELF-TEST 2.3

In the rental housing market whose demand and supply curves are shown below, what will be the effect of a law that prevents rents from rising above \$1,200 per month?



Pizza Price Controls?

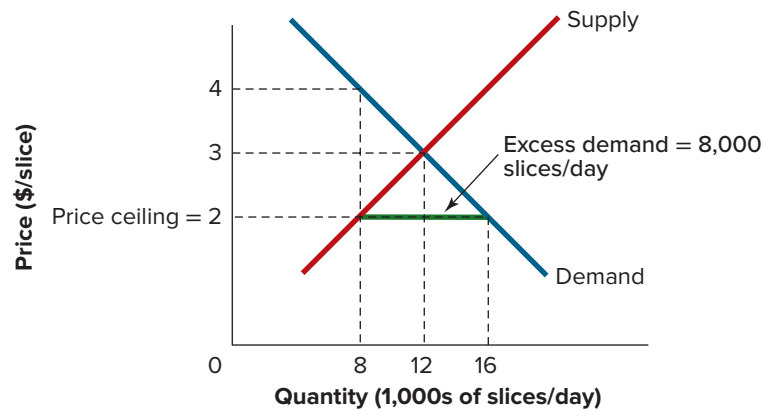
The sources of the contrast between the rent-controlled housing market and the largely unregulated food markets in New York City can be seen more vividly by trying to imagine what would happen if concern for the poor led the city's leaders to implement price controls on pizza. Suppose, for example, that the supply and demand curves for pizza are as shown in Figure 2.9 and that the city imposes a **price ceiling** of \$2 per slice, making it unlawful to charge more than that amount. At \$2 per slice, buyers want to buy 16,000 slices per day, but sellers want to sell only 8,000.

At a price of \$2 per slice, every pizza restaurant in the city will have long queues of buyers trying unsuccessfully to purchase pizza. Frustrated buyers will behave rudely to clerks, who will respond in kind. Friends of restaurant managers will begin to get preferential treatment. Devious pricing strategies will begin to emerge (such as the \$2 slice of pizza sold in combination with a \$5 cup of Coke). Pizza will be made from poorer-quality ingredients. Rumors will begin to circulate about sources of black-market pizza. And so on.

price ceiling a maximum allowable price, specified by law

FIGURE 2.9**Price Controls in the Pizza Market.**

A price ceiling below the equilibrium price of pizza would result in excess demand for pizza.



The very idea of not being able to buy a pizza seems absurd, yet precisely such things happen routinely in markets in which prices are held below the equilibrium levels. For example, prior to the collapse of communist governments, it was considered normal in those countries for people to stand in line for hours to buy bread and other basic goods, while the politically connected had first choice of those goods that were available.

RECAP ↑**MARKET EQUILIBRIUM**

Market equilibrium, the situation in which all buyers and sellers are satisfied with their respective quantities at the market price, occurs at the intersection of the supply and demand curves. The corresponding price and quantity are called the *equilibrium price* and the *equilibrium quantity*.

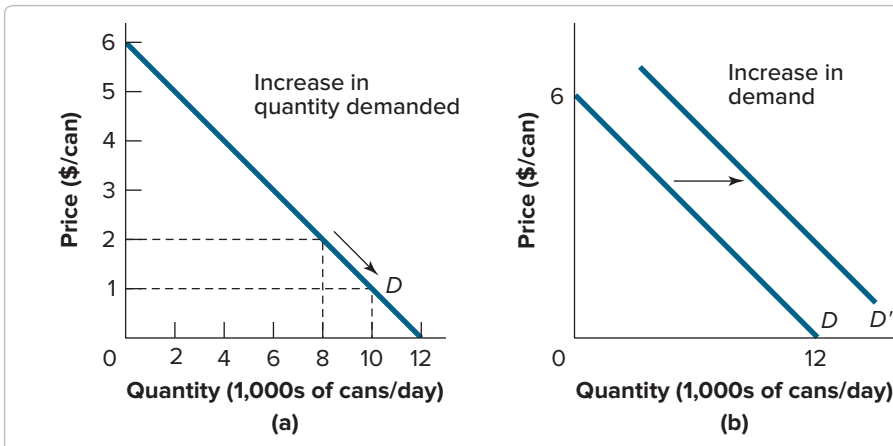
Unless prevented by regulation, prices and quantities are driven toward their equilibrium values by the actions of buyers and sellers. If the price is initially too high, so that there is excess supply, frustrated sellers will cut their price in order to sell more. If the price is initially too low, so that there is excess demand, competition among buyers drives the price upward. This process continues until equilibrium is reached.

PREDICTING AND EXPLAINING CHANGES IN PRICES AND QUANTITIES

If we know how the factors that govern supply and demand curves are changing, we can make informed predictions about how prices and the corresponding quantities will change. But when describing changing circumstances in the marketplace, we must take care to recognize some important terminological distinctions. For example, we must distinguish between the meanings of the seemingly similar expressions **change in the quantity demanded** and **change in demand**. When we speak of a “change in the quantity demanded,” this means the change in the quantity that people wish to buy that occurs in response to a change in price. For instance, Figure 2.10(a) depicts an increase in the quantity demanded that occurs in response to a reduction in the price of tuna. When the price falls from \$2 to \$1 per can, the quantity demanded rises from 8,000 to 10,000 cans per day. By contrast, when we speak of a “change in demand,” this means a *shift in the entire demand curve*. For example, Figure 2.10(b) depicts an increase in demand, meaning that at every price the quantity demanded is higher than before. In summary, a “change in the quantity demanded” refers to a movement *along* the demand curve and a “change in demand” means a *shift* of the entire curve.

change in the quantity demanded a movement along the demand curve that occurs in response to a change in price

change in demand a shift of the entire demand curve

**FIGURE 2.10****An Increase in the Quantity Demanded versus an Increase in Demand.**

(a) An increase in quantity demanded describes a downward movement along the demand curve as price falls. (b) An increase in demand describes an outward shift of the demand curve.

A similar terminological distinction applies on the supply side of the market. A **change in supply** means a shift in the entire supply curve, whereas a **change in the quantity supplied** refers to a movement along the supply curve.

Alfred Marshall's supply and demand model is one of the most useful tools of the economic naturalist. Once we understand the forces that govern the placements of supply and demand curves, we're suddenly in a position to make sense of a host of interesting observations in the world around us.

Shifts in Demand

To get a better feel for how the supply and demand model enables us to predict and explain price and quantity movements, it's helpful to begin with a few simple examples. The first one illustrates a shift in demand that results from events outside the particular market itself.

change in supply a shift of the entire supply curve

change in the quantity supplied a movement along the supply curve that occurs in response to a change in price

EXAMPLE 2.2 Complements

What will happen to the equilibrium price and quantity of tennis balls if court-rental fees decline?

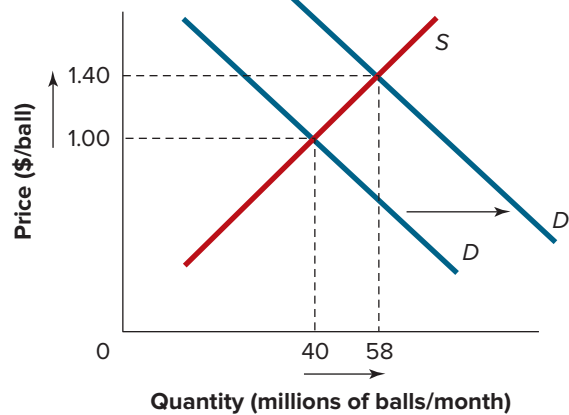
Let the initial supply and demand curves for tennis balls be as shown by the curves S and D in Figure 2.11, where the resulting equilibrium price and quantity are \$1 per ball and 40 million balls per month, respectively. Tennis courts and tennis balls are what economists call **complements**, goods that are more valuable when used in combination than when used alone. Tennis balls, for example, would be of little value if there were no tennis courts on which to play. (Tennis balls would still have *some* value even without courts—for example, to the parents who pitch them to their children for batting practice.) As tennis courts become cheaper to use, people will respond by playing more tennis, and this will increase their demand for tennis balls. A decline in court-rental fees will thus shift the demand curve for tennis balls rightward to D' . (A "rightward shift" of a demand curve also can be described as an "upward shift." These distinctions correspond, respectively, to the horizontal and vertical interpretations of the demand curve.)

Note in Figure 2.11 that, for the illustrative demand shift shown, the new equilibrium price of tennis balls, \$1.40, is higher than the original price and the new equilibrium quantity, 58 million balls per month, is higher than the original quantity.

complements two goods are complements in consumption if an increase in the price of one causes a leftward shift in the demand curve for the other (or if a decrease causes a rightward shift)

FIGURE 2.11**The Effect on the Market for Tennis Balls of a Decline in Court-Rental Fees.**

When the price of a complement falls, demand shifts right, causing equilibrium price and quantity to rise.

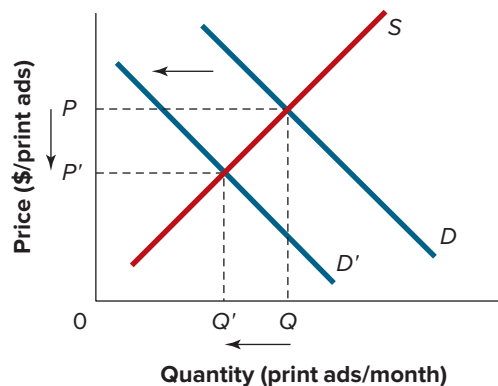
**EXAMPLE 2.3****Substitutes****What will happen to the equilibrium price and quantity of print ads as the price of digital ads falls?**

substitutes two goods are substitutes in consumption if an increase in the price of one causes a rightward shift in the demand curve for the other (or if a decrease causes a leftward shift)

Most businesses purchase a mix of digital and print ads to promote their products. Suppose the initial supply and demand curves print ads are as shown by the curves S and D in Figure 2.12, and the resulting equilibrium price and quantity are denoted P and Q . Print ads and digital ads are examples of what economists call **substitutes**, meaning that, in many applications at least, the two serve similar functions. (Many noneconomists would call them substitutes, too. Economists don't *always* choose obscure terms for important concepts!) When two goods or services are substitutes, a decrease in the price of one will cause a leftward shift in the demand curve for the other. (A “leftward shift” in a demand curve can also be described as a “downward shift.”) Diagrammatically, the demand curve for print ads shifts from D to D' in Figure 2.12.

FIGURE 2.12**The Effect on the Market for Print Ads of a Decline in the Price of Digital Ads.**

When the price of a substitute falls, demand shifts left, causing equilibrium price and quantity to fall.



As the figure shows, both the new equilibrium price, P' , and the new equilibrium quantity, Q' , are lower than the initial values, P and Q . Cheaper digital ads probably won't put print advertising companies out of business, but it will definitely lead to a drop in sales.

To summarize, economists define goods as substitutes if an increase in the price of one causes a rightward shift in the demand curve for the other. By contrast, goods are complements if an increase in the price of one causes a leftward shift in the demand curve for the other.

The concepts of substitutes and complements enable you to answer questions like the one posed in Self-Test 2.4.

SELF-TEST 2.4

How will a decline in airfares affect intercity bus fares and the price of hotel rooms in resort communities?



► Visit your instructor's Connect course and access your eBook to view this video.



Why are rotisserie chickens less expensive than fresh chickens?

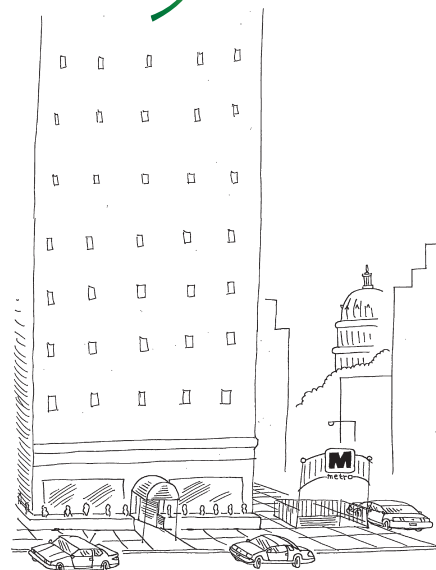
Demand curves are shifted not just by changes in the prices of substitutes and complements but also by other factors that change the amounts people are willing to pay for a given good or service. One of the most important such factors is income.

The Economic Naturalist 2.1



When the federal government implements a large pay increase for its employees, why do rents for apartments located near Washington Metro stations go up relative to rents for apartments located far away from Metro stations?

For the citizens of Washington, D.C., a substantial proportion of whom are government employees, it's more convenient to live in an apartment located 1 block from the nearest subway station than to live in one that is 20 blocks away. Conveniently located apartments thus command relatively high rents. Suppose the initial demand and supply curves for such apartments are as shown in Figure 2.13. Following a federal pay raise, some government employees who live in less convenient apartments will be willing and able to use part of their extra income to bid for more conveniently located apartments, and those who already live in such apartments will be willing and able to pay more to keep them. The effect of the pay raise is thus to shift the demand curve for conveniently located apartments to the right, as indicated by the demand curve labeled D' . As a result, both the equilibrium price and quantity of such apartments, P' and Q' , will be higher than before.



Who gets to live in the most conveniently located apartments?

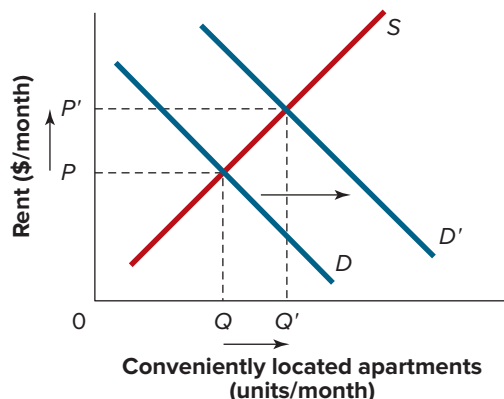


FIGURE 2.13

The Effect of a Federal Pay Raise on the Rent for Conveniently Located Apartments in Washington, D.C.

An increase in income shifts demand for a normal good to the right, causing equilibrium price and quantity to rise.

It might seem natural to ask how there could be an increase in the number of conveniently located apartments, which might appear to be fixed by the constraints of geography. But we must never underestimate the ingenuity of sellers when they confront an opportunity to make money by supplying more of something that people want. For example, if rents rose sufficiently, some landlords might respond by converting warehouse space to residential use. Or perhaps people with cars who do not place high value on living near a subway station might sell their co-op apartments to landlords, thereby freeing them for people eager to rent them. (Note that these responses constitute movements along the supply curve of conveniently located apartments, as opposed to shifts in that supply curve.)

When incomes increase, the demand curves for most goods will behave like the demand curve for conveniently located apartments, and in recognition of that fact, economists have chosen to call such goods **normal goods**.

normal good a good whose demand curve shifts rightward when the incomes of buyers increase and leftward when the incomes of buyers decrease

Not all goods are normal goods, however. In fact, the demand curves for some goods actually shift leftward when income goes up. Such goods are called **inferior goods**.

inferior good a good whose demand curve shifts leftward when the incomes of buyers increase and rightward when the incomes of buyers decrease

When would having more money tend to make you want to buy less of something? In general, this happens with goods for which there exist attractive substitutes that sell for only slightly higher prices. Apartments in unsafe, inconveniently located neighborhoods are an example. Most residents would choose to move out of such neighborhoods as soon as they could afford to, which means that an increase in income would cause the demand for such apartments to shift leftward.

SELF-TEST 2.5

How will a large pay increase for federal employees affect the rents for apartments located far away from Washington Metro stations?

Ground beef with high fat content is another example of an inferior good. For health reasons, most people prefer grades of meat with low fat content, and when they do buy high-fat meats it's usually a sign of budgetary pressure. When people in this situation receive higher incomes, they usually switch quickly to leaner grades of meat.

Preferences, or tastes, are another important factor that determines whether the purchase of a given good will satisfy the Cost-Benefit Principle. The release of each Marvel Studio film appears to kindle a powerful preference among children for Marvel action figures. When these films are released, the demand for Marvel action figures shifts sharply to the right. And the same children who can't get enough Marvel action figures suddenly lose interest in their other toys, whose respective demand curves shift to the left.

Expectations about the future are another factor that may cause demand curves to shift. If Apple users hear a credible rumor, for example, that a cheaper or significantly upgraded model will be introduced next month, the demand curve for the current model is likely to shift leftward.

Shifts in the Supply Curve

The preceding examples involved changes that gave rise to shifts in demand curves. Next, we'll look at what happens when supply curves shift. Because the supply curve is based on costs of production, anything that changes production costs will shift the supply curve, resulting in a new equilibrium quantity and price.

EXAMPLE 2.4 Increasing Opportunity Cost

What will happen to the equilibrium price and quantity of skateboards if the price of fiberglass, a substance used for making skateboards, rises?

Suppose the initial supply and demand curves for skateboards are as shown by the curves S and D in Figure 2.14, resulting in an equilibrium price and quantity of

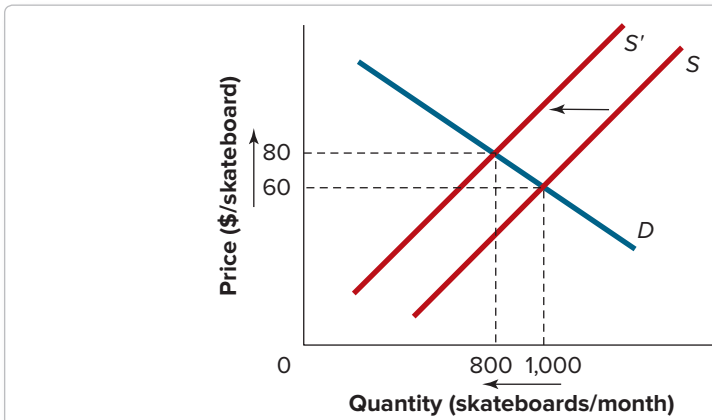


FIGURE 2.14

The Effect on the Skateboard Market of an Increase in the Price of Fiberglass.

When input prices rise, supply shifts left, causing equilibrium price to rise and equilibrium quantity to fall.

\$60 per skateboard and 1,000 skateboards per month, respectively. Since fiberglass is one of the materials used to produce skateboards, the effect of an increase in its price is to raise the marginal cost of producing skateboards. How will this affect the supply curve of skateboards?

Recall that the supply curve is upward-sloping because when the price of skateboards is low, only those potential sellers whose marginal cost of making skateboards is low can sell boards profitably, whereas at higher prices, those with higher marginal costs also can enter the market profitably. So if the cost of one of the materials used to produce skateboards rises, the number of potential sellers who can profitably sell skateboards at any given price will fall. And this, in turn, implies a leftward shift in the supply curve for skateboards. Note that a “leftward shift” in a supply curve also can be viewed as an “upward shift” in the same curve. The first corresponds to the horizontal interpretation of the supply curve, while the second corresponds to the vertical interpretation. We will use these expressions to mean exactly the same thing. The new supply curve (after the price of fiberglass rises) is the curve labeled S' in Figure 2.14.

Does an increase in the cost of fiberglass have any effect on the demand curve for skateboards? The demand curve tells us how many skateboards buyers wish to purchase at each price. Any given buyer is willing to purchase a skateboard if her reservation price for it exceeds its market price. And since each buyer’s reservation price, which is based on the benefits of owning a skateboard, does not depend on the price of fiberglass, there should be no shift in the demand curve for skateboards.

In Figure 2.14, we can now see what happens when the supply curve shifts leftward and the demand curve remains unchanged. For the illustrative supply curve shown, the new equilibrium price of skateboards, \$80, is higher than the original price, and the new equilibrium quantity, 800 per month, is lower than the original quantity. (These new equilibrium values are merely illustrative. There is insufficient information provided in the example to determine their exact values.) People who don’t place a value of at least \$80 on owning a skateboard will choose to spend their money on something else.

The effects on equilibrium price and quantity run in the opposite direction whenever marginal costs of production decline, as illustrated in the next example.

EXAMPLE 2.5 Reduction of Marginal Cost

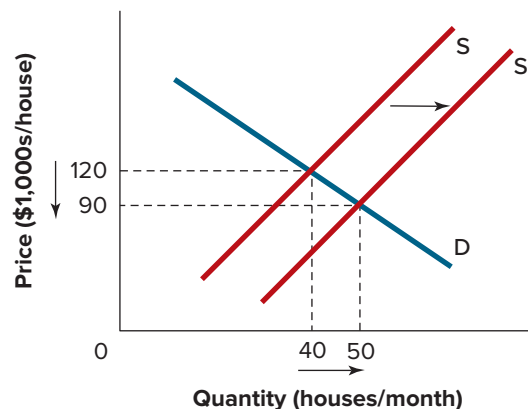
What will happen to the equilibrium price and quantity of new houses if the wage rate of carpenters falls?

Suppose the initial supply and demand curves for new houses are as shown by the curves S and D in Figure 2.15, resulting in an equilibrium price of

FIGURE 2.15

The Effect on the Market for New Houses of a Decline in Carpenters' Wage Rates.

When input prices fall, supply shifts right, causing equilibrium price to fall and equilibrium quantity to rise.



\$120,000 per house and an equilibrium quantity of 40 houses per month, respectively. A decline in the wage rate of carpenters reduces the marginal cost of making new houses, and this means that, for any given price of houses, more builders can profitably serve the market than before. Diagrammatically, this means a rightward shift in the supply curve of houses, from S to S' . (A “rightward shift” in the supply curve also can be described as a “downward shift.”)

Does a decrease in the wage rate of carpenters have any effect on the demand curve for houses? The demand curve tells us how many houses buyers wish to purchase at each price. Because carpenters are now earning less than before, the maximum amount that they are willing to pay for houses may fall, which would imply a leftward shift in the demand curve for houses. But because carpenters make up only a tiny fraction of all potential home buyers, we may assume that this shift is negligible. Thus, a reduction in carpenters' wages produces a significant rightward shift in the supply curve of houses, but no appreciable shift in the demand curve.

We see from Figure 2.15 that the new equilibrium price, \$90,000 per house, is lower than the original price and the new equilibrium quantity, 50 houses per month, is higher than the original quantity.

Examples 2.4 and 2.5 involved changes in the cost of a material, or input, in the production of the good in question—fiberglass in the production of skateboards and carpenters' labor in the production of houses. As the following example illustrates, supply curves also shift when technology changes.

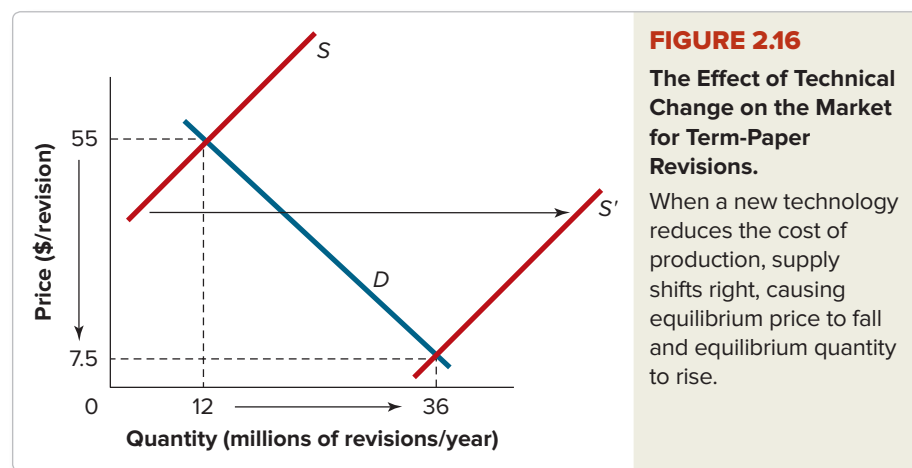
The Economic Naturalist 2.2



Why do major term papers go through so many more revisions today than in the 1970s?

Students in the dark days before word processors were in widespread use could not make even minor revisions in their term papers without having to retype their entire manuscript from scratch. The availability of word-processing technology has, of course, radically changed the picture. Instead of having to retype the entire draft, now only the changes need be entered.

In Figure 2.16, the curves labeled S and D depict the supply and demand curves for revisions in the days before word processing, and the curve S' depicts the supply curve for revisions today. As the diagram shows, the result is not only a sharp decline in the price per revision, but also a corresponding increase in the equilibrium number of revisions.



Why does written work go through so many more revisions now than in the 1970s?

Note that in The Economic Naturalist 2.2 we implicitly assumed that students purchased typing services in a market. In fact, however, many students type their own term papers. Does that make a difference? Even if no money actually changes hands, students pay a price when they revise their term papers—namely, the opportunity cost of the time it takes to perform that task. Because technology has radically reduced that cost, we would expect to see a large increase in the number of term-paper revisions even if most students type their own work.

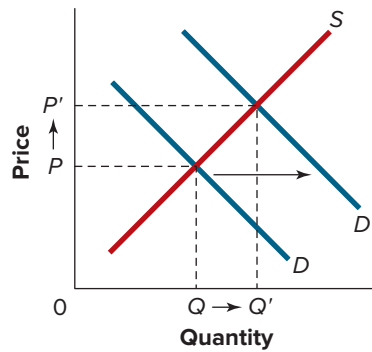
Changes in input prices and technology are two of the most important factors that give rise to shifts in supply curves. In the case of agricultural commodities, weather may be another important factor, with favorable conditions shifting the supply curves of such products to the right and unfavorable conditions shifting them to the left. (Weather also may affect the supply curves of nonagricultural products through its effects on the national transportation system.) Expectations of future price changes also may shift current supply curves, as when the expectation of poor crops from a current drought causes suppliers to withhold supplies from existing stocks in the hope of selling at higher prices in the future. Changes in the number of sellers in the market also can cause supply curves to shift.

Four Simple Rules

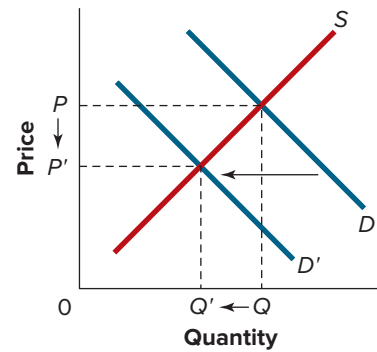
For supply and demand curves that have the conventional slopes (upward-sloping for supply curves, downward-sloping for demand curves), the preceding examples illustrate the four basic rules that govern how shifts in supply and demand affect equilibrium prices and quantities. These rules are summarized in Figure 2.17.

FIGURE 2.17

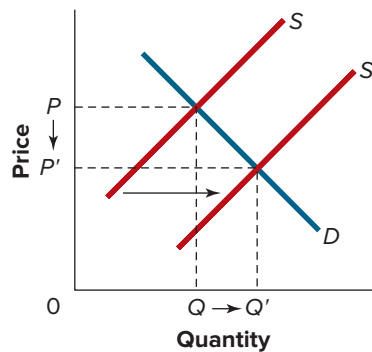
Four Rules Governing the Effects of Supply and Demand Shifts.



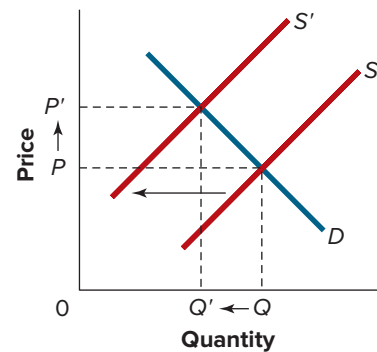
An increase in demand will lead to an increase in both the equilibrium price and quantity.



A decrease in demand will lead to a decrease in both the equilibrium price and quantity.



An increase in supply will lead to a decrease in the equilibrium price and an increase in the equilibrium quantity.



A decrease in supply will lead to an increase in the equilibrium price and a decrease in the equilibrium quantity.

RECAP ↑

FACTORS THAT SHIFT SUPPLY AND DEMAND

Factors that cause an increase (rightward or upward shift) in demand:

1. A decrease in the price of complements to the good or service.
2. An increase in the price of substitutes for the good or service.
3. An increase in income (for a normal good).
4. An increased preference by demanders for the good or service.
5. An increase in the population of potential buyers.
6. An expectation of higher prices in the future.

When these factors move in the opposite direction, demand will shift left.

Factors that cause an increase (rightward or downward shift) in supply:

1. A decrease in the cost of materials, labor, or other inputs used in the production of the good or service.
2. An improvement in technology that reduces the cost of producing the good or service.
3. An improvement in the weather (especially for agricultural products).
4. An increase in the number of suppliers.
5. An expectation of lower prices in the future.

When these factors move in the opposite direction, supply will shift left.

The qualitative rules summarized in Figure 2.17 hold for supply or demand shifts of any magnitude, provided the curves have their conventional slopes. But as the next example demonstrates, when both supply and demand curves shift at the same time, the direction in which equilibrium price or quantity changes will depend on the relative magnitudes of the shifts.

EXAMPLE 2.6 Shifts in Supply and Demand

How do shifts in *both* demand and supply affect equilibrium quantities and prices?

What will happen to the equilibrium price and quantity in the corn tortilla chip market if both of the following events occur: (1) researchers prove that the oils in which tortilla chips are fried are harmful to human health and (2) the price of corn harvesting equipment falls?

The conclusion regarding the health effects of the oils will shift the demand for tortilla chips to the left because many people who once bought chips in the belief that they were healthful will now switch to other foods. The decline in the price of harvesting equipment will shift the supply of chips to the right because additional farmers will now find it profitable to enter the corn market. In Figure 2.18(a) and 2.18(b), the original supply and demand curves are denoted by S and D , while the new curves are denoted by S' and D' . Note that in both panels the shifts lead to a decline in the equilibrium price of chips.

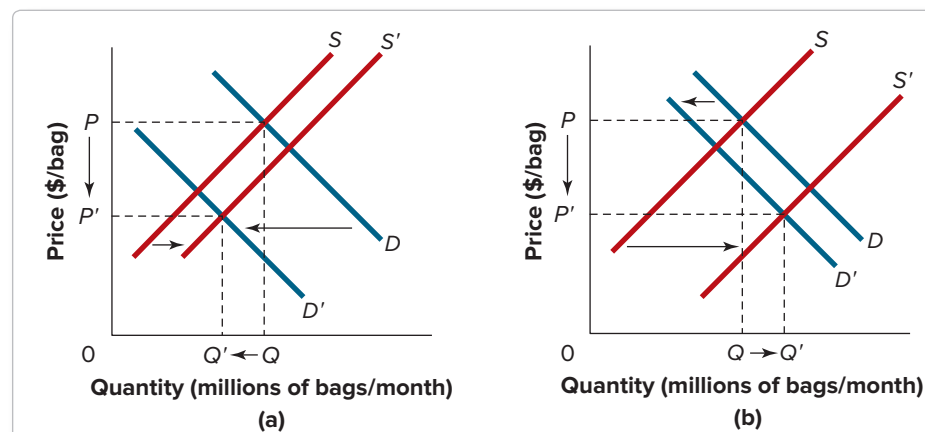


FIGURE 2.18

The Effects of Simultaneous Shifts in Supply and Demand.

When demand shifts left and supply shifts right, equilibrium price falls, but equilibrium quantity may either rise (b) or fall (a).

But note also that the effect of the shifts on equilibrium quantity cannot be determined without knowing their relative magnitudes. Taken separately, the demand shift causes a decline in equilibrium quantity, whereas the supply shift causes an increase in equilibrium quantity. The net effect of the two shifts thus depends on which of the individual effects is larger. In Figure 2.18(a), the demand shift dominates, so equilibrium quantity declines. In Figure 2.18(b), the supply shift dominates, so equilibrium quantity goes up.

The following Self-Test 2.6 asks you to consider a simple variation on the problem posed in the previous example.

SELF-TEST 2.6

What will happen to the equilibrium price and quantity in the corn tortilla chip market if both of the following events occur: (1) researchers discover that a vitamin found in corn helps protect against cancer and heart disease and (2) a swarm of locusts destroys part of the corn crop?



The Economic Naturalist 2.3

Why do the prices of some goods, like airline tickets to Europe, go up during the months of heaviest consumption, while others, like sweet corn, go down?

Seasonal price movements for airline tickets are primarily the result of seasonal variations in demand. Thus, ticket prices to Europe are highest during the summer months because the demand for tickets is highest during those months, as shown in Figure 2.19(a), where the w and s subscripts denote winter and summer values, respectively.



Why are some goods cheapest during the months of heaviest consumption, while others are most expensive during those months?

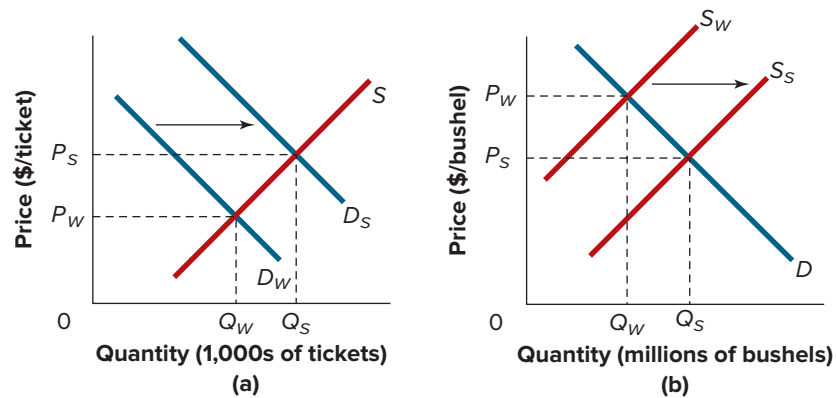


FIGURE 2.19

Seasonal Variation in the Air Travel and Corn Markets.

(a) Prices are highest during the period of heaviest consumption when heavy consumption is the result of high demand. (b) Prices are lowest during the period of heaviest consumption when heavy consumption is the result of high supply.

By contrast, seasonal price movements for sweet corn are primarily the result of seasonal variations in supply. The price of sweet corn is lowest in the summer months because its supply is highest during those months, as seen in Figure 2.19(b).

EFFICIENCY AND EQUILIBRIUM

Markets represent a highly effective system of allocating resources. When a market for a good is in equilibrium, the equilibrium price conveys important information to potential suppliers about the value that potential demanders place on that good. At the same time, the equilibrium price informs potential demanders about the opportunity cost of supplying the good. This rapid, two-way transmission of information is the reason that markets can coordinate an activity as complex as supplying New York City with food and drink, even though no one person or organization oversees the process.

But are the prices and quantities determined in market equilibrium socially optimal, in the sense of maximizing total economic surplus? That is, does equilibrium in unregulated markets always maximize the difference between the total benefits and total costs experienced by market participants? As we'll see, the answer is it depends: a market that is out of equilibrium, such as the rent-controlled New York housing market, always creates opportunities for individuals to arrange transactions that will increase their individual economic surplus. As we'll also see,

however, a market for a good that is in equilibrium makes the largest possible contribution to total economic surplus only when its supply and demand curves fully reflect all costs and benefits associated with the production and consumption of that good.

Cash on the Table

In economics we assume that all exchange is purely voluntary. This means that a transaction cannot take place unless the buyer's reservation price for the good exceeds the seller's reservation price. When that condition is met and a transaction takes place, both parties receive an economic surplus. The **buyer's surplus** from the transaction is the difference between his or her reservation price and the price he or she actually pays. The **seller's surplus** is the difference between the price he or she receives and his or her reservation price. The **total surplus** from the transaction is the sum of the buyer's surplus and the seller's surplus. It is also equal to the difference between the buyer's reservation price and the seller's reservation price.

Suppose there is a potential buyer whose reservation price for an additional slice of pizza is \$4 and a potential seller whose reservation price is only \$2. If this buyer purchases a slice of pizza from this seller for \$3, the total surplus generated by this exchange is $\$4 - \$2 = \$2$, of which $\$4 - \$3 = \$1$ is the buyer's surplus and $\$3 - \$2 = \$1$ is the seller's surplus.

A regulation that prevents the price of a good from reaching its equilibrium level unnecessarily prevents exchanges of this sort from taking place, and in the process reduces total economic surplus. Consider again the effect of price controls imposed in the market for pizza. The demand curve in Figure 2.20 tells us that if a price ceiling of \$2 per slice were imposed, only 8,000 slices of pizza per day would be sold. At that quantity, the vertical interpretations of the supply and demand curves tell us that a buyer would be willing to pay as much as \$4 for an additional slice and that a seller would be willing to sell one for as little as \$2. The difference—\$2 per slice—is the additional economic surplus that would result if an additional slice were produced and sold. As noted earlier, an extra slice sold at a price of \$3 would result in an additional \$1 of economic surplus for both buyer and seller.

When a market is out of equilibrium, it's always possible to identify mutually beneficial exchanges of this sort. When people have failed to take advantage of all mutually beneficial exchanges, we often say that there's "**cash on the table**"—the economist's metaphor for unexploited opportunities. When the price in a market is below the equilibrium price, there's cash on the table because the reservation price of sellers (marginal cost) will always be lower than the reservation price of buyers. In the absence of a law preventing buyers from paying more than \$2 per slice,

buyer's surplus the difference between the buyer's reservation price and the price he or she actually pays

seller's surplus the difference between the price received by the seller and his or her reservation price

total surplus the difference between the buyer's reservation price and the seller's reservation price

cash on the table an economic metaphor for unexploited gains from exchange

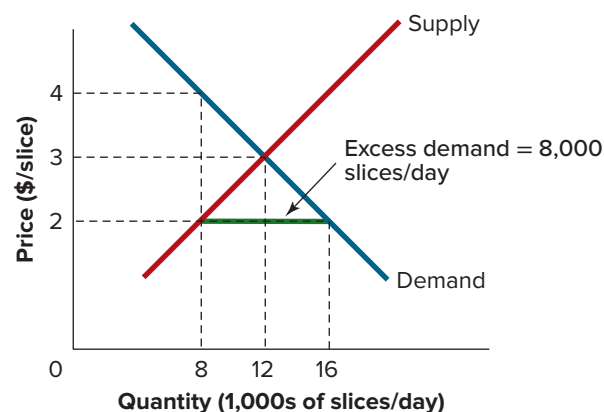


FIGURE 2.20
Price Controls in the Pizza Market.

A price ceiling below the equilibrium price of pizza would result in excess demand for pizza.

restaurant owners would quickly raise their prices and expand their production until the equilibrium price of \$3 per slice were reached. At that price, buyers would be able to get precisely the 12,000 slices of pizza they want to buy each day. All mutually beneficial opportunities for exchange would have been exploited, leaving no more cash on the table.

It should be no surprise that buyers and sellers in the marketplace have an uncanny ability to detect the presence of cash on the table. It is almost as if unexploited opportunities give off some exotic scent triggering neurochemical explosions in the olfactory centers of their brains. The desire to scrape cash off the table and into their pockets is what drives sellers in each of New York City's thousands of individual food markets to work diligently to meet their customers' demands. That they succeed to a far higher degree than participants in the city's rent-controlled housing market is plainly evident. Whatever flaws it might have, the market system moves with considerably greater speed and agility than any centralized allocation mechanisms yet devised. But as we emphasize in the following section, this does not mean that markets *always* lead to the greatest good for all.

Smart for One, Dumb for All

socially optimal quantity

the quantity of a good that results in the maximum possible economic surplus from producing and consuming the good

efficiency (or **economic efficiency**) a condition that occurs when all goods and services are produced and consumed at their respective socially optimal levels

The **socially optimal quantity** of any good is the quantity that maximizes the total economic surplus that results from producing and consuming the good. From the Cost-Benefit Principle, we know that we should keep expanding production of the good as long as its marginal benefit is at least as great as its marginal cost. This means that the socially optimal quantity is that level for which the marginal cost and marginal benefit of the good are the same.

When the quantity of a good is less than the socially optimal quantity, boosting its production will increase total economic surplus. By the same token, when the quantity of a good exceeds the socially optimal quantity, reducing its production will increase total economic surplus. **Efficiency**, or **economic efficiency**, occurs when all goods and services in the economy are produced and consumed at their respective socially optimal levels.

Efficiency is an important social goal. Failure to achieve efficiency means that total economic surplus is smaller than it could have been. Movements toward efficiency make the total economic pie larger, making it possible for everyone to have a larger slice.

Is the market equilibrium quantity of a good efficient? That is, does it maximize the total economic surplus received by participants in the market for that good? When the private market for a given good is in equilibrium, we can say that the cost *to the seller* of producing an additional unit of the good is the same as the benefit *to the buyer* of having an additional unit. If all costs of producing the good are borne directly by sellers, and if all benefits from the good accrue directly to buyers, it follows that the market equilibrium quantity of the good will equate the marginal cost and marginal benefit of the good. And this means that the equilibrium quantity also maximizes total economic surplus.

But sometimes the production of a good entails costs that fall on people other than those who sell the good. This will be true, for instance, for goods whose production generates significant levels of environmental pollution. As extra units of these goods are produced, the extra pollution harms other people besides sellers. In the market equilibrium for such goods, the benefit *to buyers* of the last good produced is, as before, equal to the cost incurred by sellers to produce that good. But since producing that good also imposes pollution costs on others, we know that the *full* marginal cost of the last unit produced—the seller's private marginal cost plus the marginal pollution cost borne by others—must be higher than the benefit of the last unit produced. So, in this case, the market equilibrium quantity of the good will be larger than the socially optimal quantity. Total economic surplus would be higher if output of the good were lower. Yet neither sellers nor buyers have any incentive to alter their behavior.

The Economic Naturalist 2.4



Why was there a shortage of toilet paper during the COVID-19 pandemic?

During the COVID-19 pandemic many consumers in search of toilet paper arrived at their local grocery stores only to be faced with empty shelves. If prices are supposed to adjust to equate the quantity supplied and the quantity demanded, then why was there a shortage of toilet paper during the pandemic?

Faced with the public health crisis brought about COVID-19 many, if not most, consumers in the United States stocked up on essential items such as toilet paper. Indeed, some consumers purchased excessive quantities of these items, and this type of hoarding behavior received much of the blame for the shortage of toilet paper during the pandemic.

But there is another important factor as well. The toilet paper market in the United States is divided into two distinct segments. There is a market for the lower-quality commercial-grade toilet paper, typically used in restaurants, schools, workplaces and other public venues, and a separate market for the higher-quality consumer-grade toilet paper, generally used in people's homes. Even though the overall use of toilet paper did not substantially change during the pandemic, there was a large shift away from commercial-grade toilet paper toward consumer-grade toilet paper. That is, the widespread stay-at-home orders led to a sharp decrease in the demand for commercial-grade toilet paper, and an equally abrupt increase in the demand for consumer-grade toilet paper, and toilet paper producers, who are accustomed to relatively stable demand, had difficulty adjusting to this rapid change.

Of course, in normal times, an increase in demand would not lead to a shortage. Instead we would expect prices to rise. During times of emergency, however, both social norms and price gouging laws limit retailers' ability to raise prices. Thus while the demand for consumer-grade toilet paper surged during the pandemic (both because of hoarding and the shift away from commercial-grade toilet paper), prices did not adjust, leading to empty shelves and anxious consumers.

Another possibility is that people other than those who buy a good may receive significant benefits from it. For instance, when someone purchases a vaccination against measles from her doctor, she not only protects herself, but also makes it less likely that others will catch this disease. From the perspective of society as a whole, we should keep increasing the number of vaccinations until their marginal cost equals their marginal benefit. The marginal benefit of a vaccination is the value of the protection it provides the person vaccinated *plus* the value of the protection it provides all others. Private consumers, however, will choose to be vaccinated only if the marginal benefit *to them* exceeds the price of the vaccination. In this case, then, the market equilibrium quantity of vaccinations will be smaller than the quantity that maximizes total economic surplus. Again, however, individuals would have no incentive to alter their behavior.

Situations like the ones just discussed provide examples of behaviors that we may call "smart for one but dumb for all." In each case, the individual actors are behaving rationally. They are pursuing their goals as best they can, and yet there remain unexploited opportunities for gain from the point of view of the whole society. The difficulty is that these opportunities cannot be exploited by individuals acting alone. In subsequent chapters, we will see how people can often organize collectively to exploit such opportunities.

RECAP ↑**EFFICIENCY AND EQUILIBRIUM**

When the supply and demand curves for a good reflect all significant costs and benefits associated with the production and consumption of that good, the market equilibrium will result in the largest possible economic surplus. But if people other than buyers benefit from the good, or if people other than sellers bear costs because of it, market equilibrium need not result in the largest possible economic surplus.

SUMMARY

- The demand curve is a downward-sloping line that tells what quantity buyers will demand at any given price. The supply curve is an upward-sloping line that tells what quantity sellers will offer at any given price. (LO1)
- Alfred Marshall's model of supply and demand explains why neither cost of production nor value to the purchaser (as measured by willingness to pay) is, by itself, sufficient to explain why some goods are cheap and others are expensive. To explain variations in price, we must examine the interaction of cost and willingness to pay. As we've seen in this chapter, goods differ in price because of differences in their respective supply and demand curves. (LO2)
- Market equilibrium occurs when the quantity buyers demand at the market price is exactly the same as the quantity that sellers offer. The equilibrium price–quantity pair is the one at which the demand and supply curves intersect. In equilibrium, market price measures both the value of the last unit sold to buyers and the cost of the resources required to produce it. (LO2)
- When the price of a good lies above its equilibrium value, there is an excess supply of that good. Excess supply motivates sellers to cut their prices and price continues to fall until equilibrium price is reached. When price lies below its equilibrium value, there is excess demand. With excess demand, frustrated buyers are motivated to offer higher prices and the upward pressure on prices persists until equilibrium is reached. A remarkable feature of the market system is that, relying only on the tendency of people to respond in self-interested ways to market price signals, it somehow manages to coordinate the actions of literally billions of buyers and sellers worldwide. When excess demand or excess supply occurs, it tends to be small and brief, except in markets where regulations prevent full adjustment of prices. (LO2)
- The basic supply and demand model is a primary tool of the economic naturalist. Changes in the equilibrium price of a good, and in the amount of it traded in the marketplace, can be predicted on the basis of shifts in its supply or demand curves. The following four rules hold for any good with a downward-sloping demand curve and an upward-sloping supply curve.
 1. An increase in demand will lead to an increase in equilibrium price and quantity.
 2. A reduction in demand will lead to a reduction in equilibrium price and quantity.
 3. An increase in supply will lead to a reduction in equilibrium price and an increase in equilibrium quantity.
 4. A decrease in supply will lead to an increase in equilibrium price and a reduction in equilibrium quantity. (LO3)
- Incomes, tastes, population, expectations, and the prices of substitutes and complements are among the factors that shift demand schedules. Supply schedules, in turn, are primarily governed by such factors as technology, input prices, expectations, the number of sellers, and, especially for agricultural products, the weather. (LO3)
- The efficiency of markets in allocating resources does not eliminate social concerns about how goods and services are distributed among different people. For example, we often lament the fact many buyers enter the market with too little income to buy even the most basic goods and services. Concern for the well-being of the poor has motivated many governments to intervene in a variety of ways to alter the outcomes of market forces. Sometimes these interventions take the form of laws that peg prices below their equilibrium levels. Such laws almost invariably generate harmful, if unintended, consequences. Programs like rent-control laws, for example, lead to severe housing shortages, black marketeering, and a rapid deterioration of the relationship between landlords and tenants. (LO4)
- If the difficulty is that the poor have too little money, the best solution is to discover ways of boosting their incomes directly. The law of supply and demand cannot be repealed by the legislature. But legislatures do have the capacity to alter the underlying forces that

govern the shape and position of supply and demand schedules. (LO4)

- When the supply and demand curves for a good reflect all significant costs and benefits associated with the production and consumption of that good, the market equilibrium price will guide people to produce and consume the quantity of the good that results in the largest possible economic surplus. This conclusion, however, does not apply if others, besides buyers, benefit from the good (as when someone benefits from his neighbor's purchase of a vaccination against measles) or if others besides sellers bear costs because of the good (as when its production generates pollution). In such cases, market equilibrium does not result in the greatest gain for all. (LO4)

KEY TERMS

buyer's reservation price
buyer's surplus
cash on the table
change in demand
change in supply
change in the quantity demanded
change in the quantity supplied
complements

demand curve
efficiency (or economic efficiency)
equilibrium
equilibrium price and equilibrium quantity
excess demand (or shortage)
excess supply (or surplus)
income effect
inferior good
market

market equilibrium
normal good
price ceiling
seller's reservation price
seller's surplus
socially optimal quantity
substitutes
substitution effect
supply curve
total surplus

REVIEW QUESTIONS

1. Explain the distinction between the horizontal and vertical interpretations of the demand curve. (LO1)
2. Why isn't knowing the cost of producing a good sufficient to predict its market price? (LO2)
3. In recent years, a government official proposed that gasoline price controls be imposed to protect the poor from rising gasoline prices. What evidence could you consult to discover whether this proposal was enacted? (LO2)
4. Distinguish between the meaning of the expressions "change in demand" and "change in the quantity demanded." (LO3)
5. Give an example of behavior you have observed that could be described as "smart for one but dumb for all." (LO4)

PROBLEMS



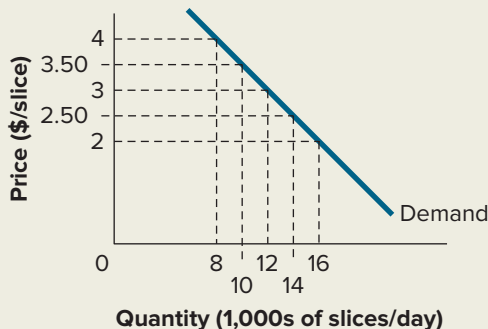
1. How would each of the following affect the U.S. market supply curve for corn? (LO1)
 - a. A new and improved crop rotation technique is discovered.
 - b. The price of fertilizer falls.
 - c. The government offers new tax breaks to farmers.
 - d. A tornado sweeps through Iowa.
2. Indicate how you think each of the following would shift demand in the indicated market. (LO1)
 - a. The incomes of buyers in the market for Adirondack vacations increase.
 - b. Buyers in the market for pizza read a study linking pepperoni consumption to heart disease.
 - c. Buyers in the market for gas-powered cars learn of an increase in the price of electric cars (a substitute for gas-powered cars).
 - d. Buyers in the market for electric cars learn of an increase in the price of electric cars.
3. An Arizona student claims to have spotted a UFO over the desert outside of Tucson. How will this claim affect the *supply* (not the quantity supplied) of binoculars in Tucson stores? (LO1)
4. Suppose that when milk sells for \$4.50 per gallon, the quantity of milk demanded is 3,250 gallons per day and the quantity of milk supplied is 3,860 gallons per day. Will the equilibrium price of milk be greater than, less than, or equal to \$4.50 per gallon? Explain. (LO2)
5. State whether the following pairs of goods are complements, or substitutes, or both. (LO3)
 - a. Washing machines and dryers.
 - b. Tennis rackets and tennis balls.

- c. Birthday cake and birthday candles.
 - d. Cloth diapers and disposable diapers.
6. How will an increase in the birth rate affect the equilibrium price of land? (LO3)
 7. What will happen to the equilibrium price and quantity of beef if the price of chickenfeed increases (assume that chicken and beef are substitutes)? (LO3)
 8. How will a new law mandating an increase in required levels of automobile insurance affect the equilibrium price and quantity in the market for new automobiles? (LO3)
 9. Predict what will happen to the equilibrium price and quantity of oranges if the following events take place. (LO3)
 - a. A study finds that a daily glass of orange juice reduces the risk of heart disease.
 - b. The price of grapefruit falls drastically.
 - c. The wage paid to orange pickers rises.
 - d. Exceptionally good weather provides a much greater than expected harvest.
 10. Suppose the current issue of *The New York Times* reports an outbreak of mad cow disease in Nebraska, as well as the discovery of a new breed of chicken that gains more weight than existing breeds that consume the same amount of food. How will these developments affect the equilibrium price and quantity of chickens sold in the United States? (LO3)
 11. Twenty-five years ago, tofu was available only from small businesses operating in predominantly Asian sections of large cities. Today tofu has become popular as a high-protein health food and is widely available in supermarkets throughout the United States.

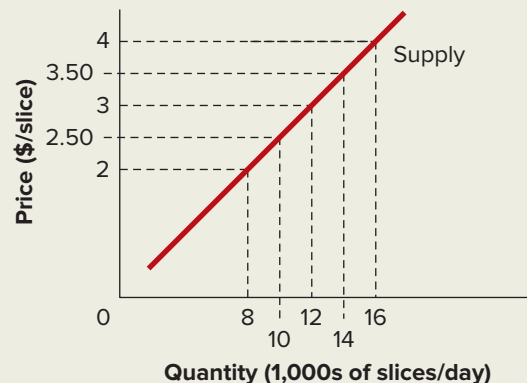
At the same time, tofu production has evolved to become factory-based using modern food-processing technologies. Draw a diagram with demand and supply curves depicting the market for tofu 25 years ago and the market for tofu today. Given the information above, what does the demand–supply model predict about changes in the volume of tofu sold in the United States between then and now? What does it predict about changes in the price of tofu? (LO3)
 12. For each of the following statements about a market equilibrium identify whether the statement is always true, never true, or sometimes true. (LO4)
 - a. A market equilibrium maximizes total economic surplus.
 - b. A market equilibrium exploits all gains achievable through collective action.
 - c. A market equilibrium leaves no unexploited opportunities for individuals.
 13. In March 2020, global crude oil prices tumbled from over \$50 a barrel to below \$23 per barrel, bringing prices to their lowest level in nearly two decades. This precipitous drop in crude oil prices was fueled by two major shocks to the oil market. First, the COVID-19 pandemic led to a massive reduction in all forms of travel as large numbers of people around the world were advised to shelter in place. Second, Russia, Saudi Arabia, and a number of other major oil-producing nations typically reach collective agreements to limit the world supply of oil in order to keep prices high, but these negotiations broke down in March 2020, prompting a sharp increase in oil production. Using a supply and demand graph, show how each of these factors affected the market price and quantity of crude oil. (LO3)

ANSWERS TO SELF-TESTS

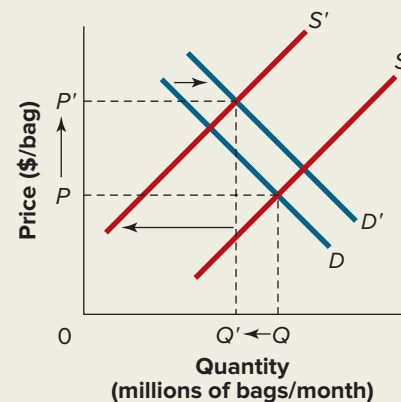
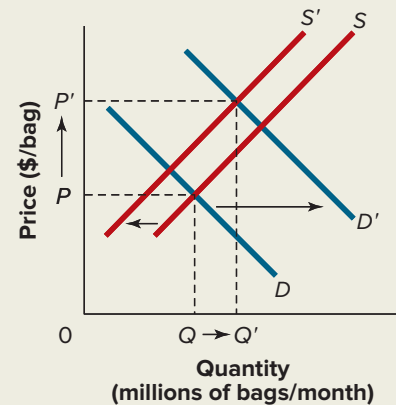
- 2.1 At a quantity of 10,000 slices per day, the marginal buyer's reservation price is \$3.50 per slice. At a price of \$2.50 per slice, the quantity demanded will be 14,000 slices per day. (LO1)



- 2.2 At a quantity of 10,000 slices per day, the marginal cost of pizza is \$2.50 per slice. At a price of \$3.50 per slice, the quantity supplied will be 14,000 slices per day. (LO1)



- 2.3 Because landlords are permitted to charge less than the maximum rent established by rent-control laws, a law that sets the maximum rent at \$1,200 will have no effect on the rents actually charged in this market, which will settle at the equilibrium value of \$800 per month. (LO2)
- 2.4 Travel by air and travel by intercity bus are substitutes, so a decline in airfares will shift the demand for bus travel to the left, resulting in lower bus fares and fewer bus trips taken. Travel by air and the use of resort hotels are complements, so a decline in airfares will shift the demand for resort hotel rooms to the right, resulting in higher hotel rates and an increase in the number of rooms rented. (LO3)
- 2.5 Apartments located far from Washington Metro stations are an inferior good. A pay increase for federal workers will thus shift the demand curve for such apartments downward, which will lead to a reduction in their equilibrium rent. (LO3)
- 2.6 The vitamin discovery shifts the demand for tortilla chips to the right and the crop losses shift the supply of tortilla chips to the left. Both shifts result in an increase in the equilibrium price of tortilla chips. But depending on the relative magnitude of the shifts, the equilibrium quantity of tortilla chips may either rise (top figure) or fall (bottom figure). (LO3)



The Algebra of Supply and Demand

In the text of this chapter, we developed supply and demand analysis in a geometric framework. The advantage of this framework is that many find it an easier one within which to visualize how shifts in either curve affect equilibrium price and quantity.

It is a straightforward extension to translate supply and demand analysis into algebraic terms. In this brief appendix, we show how this is done. The advantage of the algebraic framework is that it greatly simplifies computing the numerical values of equilibrium prices and quantities.

Consider, for example, the supply and demand curves in Figure 2A.1, where P denotes the price of the good and Q denotes its quantity. What are the equations of these curves?

The equation of a straight-line demand curve must take the general form $P = a + bQ^d$, where P is the price of the product (as measured on the vertical axis), Q^d is the quantity demanded at that price (as measured on the horizontal axis), a is the vertical intercept of the demand curve, and b is its slope. For the demand curve shown in Figure 2A.1, the vertical intercept is 16 and the slope is -2 . So the equation for this demand curve is

$$P = 16 - 2Q^d. \quad (2A.1)$$

Similarly, the equation of a straight-line supply curve must take the general form $P = c + dQ^s$, where P is again the price of the product, Q^s is the quantity supplied at that price, c is the vertical intercept of the supply curve, and d is its slope. For the supply curve shown in Figure 2A.1, the vertical intercept is 4 and the slope is also 4. So the equation for this supply curve is

$$P = 4 + 4Q^s. \quad (2A.2)$$

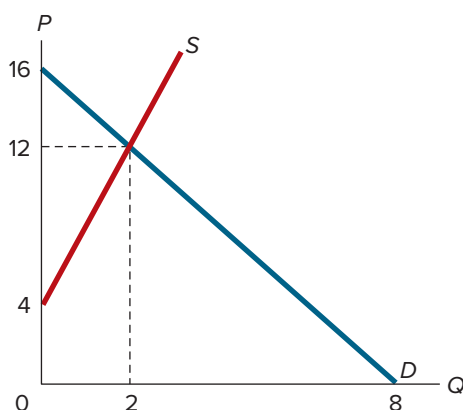


FIGURE 2A.1
Supply and Demand Curves.

If we know the equations for the supply and demand curves in any market, it is a simple matter to solve them for the equilibrium price and quantity using the method of simultaneous equations. The following example illustrates how to apply this method.

EXAMPLE 2A.1 Simultaneous Equations

If the supply and demand curves for a market are given by $P = 4 + 4Q^s$ and $P = 16 - 2Q^d$, respectively, find the equilibrium price and quantity for this market.

In equilibrium, we know that $Q^s = Q^d$. Denoting this common value as Q^* , we may then equate the right-hand sides of Equations 2A.1 and 2A.2 and solve

$$4 + 4Q^* = 16 - 2Q^*, \quad (2A.3)$$

which yields $Q^* = 2$. Substituting $Q^* = 2$ back into either the supply or demand equation gives the equilibrium price $P^* = 12$.

Of course, having already begun with the graphs of Equations 2A.1 and 2A.2 in hand, we could have identified the equilibrium price and quantity by a simple glance at Figure 2A.1. (That is why it seems natural to say that the graphical approach helps us visualize the equilibrium outcome.) As the following Self-Test 2A.1 illustrates, the advantage of the algebraic approach to finding the equilibrium price and quantity is that it is much less painstaking than having to produce accurate drawings of the supply and demand schedules.

SELF-TEST 2A.1

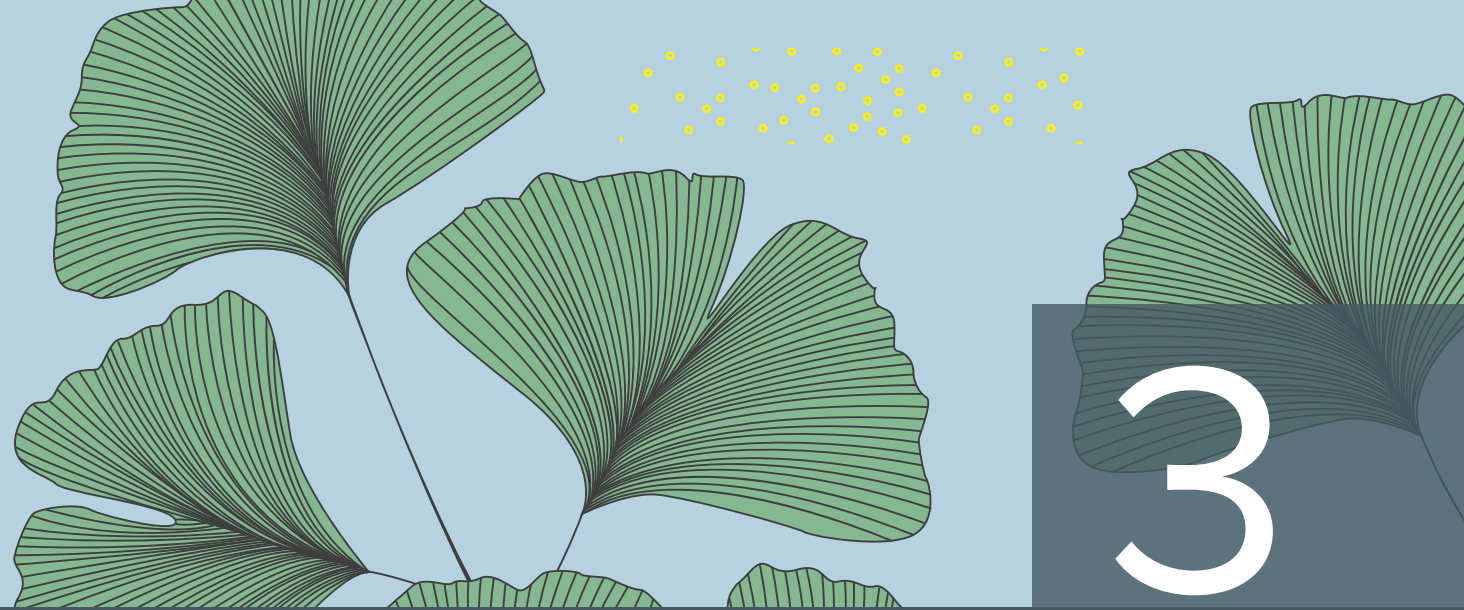
Find the equilibrium price and quantity in a market whose supply and demand curves are given by $P = 2Q^s$ and $P = 8 - 2Q^d$, respectively.

ANSWER TO APPENDIX SELF-TEST

2A.1 Let Q^* denote the equilibrium quantity. Since the equilibrium price and quantity lie on both the supply and demand curves, we equate the right-hand sides of the supply and demand equations to obtain

$$2Q^* = 8 - 2Q^*,$$

which solves for $Q^* = 2$. Substituting $Q^* = 2$ back into either the supply or demand equation gives the equilibrium price $P^* = 4$.



Demand and Elasticity

On the northern border of a large university in the East, a creek widens to form a picturesque lake, fondly remembered by generations of alumni as a popular recreation spot. Over the years, the lake had gradually silted in, and by the late 1980s, even paddling a canoe across it had become impossible. A generous alumnus then sponsored an effort to restore the lake. Heavy dredging equipment hauled out load after load of mud, and months later the lake was silt-free.

To mark the occasion, the university held a ceremony. Bands played, the president spoke, a chorus sang, and distinguished visitors applauded the donor's generosity. Hundreds of faculty and students turned out for the festivities. Spotting a good opportunity to promote their product, the proprietors of a local ice cream store set up a temporary stand at the water's edge, with a large sign: "Free Ice Cream."

Word spread. Soon scores of people were lined up waiting to try Vanilla Almond Delight, Hazelnut Cream, and Fudge Faire. The ice cream was plentiful, and because it was free, everyone could obviously afford it—or so it seemed. In fact, many people who wanted ice cream that day never got any. The reason, of course, was that they found waiting in a long line too steep a price to pay.

When a good or service is scarce, it must somehow be rationed among competing users. In most markets, monetary prices perform that task. But in the case of a stand offering free ice cream, waiting time becomes the effective rationing device. Having to stand in line is a cost, no less so than having to part with some money.

This example drives home the point that although the demand curve is usually described as a relationship between the quantity demanded of a good and its monetary price, the relationship is really a much more general one. At bottom, the demand curve is a relationship between the quantity demanded and *all* costs—monetary and nonmonetary—associated with acquiring a good.

Our task in this chapter will be to explore the demand side of the market in greater depth than was possible in Chapter 2, *Supply and Demand*. There we merely asked you to accept as an intuitively plausible claim that the quantity demanded of a

LEARNING OBJECTIVES

After reading this chapter, you should be able to:

- LO1** Relate the law of demand to the Cost-Benefit Principle.
- LO2** Apply the law of demand to consumer decision making to show how it is related to substitution and income effects.
- LO3** Discuss the relationship between the individual demand curve and the market demand curve.
- LO4** Define the price elasticity of demand and explain what determines whether demand is elastic or inelastic.
- LO5** Calculate the price elasticity of demand using information from a demand curve.
- LO6** Describe how changes in the price of a good affect total revenue and total expenditure depending on the price elasticity of demand for the good.
- LO7** Explain the cross-price elasticity of demand and the income elasticity of demand.

law of demand people do less of what they want to do as the cost of doing it rises

good or service declines as its price rises. This relationship is known as the **law of demand**, and we'll see how it emerges as a simple consequence of the assumption that people spend their limited incomes in rational ways. In the process, we'll see more clearly the dual roles of income and substitution as factors that account for the law of demand. Next, we will see how to generate market demand curves by adding the demand curves for individual buyers horizontally.

Following that, we will introduce the concept of elasticity, a measure of the extent to which quantity demanded and quantity supplied respond to variations in price, income, and other factors. In Chapter 2, *Supply and Demand*, we saw how shifts in supply and demand curves enabled us to predict the direction of change in the equilibrium values of price and quantity. An understanding of price elasticity will enable us to make even more precise statements about the effects of such changes. Finally, we will explore why some goods have higher price elasticity of demand than others and the implications of that fact for how total spending responds to changes in prices.

THE LAW OF DEMAND

With our discussion of the free ice cream offer in mind, let us restate the law of demand as follows:

Law of Demand: People do less of what they want to do as the cost of doing it rises.

By stating the law of demand this way, we can see it as a direct consequence of the Cost-Benefit Principle, which says that an activity should be pursued if (and only if) its benefits are at least as great as its costs. Recall that we measure the benefit of an activity by the highest price we'd be willing to pay to pursue it—namely, our reservation price for the activity. When the cost of an activity rises, it's more likely to exceed our reservation price, and we're therefore less likely to pursue that activity.

The law of demand applies to BMWs, cheap key rings, and “free” ice cream, not to mention manicures, medical care, and acid-free rain. It stresses that a “cost” is the sum of *all* the sacrifices—monetary and nonmonetary, implicit and explicit—we must make to engage in an activity.

The Origins of Demand

How much are you willing to pay for the latest Beyoncé album? The answer will clearly depend on how you feel about her music. To her diehard fans, buying the new release might seem absolutely essential; they'd pay a steep price indeed. But those who don't like her music may be unwilling to buy it at any price.

Wants (also called “preferences” or “tastes”) are clearly an important determinant of a consumer's reservation price for a good. But that raises the question of where wants come from. Many tastes—such as the taste for water on a hot day or for a comfortable place to sleep at night—are largely biological in origin. But many others are heavily shaped by culture, and even basic cravings may be socially molded. For example, people raised in southern India develop a taste for hot curry dishes, while those raised in France generally prefer milder foods.

Tastes for some items may remain stable for many years, but tastes for others may be highly volatile. Although books about the *Titanic* disaster have been continuously available since the vessel sank in spring 1912, not until the appearance of James Cameron's blockbuster film did these books begin to sell in large quantities. In spring 1998, 5 of the 15 books on *The New York Times* paperback bestseller list were about the *Titanic* itself or one of the actors in the film. Yet none of these books, or any other book about the *Titanic*, has made the bestseller list in the years since then. Still, echoes of the film continue to reverberate in the

marketplace. In the years since its release, for example, demand for ocean cruises has grown sharply, and several television networks have introduced shows set on cruise ships.

Peer influence provides another example of how social forces often influence demand. Indeed, it is often the most important single determinant of demand. For instance, if our goal is to predict whether a young man will purchase an illegal recreational drug, knowing how much income he has is not very helpful. Knowing the prices of whiskey and other legal substitutes for illicit drugs also tells us little. Although these factors do influence purchase decisions, by themselves they are weak predictors. But if we know that most of the young man's best friends are heavy drug users, there's a reasonably good chance that he'll use drugs as well.

Another important way in which social forces shape demand is in the relatively common desire to consume goods and services that are recognized as the best of their kind. For instance, many people want to see LeBron James play basketball not just because of his skills on the court, but because he is widely regarded as one of the best basketball players of all time.

Consider, too, the decision of how much to spend on an interview suit. Employment counselors never tire of reminding us that making a good first impression is extremely important when you go for a job interview. At the very least, that means showing up in a suit that looks good. But looking good is a relative concept. If everyone else shows up in a \$200 suit, you'll look good if you show up in a \$300 suit. But you won't look as good in that same \$300 suit if everyone else shows up in suits costing \$1,000. The amount you'll choose to spend on an interview suit, then, clearly depends on how much others in your circle are spending.

Needs versus Wants

In everyday language, we distinguish between goods and services people need and those they merely want. For example, we might say that someone wants a ski vacation in Utah, but what he really needs is a few days off from his daily routine; or that someone wants a house with a view, but what she really needs is shelter from the elements. Likewise, because people need protein to survive, we might say that a severely malnourished person needs more protein. But it would strike us as odd to say that anyone—even a malnourished person—needs more prime filet of beef because health can be restored by consuming far less expensive sources of protein.

Economists like to emphasize that once we have achieved bare subsistence levels of consumption—the amount of food, shelter, and clothing required to maintain our health—we can abandon all reference to needs and speak only in terms of wants. This linguistic distinction helps us think more clearly about the true nature of our choices.

For instance, someone who says “Californians don't have nearly as much water as they need” will tend to think differently about water shortages than someone who says “Californians don't have nearly as much water as they want when the price of water is low.” The first person is likely to focus on regulations to prevent people from watering their lawns, or on projects to capture additional runoff from the Sierra Nevada. The second person is more likely to focus on the artificially low price of water in California. Whereas remedies of the first sort are often costly and extremely difficult to implement, raising the price of water is both simple and effective.

The Economic Naturalist 3.1



Why does California experience chronic water shortages?

Some might respond that the state must serve the needs of a large population with a relatively low average annual rainfall. Yet other states, like New Mexico, have even less rainfall per person and do not experience water shortages nearly



Why do farmers grow water-intensive crops like rice in an arid state like California?

as often as California. California's problem exists because local governments sell water at extremely low prices, which encourages Californians to use water in ways that make no sense for a state with low rainfall. For instance, rice, which is well suited for conditions in high-rainfall states like South Carolina, requires extensive irrigation in California. But because California farmers can obtain water so cheaply, they plant and flood hundreds of thousands of acres of rice paddies each spring in the Central Valley. Two thousand tons of water are needed to produce one ton of rice, but many other grains can be produced with only half that amount. If the price of California water were higher, farmers would simply switch to other grains.

Likewise, cheap water encourages homeowners in Los Angeles and San Diego to plant water-intensive lawns and shrubs, like the ones common in the East and Midwest. By contrast, residents of cities like Santa Fe, New Mexico, where water prices are high, choose native plantings that require little or no watering.

APPLYING THE LAW OF DEMAND

The real payoff from learning the law of demand lies in using it to make sense of the world around you. To encourage you in your efforts to become an economic naturalist, we turn now to a sequence of Economic Naturalist examples in this vein.

Substitution at Work

In the first of these examples, we focus on the role of substitution. When the price of a good or service goes up, rational consumers generally turn to less expensive substitutes. Can't meet the payments on a new car? Then buy a used one, or rent an apartment on a bus or subway line. French restaurants too pricey? Then go out for Chinese, or eat at home more often. National Football League tickets too high? Watch the game on television, or read a book. Can't afford a book? Check one out of the library, or download some reading matter from the Internet. Once you begin to see substitution at work, you'll be amazed by the number and richness of the examples that confront you every day.



The Economic Naturalist 3.2

Why would Jeff Bezos live in a smaller house in Manhattan than in Medina, Washington?

Among his many residences, Amazon CEO Jeff Bezos owns a 29,000-square-foot home in Medina, Washington. Bezos also owns a 17,000-square-foot apartment in Manhattan. Although both residences are enormous, Bezos's apartment in Manhattan is much smaller than his home in Medina. Bezos is one of the richest people in the world, so why would he choose to purchase a much smaller home in Manhattan than Medina?

For people trying to decide how large a home to buy, the most obvious difference between Manhattan and Medina is the huge difference in housing prices. The cost of land alone is several times higher in Manhattan than in Medina, and

construction costs are also much higher. So even though Bezos could *afford* to purchase a 29,000-square-foot home in Manhattan, housing prices are so high that he simply chooses to purchase a smaller home and spend his fortune in other ways.

Here we note in passing that an additional factor in Bezos's decision may have been the link between context and evaluation: A house seems small only if it is small relative to other houses in the same local environment. Because Manhattan prices are so high, others choose to build smaller houses there, too, so a 17,000-square-foot house in Manhattan is a larger dwelling, in relative terms, than a 29,000-square-foot house in Medina. We will discuss this point more thoroughly in Chapter 8, *An Introduction to Behavioral Economics*.

An especially vivid illustration of substitution occurred during the late 1970s, when fuel shortages brought on by interruptions in the supply of oil from the Middle East led to sharp increases in the price of gasoline and other fuels. In a variety of ways—some straightforward, others remarkably ingenious—consumers changed their behavior to economize on the use of energy. They formed carpools; switched to public transportation; bought four-cylinder cars; moved closer to work; took fewer trips; turned down their thermostats; installed insulation, storm windows, and solar heaters; and bought more efficient appliances. Many people even moved farther south to escape high winter heating bills.

As the next example points out, consumers not only abandon a good in favor of substitutes when it gets more expensive, but they also return to that good when prices return to their original levels.

The Economic Naturalist 3.3



Why did people turn to four-cylinder cars in the 1970s, only to shift back to six- and eight-cylinder cars in the 1990s?

In 1973, the price of gasoline was 38 cents per gallon. The following year the price shot up to 52 cents per gallon in the wake of a major disruption of oil supplies. A second disruption in 1979 drove the 1980 price to \$1.19 per gallon. These sharp increases in the price of gasoline led to big increases in the demand for cars with four-cylinder engines, which delivered much better fuel economy than the six- and eight-cylinder cars most people had owned. After 1980, however, fuel supplies stabilized, and prices rose only slowly, reaching \$1.40 per gallon by 1999. Yet despite the continued rise in the price of gasoline, the switch to smaller engines did not continue. By the late 1980s, the proportion of cars sold with six- and eight-cylinder engines began rising again. Why this reversal?

The key to explaining these patterns is to focus on changes in the **real price** of gasoline. When someone decides how big an automobile engine to choose, what matters is not the **nominal price** of gasoline, but the price of gasoline *relative* to all other goods. After all, for a consumer faced with a decision of whether to spend \$1.40 for a gallon of gasoline, the important question is how much satisfaction she would get out of the other things she could purchase with the same money. Even though the price of gasoline continued to rise slowly in nominal, or dollar, terms through the 1980s and 1990s, it declined sharply relative to the price of other goods. Indeed, in terms of real purchasing power, the 1999 price was actually slightly lower than the 1973 price. (That is, in 1999 \$1.40 bought

real price the dollar price of a good relative to the average dollar price of all other goods

nominal price the absolute price of a good in dollar terms

slightly fewer goods and services than 38 cents bought in 1973.) It is this decline in the real price of gasoline that accounts for the reversal of the trend toward smaller engines.

McGraw Hill connect

► Visit your instructor's Connect course and access your eBook to view this video.



Why are smaller automobile engines more common in Europe than in the United States?

A sharp decline in the real price of gasoline also helps account for the explosive growth in sport-utility vehicles in the 1990s. Almost 4 million SUVs were sold in the United States in 2001, up from only 750,000 in 1990. Some of them—like the Ford Excursion—weigh more than 7,500 pounds (three times as much as a Honda Civic) and get less than 10 miles per gallon on city streets. Vehicles like these would have been dismal failures during the 1970s, but they were by far the hottest sellers in the cheap-energy environment of 2001.

In 2004, gasoline prices yet again began to rise sharply in real terms, and by 2012 had reached almost \$5 per gallon in some parts of the country. Just as expected, the patterns of vehicle purchases began to shift almost immediately. Large SUVs, in high demand just months earlier, began selling at deep discounts. And with long waiting lists for fuel-efficient cars such as the Toyota Prius, buyers not only seldom received discounts, but they frequently paid even more than the sticker price. More recently, with the dramatic fall in gasoline prices since 2012, sales for large SUVs and other light trucks are again on the rise, while sales for more fuel-efficient passenger cars have reached their lowest point in decades.

Here's another closely related example of the influence of price on spending decisions.



The Economic Naturalist 3.4

Why are automobile engines smaller in England than in the United States?



Does the quantity of horsepower demanded depend on gasoline prices?

In England, the most popular model of BMW's 5-series car is the 516i, whereas in the United States it is the 530i. The engine in the 516i is almost 50 percent smaller than the engine in the 530i. Why this difference?

In both countries, BMWs appeal to professionals with roughly similar incomes, so the difference cannot be explained by differences in purchasing power. Rather, it is the direct result of the heavy tax the British levy on gasoline. With tax, a gallon of gasoline sells for more than two times the price in the United States. This difference encourages the British to choose smaller, more fuel-efficient engines.

The Importance of Income Differences

The most obvious difference between the rich and the poor is that the rich have higher incomes. To explain why the wealthy generally buy larger houses than the poor, we need not assume that the wealthy feel more strongly about housing than the poor. A much simpler explanation is that the total satisfaction a person derives from her housing, as with most other goods, increases with the amount that she consumes.

As the next example illustrates, income influences the demand not only for housing and other goods, but also for quality of service.

The Economic Naturalist 3.5

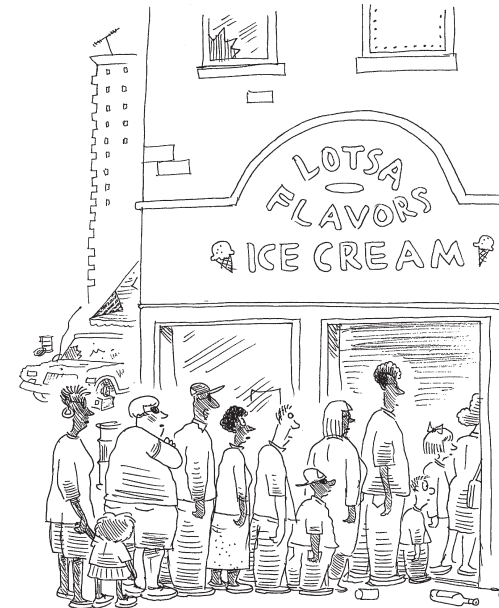


Why are waiting lines longer in poorer neighborhoods?

As part of a recent promotional campaign, a Baskin-Robbins retailer offered free ice cream at two of its franchise stores. The first was located in a high-income neighborhood, the second in a low-income neighborhood. Why was the queue for free ice cream longer in the low-income neighborhood?

Residents of both neighborhoods must decide whether to stand in line for free ice cream or go to some other store and avoid the line by paying the usual price. If we make the plausible assumption that people with higher incomes are more willing than others to pay to avoid standing in line, we should expect to see shorter lines in the high-income neighborhood.

Similar reasoning helps explain why lines are shorter in grocery stores that cater to high-income consumers. Keeping lines short at *any* grocery store means hiring more clerks, which means charging higher prices. High-income consumers are more likely than others to be willing to pay for shorter lines.



Why are lines longer in low-income neighborhoods?

RECAP ↑

APPLYING THE LAW OF DEMAND

Application of the law of demand highlights the important roles of income and substitution in explaining differences in consumption patterns—among individuals, among communities, and across time. The rule also highlights the fact that real, as opposed to nominal, prices and income are what matter.

INDIVIDUAL AND MARKET DEMAND CURVES

If we know what each individual's demand curve for a good looks like, how can we use that information to construct the market demand curve for the good? We must add the individual demand curves together, a process that is straightforward but requires care.

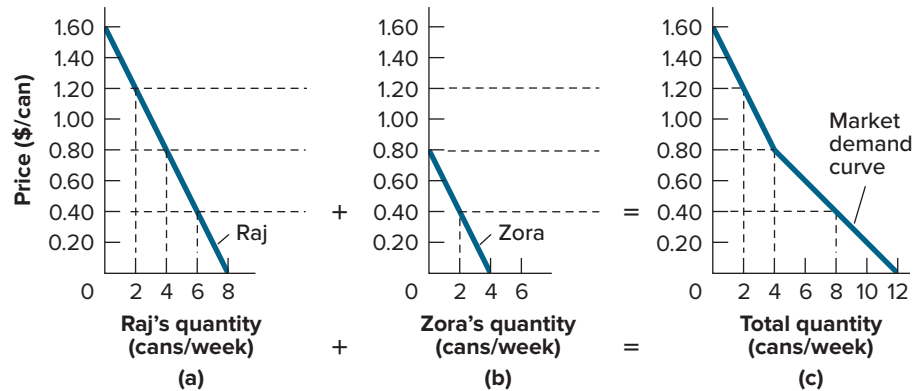
Horizontal Addition

Suppose that there are only two buyers—Raj and Zora—in the market for canned tuna and that their demand curves are as shown in Figure 3.1(a) and (b). To construct the market demand curve for canned tuna, we simply announce a sequence of prices and then add the quantity demanded by each buyer at each price. For example, at a price of 40 cents per can, Raj demands six cans per week (a) and Zora demands two cans per week (b), for a market demand of eight cans per week (c).

The process of adding individual demand curves to get the market demand curve is known as *horizontal addition*, a term used to emphasize that we are adding quantities, which are measured on the horizontal axes of individual demand curves.

FIGURE 3.1**Individual and Market Demand Curves for Canned Tuna.**

The quantity demanded at any price on the market demand curve (c) is the sum of the individual quantities demanded at that price, (a) and (b).

**SELF-TEST 3.1**

The buyers' side of the market for movie tickets consists of two consumers whose demands are as shown in the diagram below. Graph the market demand curve for this market.

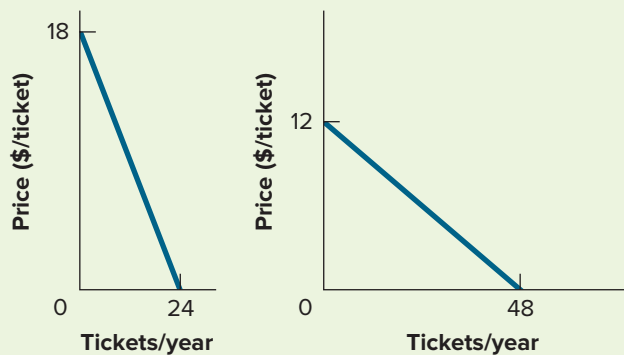
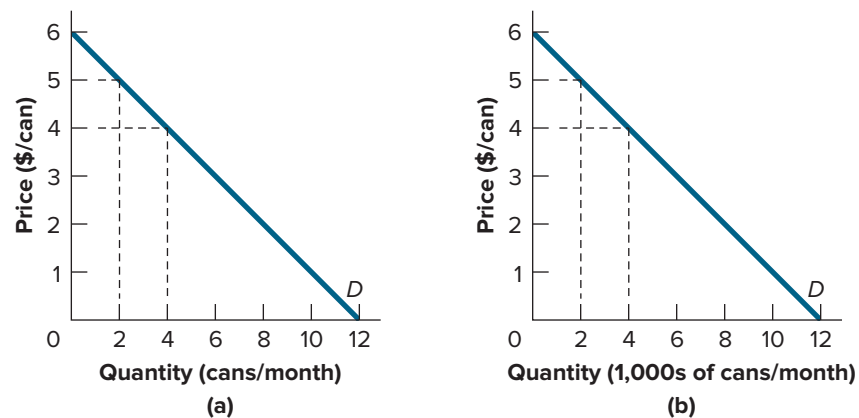


Figure 3.2 illustrates the special case in which each of 1,000 consumers in the market has the same demand curve (a). To get the market demand curve (b) in this case, we simply multiply each quantity on the representative individual demand curve by 1,000.

FIGURE 3.2**The Individual and Market Demand Curves When All Buyers Have Identical Demand Curves.**

When individual demand curves are identical, we get the market demand curve (b) by multiplying each quantity on the individual demand curve (a) by the number of consumers in the market.



ELASTICITY

Many illicit-drug users commit crimes to finance their addiction. The connection between drugs and crime has led to calls for more vigorous efforts to stop the smuggling of illicit drugs. But can such efforts reduce the likelihood that your iPhone or laptop computer will be stolen? If attempts to reduce the supply of illicit drugs are successful, our basic supply and demand analysis tells us that the supply curve for drugs will shift to the left and the market price of drugs will increase. Given that demand curves are downward-sloping, drug users will respond by consuming a smaller quantity of drugs. But the amount of crime drug users commit depends not on the *quantity* of drugs they consume, but rather on their *total expenditure* on drugs. Depending on the specific characteristics of the demand curve for illicit drugs, a price increase might reduce total expenditure on drugs, but it also could raise total expenditure.

Suppose, for example, that extra border patrols shift the supply curve in the market for illicit drugs to the left, as shown in Figure 3.3. As a result, the equilibrium quantity of drugs would fall from 50,000 to 40,000 ounces per day and the price of drugs would rise from \$50 to \$80 per ounce. The total amount spent on drugs, which was \$2,500,000 per day (50,000 ounces/day $\times$ \$50/ounce), would rise to \$3,200,000 per day (40,000 ounces/day $\times$ \$80/ounce). In this case, then, efforts to stem the supply of drugs would actually increase the likelihood of your laptop being stolen.

Other benefits from stemming the flow of illicit drugs might still outweigh the resulting increase in crime. But knowing that the policy might increase drug-related crime would clearly be useful to law-enforcement authorities.



Could reducing the supply of illegal drugs cause an increase in drug-related burglaries?

PRICE ELASTICITY OF DEMAND

When the price of a good or service rises, the quantity demanded falls. But to predict the effect of the price increase on total expenditure, we also must know by how much quantity falls. The quantity demanded of some goods such as salt is not very sensitive to changes in price. Indeed, even if the price of salt were to double, or to fall by half, most people would hardly alter their consumption of it. For other goods, however, the quantity demanded is extremely responsive to changes in price. For example, when a luxury tax was imposed on yachts in the early 1990s, purchases of yachts plummeted sharply. (Refer to The Economic Naturalist 3.7 example presented later in this chapter.)

Price Elasticity Defined

The **price elasticity of demand** for a good is a measure of the responsiveness of the quantity demanded of that good to changes in its price. Formally, the price elasticity of demand for a good is defined as the percentage change in the quantity demanded that results from a 1 percent change in its price. For example, if the price of beef falls

price elasticity of demand
the percentage change in the quantity demanded of a good or service that results from a 1 percent change in its price

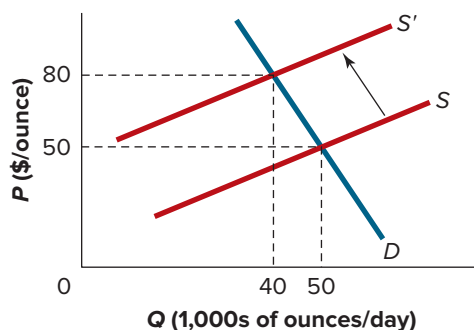


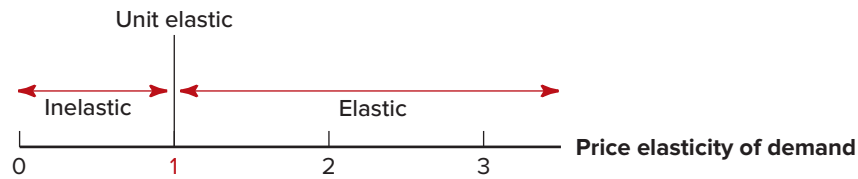
FIGURE 3.3

The Effect of Extra Border Patrols on the Market for Illicit Drugs.

Extra patrols shift supply leftward and reduce the quantity demanded, but they may actually increase the total amount spent on drugs.

FIGURE 3.4**Elastic and Inelastic Demand.**

Demand for a good is called elastic, unit elastic, or inelastic with respect to price if the price elasticity is greater than 1, equal to 1, or less than 1, respectively.



by 1 percent and the quantity demanded rises by 2 percent, then the price elasticity of demand for beef has a value of -2 .

Although the definition just given refers to the response of quantity demanded to a 1 percent change in price, it also can be adapted to other variations in price, provided they're relatively small. In such cases, we calculate the price elasticity of demand as the percentage change in quantity demanded divided by the corresponding percentage change in price. Thus, if a 2 percent reduction in the price of pork led to a 6 percent increase in the quantity of pork demanded, the price elasticity of demand for pork would be

$$\frac{\text{Percentage change in quantity demanded}}{\text{Percentage change in price}} = \frac{6 \text{ percent}}{-2 \text{ percent}} = -3. \quad (3.1)$$

Strictly speaking, the price elasticity of demand will always be negative (or zero) because price changes are always in the opposite direction from changes in quantity demanded. So for convenience, we drop the negative sign and speak of price elasticities in terms of absolute value. The demand for a good is said to be **elastic** with respect to price if the absolute value of its price elasticity is greater than 1. It is said to be **inelastic** if the absolute value of its price elasticity is less than 1. Finally, demand is said to be **unit elastic** if the absolute value of its price elasticity is equal to 1. (See Figure 3.4.)

elastic the demand for a good is elastic with respect to price if its price elasticity of demand is greater than 1

inelastic the demand for a good is inelastic with respect to price if its price elasticity of demand is less than 1

unit elastic the demand for a good is unit elastic with respect to price if its price elasticity of demand equals 1

EXAMPLE 3.1 Elasticity of Demand

What is the elasticity of demand for pizza?

When the price of pizza is \$1 per slice, buyers wish to purchase 400 slices per day, but when price falls to \$0.97 per slice, the quantity demanded rises to 404 slices per day. At the original price, what is the price elasticity of demand for pizza? Is the demand for pizza elastic with respect to price?

The fall in price from \$1 to \$0.97 is a decrease of 3 percent. The rise in quantity demanded from 400 slices to 404 slices is an increase of 1 percent. The price elasticity of demand for pizza is thus $(1 \text{ percent})/(3 \text{ percent}) = 1/3$. So when the initial price of pizza is \$1, the demand for pizza is not elastic with respect to price; it is inelastic.

SELF-TEST 3.2

What is the elasticity of demand for season ski passes?

When the price of a season ski pass is \$400, buyers, whose demand curve for passes is linear, wish to purchase 10,000 passes per year, but when price falls to \$380, the quantity demanded rises to 12,000 passes per year. At the original price, what is the price elasticity of demand for ski passes? Is the demand for ski passes elastic with respect to price?

Determinants of Price Elasticity of Demand

What factors determine the price elasticity of demand for a good or service? To answer this question, recall that before a rational consumer buys any product, the purchase decision must first satisfy the Cost-Benefit Principle. For instance, consider a good (such as a dorm refrigerator) that you buy only one unit of (if you buy it at all). Suppose that, at the current price, you have decided to buy it. Now imagine that the price goes up by 10 percent. Will a price increase of this magnitude be likely to make you change your mind? The answer will depend on factors like the following.

Substitution Possibilities

When the price of a product you want to buy goes up significantly, you're likely to ask yourself, "Is there some other good that can do roughly the same job, but for less money?" If the answer is yes, then you can escape the effect of the price increase by simply switching to the substitute product. But if the answer is no, you are more likely to stick with your current purchase.

These observations suggest that demand will tend to be more elastic with respect to price for products for which close substitutes are readily available. Salt, for example, has no close substitutes, which is one reason that the demand for it is highly inelastic. Note, however, that while the quantity of salt people demand is highly insensitive to price, the same cannot be said of the demand for any *specific brand* of salt. After all, despite what salt manufacturers say about the special advantages of their own labels, consumers tend to regard one brand of salt as a virtually perfect substitute for another. Thus, if Morton were to raise the price of its salt significantly, many people would simply switch to some other brand.

The vaccine against rabies is another product for which there are essentially no attractive substitutes. A person who is bitten by a rabid animal and does not take the vaccine faces a certain and painful death. Most people in that position would pay any price they could afford rather than do without the vaccine.

Budget Share

Suppose the price of key rings suddenly were to double. How would that affect the number of key rings you buy? If you're like most people, it would have no effect at all. Think about it—a doubling of the price of a \$1 item that you buy only every few years is simply nothing to worry about. By contrast, if the price of the new car you were about to buy suddenly doubled, you would definitely want to check out possible substitutes such as a used car or a smaller new model. You also might consider holding on to your current car a little longer. The larger the share of your budget an item accounts for, the greater is your incentive to look for substitutes when the price of the item rises. Big-ticket items, therefore, tend to have higher price elasticities of demand.

Time

Home appliances come in a variety of models, some more energy-efficient than others. As a general rule, the more efficient an appliance is, the higher its price. Suppose that you were about to buy a new air conditioner and electric rates suddenly rose sharply. It would probably be in your interest to buy a more efficient machine than you'd originally planned. However, what if you'd already bought a new air conditioner before you learned of the rate increase? You wouldn't think it worthwhile to discard the machine right away and replace it with a more efficient model. Rather, you'd wait until the machine wore out, or until you moved, before making the switch.

As this example illustrates, substitution of one product or service for another takes time. Some substitutions occur in the immediate aftermath of a price increase, but many others take place years or even decades later. For this reason, the price elasticity of demand for any good or service will be higher in the long run than in the short run.



► Visit your instructor's Connect course and access your eBook to view this video.



Why do people buy the same amount of (table) salt—even when the price rises substantially?

RECAP ↑

FACTORS THAT INFLUENCE PRICE ELASTICITY

The price elasticity of demand for a good or service tends to be larger when substitutes for the good are more readily available, when the good’s share in the consumer’s budget is larger, and when consumers have more time to adjust to a change in price.

Some Representative Elasticity Estimates

The entries in Table 3.1 show that the price elasticities of demand for different products often differ substantially—in this sample, ranging from a high of 3.5 for public transportation to a low of 0.1 for food. This variability is explained in part by the determinants of elasticity just discussed. Note, for example, that the price elasticity of demand for green peas is more than nine times that for coffee, reflecting the fact that there are many more close substitutes for green peas than for coffee.

Note also the contrast between the low price elasticity of demand for food and the high price elasticity of demand for green peas. Unlike green peas, food occupies a substantial share of most family budgets and there are few substitutes for broad spending categories like food.

TABLE 3.1
Historical Price Elasticity of Demand Estimates for Selected Products

Product or service	Estimated price elasticity
Food	0.1
Coffee	0.3
Magazines and newspapers	0.3
Housing	0.6
Tobacco	0.6
Clothing	0.6
Medical care	0.8
Oil	0.9
Motor vehicles	1.1
Beer	1.2
Furniture	1.3
Restaurant meals	1.6
Household electricity	1.9
Boats, pleasure aircraft	2.4
Green peas	2.8
Public transportation	3.5

Sources: K. Elzinga, “The Beer Industry,” *The Structure of American Industry*, ed. Walter Adams (New York: Macmillan, 1977); R. Fisher, *State and Local Public Finance* (Chicago, IL: Irwin, 1996); H. S. Houthakker and L. Taylor, *Consumer Demand in the United States: Analyses and Projections*, 2nd ed. (Cambridge, MA: Harvard University Press, 1970); A. Mansur and J. Whalley, “Numerical Specification of Applied General Equilibrium Models: Estimation, Calibration, and Data,” *Applied General Equilibrium Analysis*, ed. Herbert Scarf and John Shoven (New York: Cambridge University Press, 1984); J. Möller, “Income and Price Elasticities in Different Sectors of the Economy—An Analysis of Structural Change for Germany, the U.K., and the U.S.A.” December 1998; L. Taylor, “The Demand for Electricity: A Survey,” *Bell Journal of Economics*, Spring 1975; H. Theil, C. Chung, and J. Seale, “Advances in Econometrics,” Supplement I, *International Evidence on Consumption Patterns* (Greenwich, CT: JAI Press, 1989).

Using Price Elasticity of Demand

An understanding of the factors that govern price elasticity of demand is necessary not only to make sense of consumer behavior, but also to design effective public policy. Consider, for example, the debate about how taxes affect smoking among teenagers.

The Economic Naturalist 3.6



Will a higher tax on cigarettes curb teenage smoking?

Consultants hired by the tobacco industry have testified in Congress against higher cigarette taxes aimed at curbing teenage smoking. The main reason teenagers smoke is that their friends smoke, these consultants testified, and they concluded that higher taxes would have little effect. Does the consultants' testimony make economic sense?

The consultants are almost certainly right that peer influence is the most important determinant of teen smoking. But that does not imply that a higher tax on cigarettes would have little impact on adolescent smoking rates. Because most teenagers have little money to spend at their own discretion, cigarettes constitute a significant share of a typical teenage smoker's budget. The price elasticity of demand is thus likely to be far from negligible. For at least some teenage smokers, a higher tax would make smoking unaffordable. And even among those who could afford the higher prices, at least some others would choose to spend their money on other things rather than pay the higher prices.

Given that the tax would affect at least *some* teenage smokers, the consultants' argument begins to unravel. If the tax deters even a small number of smokers directly through its effect on the price of cigarettes, it will also deter others indirectly, by reducing the number of peer role models who smoke. And those who refrain because of these indirect effects will in turn no longer influence others to smoke, and so on. So even if the direct effect of higher cigarette taxes on teen smoking is small, the cumulative effects may be extremely large. The mere fact that peer pressure may be the primary determinant of teen smoking therefore does not imply that higher cigarette taxes will have no significant impact on the number of teens who smoke.



Do high cigarette prices discourage teen smoking?

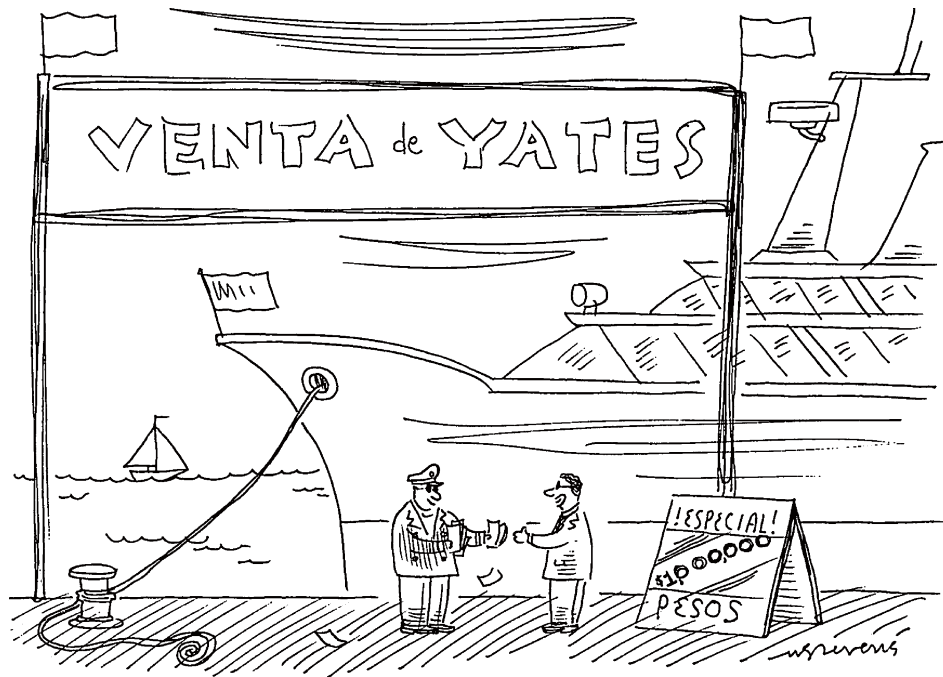
The Economic Naturalist 3.7



Why was the luxury tax on yachts such a disaster?

In 1990, Congress imposed a luxury tax on yachts costing more than \$100,000, along with similar taxes on a handful of other luxury goods. Before these taxes were imposed, the Joint Committee on Taxation estimated that they would yield more than \$31 million in revenue in 1991. However, the tax actually generated only a bit more than half that amount, \$16.6 million.¹ Several years later, the Joint

¹For an alternative view, see Dennis Zimmerman, "The Effect of the Luxury Excise Tax on the Sale of Luxury Boats," Congressional Research Service, February 10, 1992.



Why did the luxury tax on yachts backfire?

Economic Committee estimated that the tax on yachts had led to a loss of 7,600 jobs in the U.S. boating industry. Taking account of lost income taxes and increased unemployment benefits, the U.S. government actually came out \$7.6 million behind in fiscal 1991 as a result of its luxury taxes—almost \$39 million worse than the initial projection. What went wrong?

The 1990 law imposed no luxury taxes on yachts built and purchased outside the United States. What Congress failed to consider was that foreign-built yachts are almost perfect substitutes for yachts built and purchased in the United States. And, no surprise,

when prices on domestic yachts went up because of the tax, yacht buyers switched in droves to foreign models. A tax imposed on a good with a high price elasticity of demand stimulates large rearrangements of consumption but yields little revenue. Had Congress done the economic analysis properly, it would have predicted that this particular tax would be a big loser. Facing angry protests from unemployed New England shipbuilders, Congress repealed the luxury tax on yachts in 1993.

A GRAPHICAL INTERPRETATION OF PRICE ELASTICITY

For small changes in price, price elasticity of demand is the proportion by which quantity demanded changes divided by the corresponding proportion by which price changes. This formulation enables us to construct a simple expression for the price elasticity of demand for a good using only minimal information about its demand curve.

Look at Figure 3.5. P represents the current price of a good and Q the quantity demanded at that price. ΔP represents a small change in the current price, and the resulting change in quantity demanded is given by ΔQ . The expression $\Delta P/P$ will then stand for the proportion by which price changes and $\Delta Q/Q$ will stand for the corresponding proportion by which quantity changes. These two expressions, along with our definition of the price elasticity of demand (Equation 3.1), give us the formula for price elasticity

$$\text{Price elasticity} = \epsilon = \frac{\Delta Q/Q}{\Delta P/P}. \quad (3.2)$$

Suppose, for example, that 20 units were sold at the original price of 100 and that when price rose to 105, quantity demanded fell to 15 units. Neglecting the

**FIGURE 3.5****A Graphical Interpretation of Price Elasticity of Demand.**

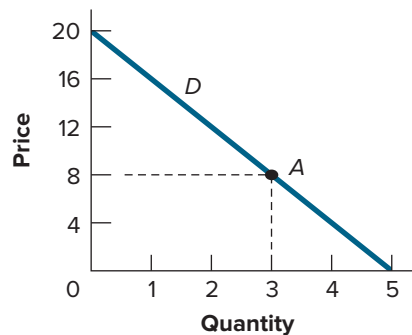
Price elasticity of demand at any point along a straight-line demand curve is the ratio of price to quantity at that point times the reciprocal of the slope of the demand curve.

negative sign of the quantity change, we would then have $\Delta Q/Q = 5/20$ and $\Delta P/P = 5/100$ which yields $\epsilon = (5/20)/(5/100) = 5$.

One attractive feature of this formula is that it has a straightforward graphical interpretation. Thus, if we want to calculate the price elasticity of demand at point A on the demand curve shown in Figure 3.5, we can begin by rewriting the right-hand side of Equation 3.2 as $(P/Q) \times (\Delta Q/\Delta P)$. And since the slope of the demand curve is equal to $\Delta P/\Delta Q$, $\Delta Q/\Delta P$ is the reciprocal of that slope: $\Delta Q/\Delta P = 1/\text{slope}$. The price elasticity of demand at point A, denoted ϵ_A , therefore has the following simple formula

$$\epsilon_A = \frac{P}{Q} \times \frac{1}{\text{slope}}. \quad (3.3)$$

To demonstrate how convenient this graphical interpretation of elasticity can be, suppose we want to find the price elasticity of demand at point A on the demand curve in Figure 3.6. The slope of this demand curve is the ratio of its vertical intercept to its horizontal intercept: $20/5 = 4$. So $1/\text{slope} = 1/4$. (Actually, the slope is -4 , but we again ignore the minus sign for convenience since price elasticity of demand always has the same sign.) The ratio P/Q at point A is $8/3$, so the price elasticity at point A is equal to $(P/Q) \times (1/\text{slope}) = (8/3) \times (1/4) = 2/3$. This means that when the price of the good is 8, a 3 percent reduction in price will lead to a 2 percent increase in quantity demanded.

**FIGURE 3.6****Calculating Price Elasticity of Demand.**

The price elasticity of demand at A is given by $(P/Q) \times (1/\text{slope}) = (8/3) \times (1/4) = 2/3$.

SELF-TEST 3.3

What is the price elasticity of demand when $P = 4$ on the demand curve in Figure 3.6?

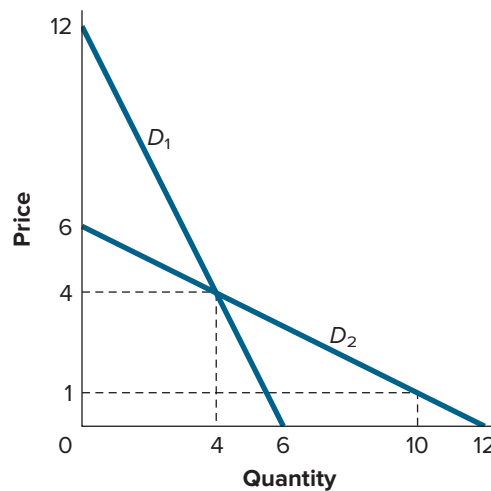
EXAMPLE 3.2 The Relationship between Elasticity and Slope

For the demand curves D_1 and D_2 shown in Figure 3.7, calculate the price elasticity of demand when $P = 4$. What is the price elasticity of demand on D_2 when $P = 1$?

These elasticities can be calculated easily using the formula $\epsilon = (P/Q) \times (1/\text{slope})$. The slope of D_1 is the ratio of its vertical intercept to its horizontal intercept: $12/6 = 2$. So $(1/\text{slope})$ is $1/2$ for D_1 . Similarly, the slope of D_2 is the ratio of its vertical intercept to its horizontal intercept: $6/12 = 1/2$. So the reciprocal of the slope of D_2 is 2. For both demand curves, $Q = 4$ when $P = 4$, so $(P/Q) = 4/4 = 1$ for each. Thus the price elasticity of demand when $P = 4$ is $(1) \times (1/2) = 1/2$ for D_1 and $(1) \times (2) = 2$ for D_2 . When $P = 1$, $Q = 10$ on D_2 , so $(P/Q) = 1/10$. Thus price elasticity of demand $= (1/10) \times (2) = 1/5$ when $P = 1$ on D_2 .

FIGURE 3.7**Price Elasticity and the Steepness of the Demand Curve.**

When price and quantity are the same, price elasticity of demand is always greater for the less steep of two demand curves.



This example illustrates a general rule: If two demand curves have a point in common, the steeper curve must be the less price-elastic of the two with respect to price at that point. However, this does not mean that the steeper curve is less elastic at *every* point. Thus, we saw that at $P = 1$, price elasticity of demand on D_2 was only $1/5$, or less than half the corresponding elasticity on the steeper D_1 at $P = 4$.

Price Elasticity Changes along a Straight-Line Demand Curve

As a glance at our elasticity formula makes clear, price elasticity has a different value at every point along a straight-line demand curve. The slope of a straight-line demand curve is constant, which means that $1/\text{slope}$ is also constant. But the price-quantity ratio P/Q declines as we move down the demand curve. The elasticity of demand thus declines steadily as we move downward along a straight-line demand curve.

Since price elasticity is the percentage change in quantity demanded divided by the corresponding percentage change in price, this pattern makes sense. After all, a price movement of a given absolute size is small in percentage terms when it occurs near the top of the demand curve, where price is high, but large in percentage terms when it occurs near the bottom of the demand curve, where price is low. Likewise, a

quantity movement of a given absolute value is large in percentage terms when it occurs near the top of the demand curve, where quantity is low, and small in percentage terms when it occurs near the bottom of the curve, where quantity is high.

The graphical interpretation of elasticity also makes it easy to see why the price elasticity of demand at the midpoint of any straight-line demand curve must always be 1. Consider, for example, the price elasticity of demand at point A on the demand curve D shown in Figure 3.8. At that point, the ratio P/Q is equal to $6/3 = 2$. The slope of this demand curve is the ratio of its vertical intercept to its horizontal intercept, $12/6 = 2$. So $(1/\text{slope}) = 1/2$ (again, we neglect the negative sign for simplicity). Inserting these values into the graphical elasticity formula yields $\epsilon_A = (P/Q) \times (1/\text{slope}) = (2) \times (1/2) = 1$.

This result holds not just for Figure 3.8, but also for any other straight-line demand curve.² A glance at the formula also tells us that since P/Q declines as we move downward along a straight-line demand curve, price elasticity of demand must be less than 1 at any point below the midpoint. By the same token, price elasticity must be greater than 1 for any point above the midpoint. Figure 3.9 summarizes these findings by denoting the elastic, inelastic, and unit elastic portions of any straight-line demand curve.

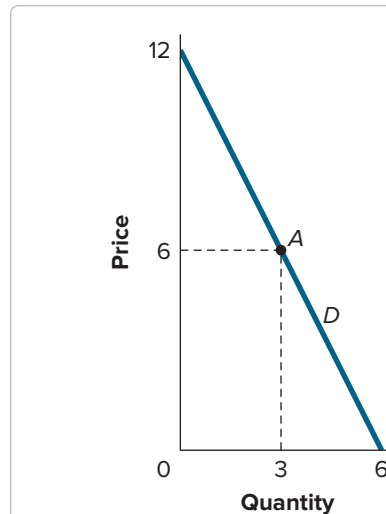


FIGURE 3.8

Elasticity at the Midpoint of a Straight-Line Demand Curve.

The price elasticity of demand at the midpoint of any straight-line demand curve always takes the value 1.

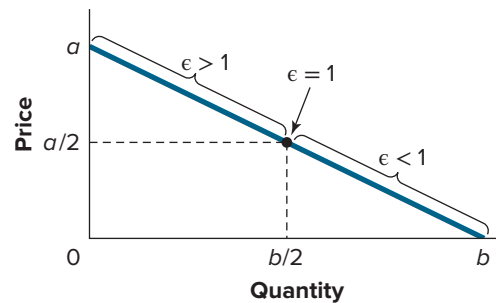


FIGURE 3.9

Price Elasticity Regions along a Straight-Line Demand Curve.

Demand is elastic on the top half, unit elastic at the midpoint, and inelastic on the bottom half of a straight-line demand curve.

Two Special Cases

There are two important exceptions to the general rule that elasticity declines along straight-line demand curves. First, the horizontal demand curve in Figure 3.10(a) has a slope of zero, which means that the reciprocal of its slope is infinite. Price elasticity of demand is thus infinite at every point along a horizontal demand curve. Such demand curves are said to be **perfectly elastic**.

Second, the demand curve in Figure 3.10(b) is vertical, which means that its slope is infinite. The reciprocal of its slope is thus equal to zero. Price elasticity of demand is thus exactly zero at every point along the curve. For this reason, vertical demand curves are said to be **perfectly inelastic**.

perfectly elastic demand

demand is perfectly elastic with respect to price if price elasticity of demand is infinite

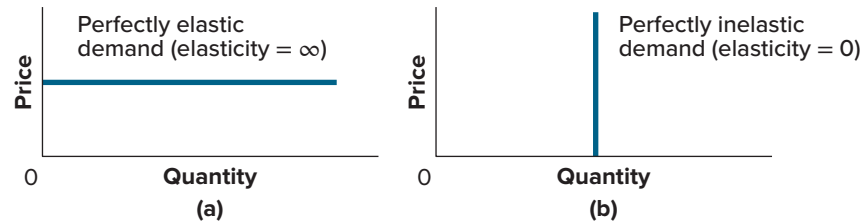
perfectly inelastic demand

demand is perfectly inelastic with respect to price if price elasticity of demand is zero

²To see why, note that at the midpoint of any such curve, P is exactly half the vertical intercept of the demand curve and Q is exactly half the horizontal intercept. Since the ratio of the vertical intercept to the horizontal intercept is the slope of the demand curve, the ratio (P/Q) must also be equal to the slope of the demand curve. And this means that $(1/\text{slope})$ will always be equal to (Q/P) . Thus, the product $(P/Q) \times (1/\text{slope}) = (P/Q) \times (Q/P)$ will always be exactly 1 at the midpoint of any straight-line demand curve.

FIGURE 3.10**Perfectly Elastic and Perfectly Inelastic Demand Curves.**

The horizontal demand curve (a) is perfectly elastic, or infinitely elastic, at every point. Even the slightest increase in price leads consumers to desert the product in favor of substitutes. The vertical demand curve (b) is perfectly inelastic at every point. Consumers do not, or cannot, switch to substitutes even in the face of large increases in price.

**RECAP****CALCULATING PRICE ELASTICITY OF DEMAND**

The price elasticity of demand for a good is the percentage change in the quantity demanded that results from a 1 percent change in its price. Mathematically, the elasticity of demand at a point along a demand curve is equal to $(P/Q) \times (1/\text{slope})$, where P and Q represent price and quantity and $(1/\text{slope})$ is the reciprocal of the slope of the demand curve at that point. Demand is elastic with respect to price if the absolute value of its price elasticity exceeds 1; inelastic if price elasticity is less than 1; and unit elastic if price elasticity is equal to 1.

ELASTICITY AND TOTAL EXPENDITURE

Sellers of goods and services have a strong interest in being able to answer questions like “Will consumers spend more on my product if I sell more units at a lower price or fewer units at a higher price?” As it turns out, the answer to this question depends critically on the price elasticity of demand. To see why, let’s first examine how the total amount spent on a good varies with the price of the good.

The total daily expenditure on a good is simply the daily number of units bought times the price for which it sells. The market demand curve for a good tells us the quantity that will be sold at each price. We can thus use the information on the demand curve to show how the total amount spent on a good will vary with its price.

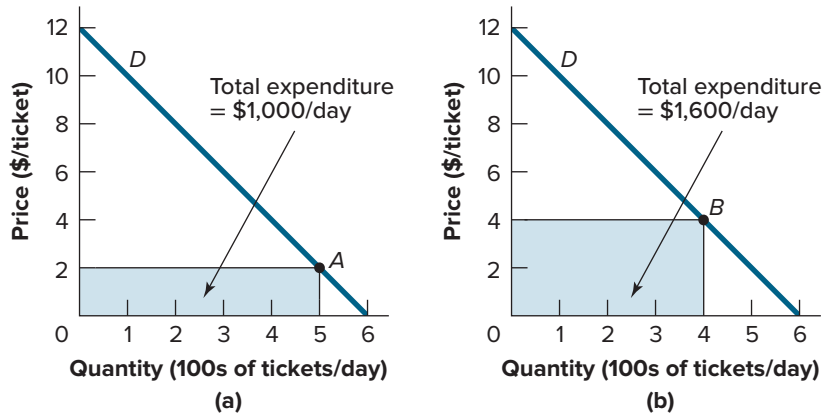
To illustrate, let’s calculate how much moviegoers will spend on tickets each day if the demand curve is as shown in Figure 3.11 and the price is \$2 per ticket (a). The demand curve tells us that at a price of \$2 per ticket, 500 tickets per day will be sold, so total expenditure at that price will be \$1,000 per day. If tickets sell not for \$2 but for \$4 apiece, 400 tickets will be sold each day (b), so total expenditure at the higher price will be \$1,600 per day.

Note that the total amount consumers spend on a product each day must equal the total amount sellers of the product receive. That is, the terms **total expenditure** and **total revenue** are simply two sides of the same coin:

Total expenditure = Total revenue: The dollar amount that consumers spend on a product ($P \times Q$) is equal to the dollar amount that sellers receive.

It might seem that an increase in the market price of a product should always result in an increase in the total revenue received by sellers. Although that happened in the case we just saw, it needn’t always be so. The law of demand tells us that when the price of a good rises, people will buy less of it. The two factors that

total expenditure (or total revenue) the dollar amount that consumers spend on a product ($P \times Q$) is equal to the dollar amount that sellers receive

**FIGURE 3.11****The Demand Curve for Movie Tickets.**

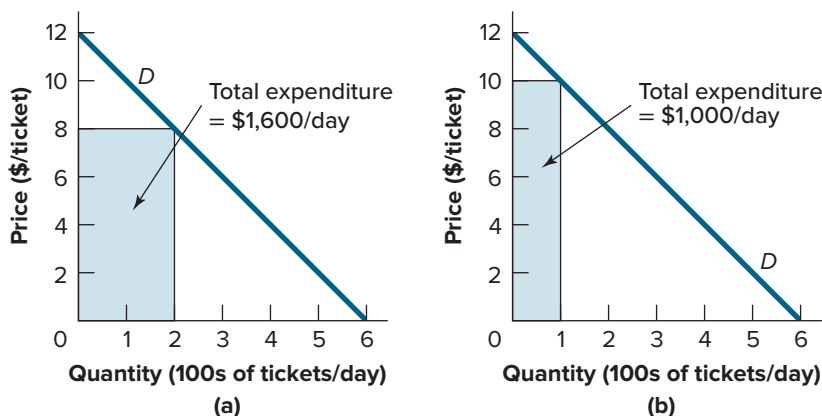
An increase in price from \$2 to \$4 per ticket increases total expenditure on tickets.

govern total revenue—price and quantity—will thus always move in opposite directions as we move along a demand curve. When price goes up and quantity goes down, the product of the two may go either up or down.

Note, for example, that for the demand curve shown in Figure 3.12 (which is the same as the one in Figure 3.11), a rise in price from \$8 per ticket (a) to \$10 per ticket (b) will cause total expenditure on tickets to go down. Thus people will spend \$1,600 per day on tickets at a price of \$8, but only \$1,000 per day at a price of \$10.

The general rule illustrated by Figures 3.11 and 3.12 is that a price increase will produce an increase in total revenue whenever it is greater, in percentage terms, than the corresponding percentage reduction in quantity demanded. Although the two price increases (from \$2 to \$4 and from \$8 to \$10) were of the same absolute value—\$2 in each case—they are much different when expressed as a percentage of the original price. An increase from \$2 to \$4 represents a 100 percent increase in price, whereas an increase from \$8 to \$10 represents only a 25 percent increase in price. And although the quantity reductions caused by the two price increases were also equal in absolute terms, they too are very different when expressed as percentages of the quantities originally sold. Thus, although the decline in quantity demanded was 100 tickets per day in each case, it was just a 20 percent reduction in the first case (from 500 units to 400 in Figure 3.11) but a 50 percent reduction in the second (from 200 units to 100 in Figure 3.12). In the second case, the negative effect on total expenditure of the 50 percent quantity reduction outweighed the positive effect of the 25 percent price increase. The reverse happened in the first case: The 100 percent increase in price (from \$2 to \$4) outweighed the 20 percent reduction in quantity (from 5 units to 4 units).

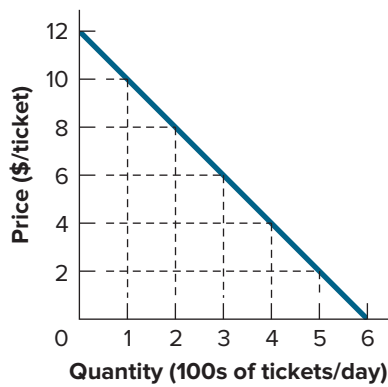
The following example provides further insight into the relationship between total revenue and price.

**FIGURE 3.12****The Demand Curve for Movie Tickets.**

An increase in price from \$8 to \$10 per ticket results in a fall in total expenditure on tickets.

EXAMPLE 3.3 Elasticity and Total Expenditure

FIGURE 3.13
The Demand Curve for
Movie Tickets.



For the demand curve shown in Figure 3.13, draw a separate graph showing how total expenditure varies with the price of movie tickets.

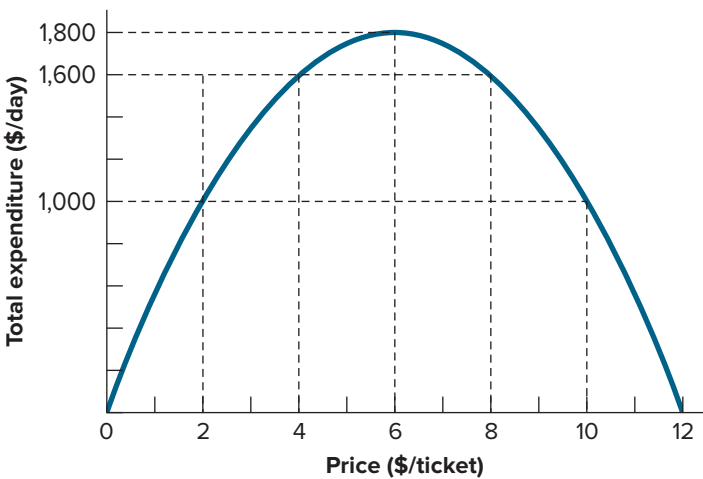
The first step in constructing this graph is to calculate total expenditure for each price shown in the graph and record the results, as in Table 3.2. The next step is to plot total expenditure at each of the price points on a graph, as in Figure 3.14. Finally, sketch the curve by joining these points. (If greater accuracy is required, you can use a larger sample of points than the one shown in Table 3.2.)

Note in Figure 3.14 that as the price per ticket increases from \$0 to \$6, total expenditure increases. But as the price rises from \$6 to \$12, total expenditure decreases. Total expenditure reaches a maximum of \$1,800 per day at a price of \$6.

TABLE 3.2
Total Expenditure as a Function of Price

Price (\$/ticket)	Total expenditure (\$/day)
12	0
10	1,000
8	1,600
6	1,800
4	1,600
2	1,000
0	0

FIGURE 3.14
Total Expenditure as a
Function of Price.
For a good whose demand
curve is a straight line, total
expenditure reaches a
maximum at the price
corresponding to the midpoint
of the demand curve.



The pattern observed in the preceding example holds true in general. For a straight-line demand curve, total expenditure is highest at the price that lies on the midpoint of the demand curve.

Bearing in mind these observations about how expenditure varies with price, let's return to the question of how the effect of a price change on total expenditure depends on the price elasticity of demand. Suppose, for example, that the business manager of a rock band knows she can sell 5,000 tickets to the band's weekly summer concerts if she sets the price at \$20 per ticket. If the elasticity of demand for tickets is equal to 3, will total ticket revenue go up or down in response to a 10 percent increase in the price of tickets?

Total revenue from tickets sold is currently $(\$20/\text{ticket}) \times (5,000 \text{ tickets/week}) = \$100,000$ per week. The fact that the price elasticity of demand for tickets is 3 implies that a 10 percent increase in price will produce a 30 percent reduction in the number of tickets sold, which means that quantity will fall to 3,500 tickets per week. Total expenditure on tickets will therefore fall to $(3,500 \text{ tickets/week}) \times (\$22/\text{ticket}) = \$77,000$ per week, which is significantly less than the current spending total.

What would have happened to total expenditure if the band manager had *reduced* ticket prices by 10 percent, from \$20 to \$18? Again assuming a price elasticity of 3, the result would have been a 30 percent increase in tickets sold—from 5,000 per week to 6,500 per week. The resulting total expenditure would have been $(\$18/\text{ticket}) \times (6,500 \text{ tickets/week}) = \$117,000$ per week, significantly more than the current total.

These examples illustrate the following important rule about how price changes affect total expenditure for an elastically demanded good:

Rule 1: When price elasticity of demand is greater than 1, changes in price and changes in total expenditure always move in opposite directions.

Let's look at the intuition behind this rule. Total expenditure is the product of price and quantity. For an elastically demanded product, the percentage change in quantity will be larger than the corresponding percentage change in price. Thus the change in quantity will more than offset the change in revenue per unit sold.

Now let's see how total spending responds to a price increase when demand is *inelastic* with respect to price. Consider a case like the one just considered except that the elasticity of demand for tickets is not 3 but 0.5. How will total expenditure respond to a 10 percent increase in ticket prices? This time the number of tickets sold will fall by only 5 percent to 4,750 tickets per week, which means that total expenditure on tickets will rise to $(4,750 \text{ tickets/week}) \times (\$22/\text{ticket}) = \$104,500$ per week, or \$4,500 per week more than the current expenditure level.

In contrast, a 10 percent price reduction (from \$20 to \$18 per ticket) when price elasticity is 0.5 would cause the number of tickets sold to grow by only 5 percent, from 5,000 per week to 5,250 per week, resulting in total expenditure of $(\$18/\text{ticket}) \times (5,250 \text{ tickets/week}) = \$94,500$ per week, significantly less than the current total.

As these examples illustrate, the effect of price changes on total expenditure when demand is inelastic is precisely the opposite of what it was when demand was elastic:

Rule 2: When price elasticity of demand is less than 1, changes in price and changes in total expenditure always move in the same direction.

Again, the intuition behind this rule is straightforward. For a product whose demand is inelastic with respect to price, the percentage change in quantity demanded will be smaller than the corresponding percentage change in price. The change in revenue per unit sold (price) will thus more than offset the change in the number of units sold.

The relationship between elasticity and the effect of a price change on total revenue is summarized in Figure 3.15, where the symbol ϵ is used to denote elasticity.

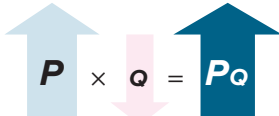
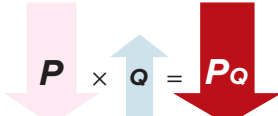
If demand is . . .	A price increase will . . .	A price reduction will . . .
elastic ($\epsilon > 1$)	reduce total expenditure 	increase total expenditure 
inelastic ($\epsilon < 1$)	increase total expenditure 	reduce total expenditure 

FIGURE 3.15**Elasticity and the Effect of a Price Change on Total Expenditure.**

Recall that in the example with which we began this chapter, an increase in the price of drugs led to an increase in the total amount spent on drugs. That will happen whenever the demand for drugs is inelastic with respect to price, as it was in that example. Had the demand for drugs instead been elastic with respect to price, the drug supply interruption would have led to a reduction in total expenditure on drugs.

INCOME ELASTICITY AND CROSS-PRICE ELASTICITY OF DEMAND

The elasticity of demand for a good can be defined not only with respect to its own price but also with respect to the prices of substitutes or complements, or even to income. For example, the elasticity of demand for peanuts with respect to the price of cashews—also known as the **cross-price elasticity of demand** for peanuts with respect to cashew prices—is the percentage by which the quantity of peanuts demanded changes in response to a 1 percent change in the price of cashews. The **income elasticity of demand** for peanuts is the percentage by which the quantity demanded of peanuts changes in response to a 1 percent change in income.

Unlike the elasticity of demand for a good with respect to its own price, these other elasticities may be either positive or negative, so it is important to note their algebraic signs carefully. The income elasticity of demand for inferior goods, for example, is negative, whereas the income elasticity of demand for normal goods is positive. When the cross-price elasticity of demand for two goods is positive—as in the peanuts–cashews example—the two goods are substitutes. When it is negative, the two goods are complements. The elasticity of demand for tennis racquets with respect to court rental fees, for example, is less than zero.

cross-price elasticity of demand the percentage by which the quantity demanded of the first good changes in response to a 1 percent change in the price of the second

income elasticity of demand the percentage by which a good's quantity demanded changes in response to a 1 percent change in income

SELF-TEST 3.4

If a 10 percent increase in income causes the number of students who choose to attend private universities to go up by 5 percent, what is the income elasticity of demand for private universities?

RECAP ↑**CROSS-PRICE AND INCOME ELASTICITIES**

When the cross-price elasticity of demand for one good with respect to the price of another good is positive, the two goods are substitutes; when the cross-price elasticity of demand is negative, the two goods are complements. A normal good has positive income elasticity of demand and an inferior good has negative income elasticity of demand.

SUMMARY

- The law of demand states that people do less of what they want to do as the cost of doing it rises. The law of demand is a direct consequence of the Cost-Benefit Principle, which says that an activity should be pursued if (and only if) its benefits are at least as great as its costs. (LO1)
- As the price of a good rises, people tend to choose less expensive alternatives. The ability to substitute one good for another is an important factor behind the law of demand. (LO2)
- For normal goods, the income effect is a second important reason that demand curves slope downward. When the price of such a good rises, not only does it become less attractive relative to its substitutes, but the consumer also has less real purchasing power, and this, too, reduces the quantity demanded. (LO2)
- The demand curve is a schedule that shows the amounts of a good people want to buy at various prices. Demand curves can be used to summarize the price–quantity relationship for a single individual, but more commonly we employ them to summarize that relationship for an entire market. At any quantity along a demand curve, the corresponding price represents the amount by which the consumer (or consumers) would benefit from having an additional unit of the product. For this reason, the demand curve is sometimes described as a summary of the benefit side of the market. (LO3)
- The price elasticity of demand is a measure of how strongly buyers respond to changes in price. It is the percentage change in quantity demanded that occurs in response to a 1 percent change in price. The demand for a good is called elastic with respect to price if the absolute value of its price elasticity is more than 1, inelastic if its price elasticity is less than 1, and unit elastic if its price elasticity is equal to 1. (LO4)
- Goods such as salt, which occupy only a small share of the typical consumer’s budget and have few or no good substitutes, tend to have low price elasticity of demand. Goods like new cars of a particular make and model, which occupy large budget shares and have many attractive substitutes, tend to have high price elasticity of demand. Price elasticity of demand is higher in the long run than in the short run because people often need time to adjust to price changes. (LO4)
- The price elasticity of demand at a point along a demand curve also can be expressed as the formula $\epsilon = (\Delta Q/Q)/(\Delta P/P)$. Here, P and Q represent price and quantity at that point and ΔP and ΔQ represent small changes in price and quantity. For straight-line demand curves, this formula can also be expressed as $\epsilon = (P/Q) \times (1/\text{slope})$. These formulations tell us that price elasticity declines in absolute terms as we move down a straight-line demand curve. (LO5)
- A cut in price will increase total spending on a good if demand is elastic but reduce it if demand is inelastic. An increase in price will increase total spending on a good if demand is inelastic but reduce it if demand is elastic. Total expenditure on a good reaches a maximum when price elasticity of demand is equal to 1. (LO6)
- Analogous formulas are used to define the elasticity of demand for a good with respect to income and the prices of other goods. In each case, elasticity is the percentage change in quantity demanded divided by the corresponding percentage change in income or price. (LO7)

KEY TERMS

cross-price elasticity of demand
elastic
income elasticity of demand
inelastic

law of demand
nominal price
perfectly elastic demand
perfectly inelastic demand

price elasticity of demand
real price
total expenditure (or total revenue)
unit elastic

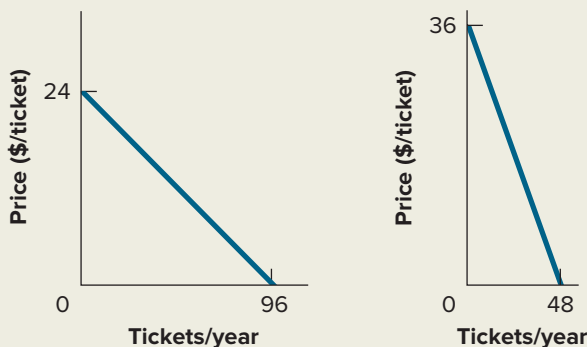
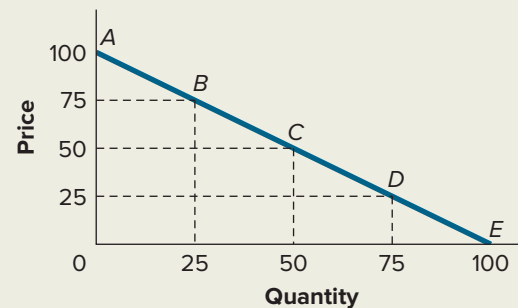
REVIEW QUESTIONS

1. Why do economists prefer to speak of demands arising out of “wants” rather than “needs”? (LO1)
2. Explain why a good or service that is offered at a monetary price of zero is unlikely to be a truly “free” good from an economic perspective. (LO2)
3. Why does a consumer’s price elasticity of demand for a good depend on the fraction of the consumer’s income spent on that good? (LO4)
4. Why does the price elasticity of demand for a good decline as we move down along a straight-line demand curve? (LO5)
5. Under what conditions will an increase in the price of a product lead to a reduction in total spending for that product? (LO6)
6. Why do economists pay little attention to the algebraic sign of the elasticity of demand for a good with respect to its own price, yet pay careful attention to the algebraic sign of the elasticity of demand for a good with respect to another good’s price? (LO7)

PROBLEMS



1. Which of the following factors would affect a buyer’s reservation price for a given good or service: social influence, the price of the good, or the cost of producing the item? (LO1)
2. In which type of restaurant would you expect the service to be more prompt and courteous: an expensive gourmet restaurant or an inexpensive diner? Explain. (LO2)
3. Ishana lives in Princeton, New Jersey, and commutes by train each day to her job in New York City (20 round-trips per month). When the price of a round-trip goes up from \$10 to \$20, she responds by consuming exactly the same number of trips as before, while spending \$200 per month less on restaurant meals. Does the fact that her quantity of train travel is completely unresponsive to the price increase imply that Ishana is not a rational consumer? (LO2)
- 4*. The buyers’ side of the market for amusement park tickets consists of two consumers whose demands are as shown in the diagram below. Graph the market demand curve for this market. (LO3)
5. Is the demand for a particular brand of car, like a Chevrolet, likely to be more or less price-elastic than the demand for all cars? Explain. (LO4)
6. Calculate the price elasticity of demand (in absolute value) at points A, B, C, D, and E on the demand curve below. (LO5)
7. The schedule below shows the number of packs of bagels bought in Davis, California, each day at a variety of prices. (LO5, LO6)

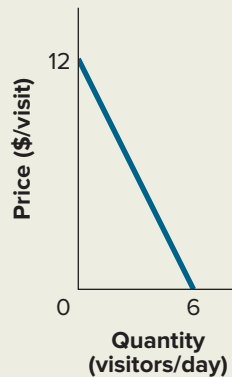


Price of bagels (\$/pack)	Number of packs purchased per day
6	0
5	3,000
4	6,000
3	9,000
2	12,000
1	15,000
0	18,000

- a. Graph the daily demand curve for packs of bagels in Davis.
- b. Calculate the price elasticity of demand at the point on the demand curve at which the price of bagels is \$4 per pack.

*Denotes more difficult problem.

- c. If all bagel shops increased the price of bagels from \$4 per pack to \$5 per pack, what would happen to total revenues?
 - d. Calculate the price elasticity of demand at a point on the demand curve where the price of bagels is \$1 per pack.
 - e. If bagel shops increased the price of bagels from \$1 per pack to \$2 per pack, what would happen to total revenues?
8. Suppose, while rummaging through your uncle's closet, you found the original painting of *Dogs*

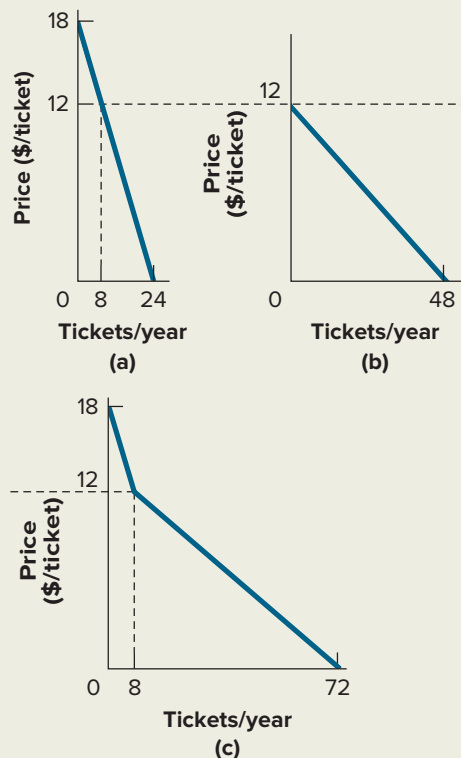


Playing Poker, a valuable piece of art. You decide to set up a display in your uncle's garage. The demand curve to see this valuable piece of art is as shown in the diagram. What price should you charge if your goal is to maximize your revenues from tickets sold? On a graph, show the inelastic and elastic regions of the demand curve. (LO5, LO6)

- 9* Suppose that, in an attempt to induce citizens to conserve energy, the government enacted regulations requiring that all air conditioners be more efficient in their use of electricity. After this regulation was implemented, government officials were then surprised to discover that people used even more electricity than before. Using the concept of price elasticity, explain how this increase might have occurred. (LO4, LO7)
10. A 2 percent increase in the price of milk causes a 4 percent reduction in the quantity demanded of chocolate syrup. What is the cross-price elasticity of demand for chocolate syrup with respect to the price of milk? Are the two goods complements or substitutes? (LO7)

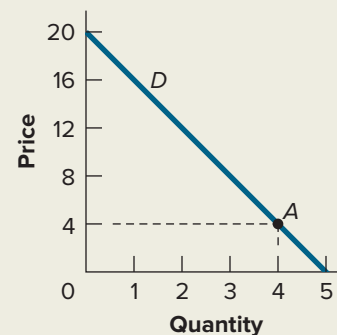
ANSWERS TO SELF-TESTS

- 3.1 Adding the two individual demand curves, (a) and (b), horizontally yields the market demand curve (c). (LO3)



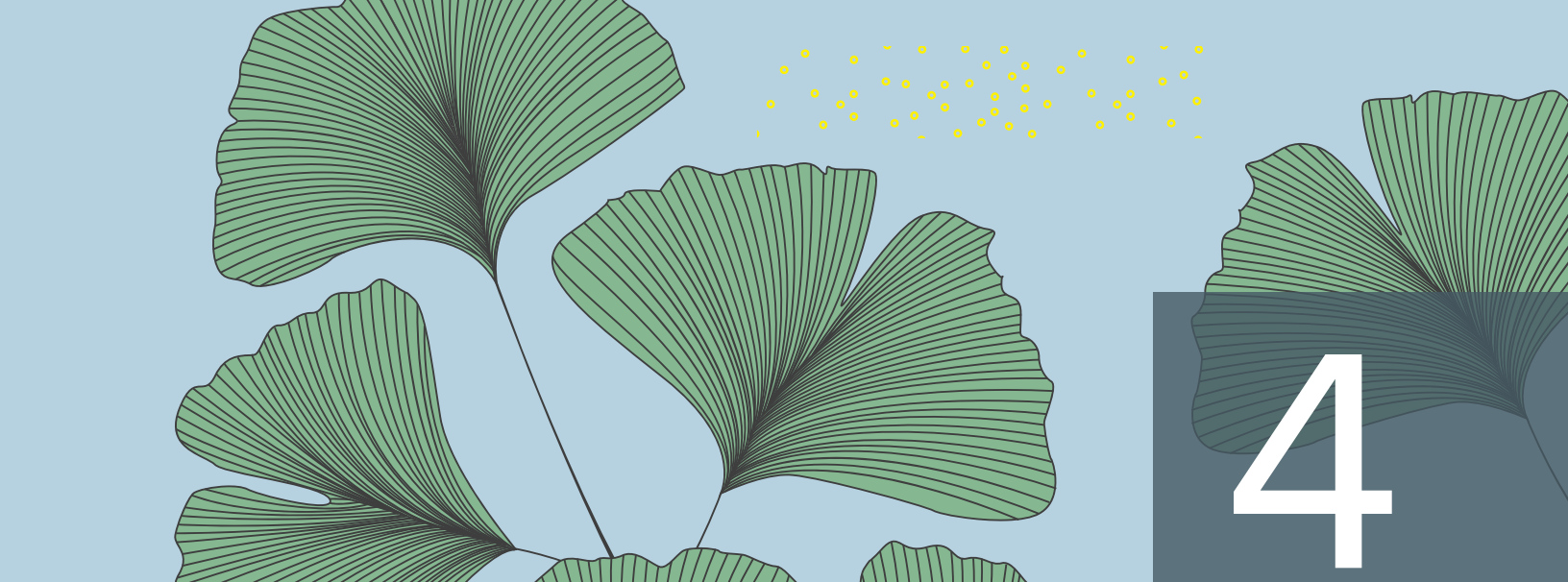
- 3.2 In response to a 5 percent reduction in the price of ski passes, the quantity demanded increased by 20 percent. The price elasticity of demand for ski passes is thus (20 percent)/(5 percent) = 4, and that means that at the initial price of \$400, the demand for ski passes is elastic with respect to price. (LO4)

- 3.3 At point A in the accompanying diagram, $P/Q = 4/4 = 1$. The slope of this demand curve is $20/5 = 4$, so $\epsilon = 1 \times (1/\text{slope}) = 1/4$. (LO5)



- 3.4 Income elasticity = Percentage change in quantity demanded/Percentage change in income = 5 percent/10 percent = 0.5. (LO7)

*Denotes more difficult problem.



4

LEARNING OBJECTIVES

After reading this chapter, you should be able to:

- LO1** Explain how opportunity cost is related to the supply curve.
- LO2** Discuss the relationship between the supply curve for an individual firm and the market supply curve for an industry.
- LO3** Explain why price equals marginal cost at the profit-maximizing output level.
- LO4** Match the determinants of supply with the factors that affect individual firms' costs.
- LO5** Define and calculate price elasticity of supply.

Perfectly Competitive Supply

Cars that took more than 50 hours to assemble in the 1970s are now built in less than 8 hours. Similar productivity growth has occurred in many other manufacturing industries. Yet in many service industries, productivity has grown only slowly, if at all. For example, the London Philharmonic Orchestra performs Beethoven's Fifth Symphony with no fewer musicians today than it did in 1850. And it still takes a barber about half an hour to cut someone's hair, just as it always has.

Given the spectacular growth in manufacturing workers' productivity, it's no surprise that their real wages have risen more than fivefold during the last century. But why have real wages for service workers risen just as much? If barbers and musicians are no more productive than they were at the turn of the century, why are they now paid five times as much?

An answer is suggested by the observation that the opportunity cost of pursuing any given occupation is the most one could have earned in some other occupation. Most people who become barbers or musicians could instead have chosen jobs in manufacturing. If workers in service industries were not paid roughly as much as they could have earned in other occupations, many of them would not have been willing to work in service industries in the first place.

The trajectories of wages in manufacturing and service industries illustrate the intimate link between the prices at which goods and services are offered for sale in the market and the opportunity cost of the resources required to produce them.

In Chapter 3, *Demand and Elasticity*, we saw that the demand curve is a schedule that tells how many units buyers wish to purchase at different prices. Our task here is to gain insight into the factors that shape the supply curve, the schedule that tells how many units suppliers wish to sell at different prices.

Although the demand side and the supply side of the market are different in several ways, many of these differences are superficial. Indeed, the behavior of both buyers and sellers is, in an important sense, fundamentally the same. After all, the two groups confront essentially similar questions—in the buyer's case,



Why are barbers paid five times as much now as in 1900, even though they can't cut hair any faster than they could then?

“Should I buy another unit?” and in the seller’s, “Should I sell another unit?” What is more, buyers and sellers use the same criterion for answering these questions. Thus, a rational consumer will buy another unit if its benefit exceeds its cost and a rational seller will sell another unit if the cost of making it is less than the extra revenue he can get from selling it (the familiar Cost-Benefit Principle again).

THINKING ABOUT SUPPLY: THE IMPORTANCE OF OPPORTUNITY COST

Do you live in a state that requires refundable soft drink container deposits? If so, you’ve probably noticed that some people always redeem their own containers while other people pass up this opportunity, leaving their used containers to be recycled by others. Recycling used containers is a service and its production obeys the same logic that applies to the production of other goods and services. The following sequence of recycling examples shows how the supply curve for a good or service is rooted in the individual’s choice of whether to produce it.

EXAMPLE 4.1 Opportunity Cost and Supply

How much time should Ushi spend recycling soft drink containers?

Ushi is trying to decide how to divide his time between his job as a dishwasher in the dining hall, which pays \$6 an hour for as many hours as he chooses to work, and gathering soft drink containers to redeem for deposit, in which case his pay depends on both the deposit per container and the number of containers he finds. Earnings aside, Ushi is indifferent between the two tasks, and the number of containers he’ll find depends on the number of hours per day he searches, as shown in the table below.

Search time (hours/day)	Total number of containers found	Additional number of containers found
0	0	
1	600	600
2	1,000	400
3	1,300	300
4	1,500	200
5	1,600	100

If the containers may be redeemed for 2 cents each, how many hours should Ushi spend searching for containers?

For each additional hour Ushi spends searching for soft drink containers, he loses the \$6 he could have earned as a dishwasher. This is his hourly opportunity cost of searching for soft drink containers. His benefit from each hour spent searching for containers is the number of additional containers he finds (shown in column 3 of the table) times the deposit he collects per container. Since he can redeem each container for 2 cents, his first hour spent collecting containers will yield earnings of $600(\$0.02) = \12 , or \$6 more than he could have earned as a dishwasher.

By the Cost-Benefit Principle, then, Ushi should spend his first hour of work each day searching for soft drink containers rather than washing dishes. A second hour searching for containers will yield 400 additional containers, for additional earnings of \$8, so it too satisfies the cost-benefit test. A third hour spent searching

yields 300 additional containers, for $300(\$0.02) = \6 of additional earnings. Since this is exactly what Ushi could have earned washing dishes, he's indifferent between spending his third hour of work each day on one task or the other. For the sake of discussion, however, we'll assume that he resolves ties in favor of searching for containers, in which case he'll spend three hours each day searching for containers.

Using the data provided in Example 4.1, what is the lowest redemption price that would induce Ushi to spend at least one hour per day recycling? Since he'll find 600 containers in his first hour of search, a 1 cent deposit on each container would enable him to match his \$6 per hour opportunity cost. More generally, if the redemption price is p , and the next hour spent searching yields ΔQ additional containers, then Ushi's additional earnings from searching the additional hour will be $p(\Delta Q)$. This means that the smallest redemption price that will lead Ushi to search another hour must satisfy the equation

$$p(\Delta Q) = \$6. \quad (4.1)$$

How high would the redemption price of containers have to be to induce Ushi to search for a second hour? Since he can find $\Delta Q = 400$ additional containers if he searches for a second hour, the smallest redemption price that will lead him to do so must satisfy $p(400) = \$6$, which solves for $p = 1.5$ cents.

SELF-TEST 4.1

In the example above, calculate the lowest container redemption prices that will lead Ushi to search a third, fourth, and fifth hour.

By searching for soft drink containers, Ushi becomes, in effect, a supplier of container-recycling services. In Self-Test 4.1, we saw that Ushi's reservation prices for his third, fourth, and fifth hours of container search are 2, 3, and 6 cents, respectively. Having calculated these reservation prices, we can now plot his supply curve of container-recycling services. This curve, which plots the redemption price per container on the vertical axis and the number of containers recycled each day on the horizontal axis, is shown in Figure 4.1. Ushi's individual supply curve of container-recycling services tells us the number of containers he is willing to recycle at various redemption prices.

The supply curve shown in Figure 4.1 is upward-sloping, just like those we saw in Chapter 2, *Supply and Demand*. There are exceptions to this general rule, but sellers of most goods will offer higher quantities at higher prices.

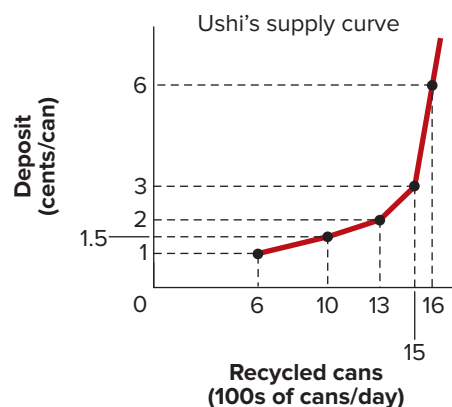


FIGURE 4.1

An Individual Supply Curve for Recycling Services.

When the deposit price increases, it becomes attractive to abandon alternative pursuits to spend more time searching for soft drink containers.

INDIVIDUAL AND MARKET SUPPLY CURVES

The relationship between the individual and market supply curves for a product is analogous to the relationship between the individual and market demand curves. The quantity that corresponds to a given price on the market demand curve is the sum of the quantities demanded at that price by all individual buyers in the market. Likewise, the quantity that corresponds to any given price on the market supply curve is the sum of the quantities supplied at that price by all individual sellers in the market.

Suppose, for example, that the supply side of the recycling-services market consists only of Ushi and his identical twin, Yoshi, whose individual supply curve is the same as Ushi's. To generate the market supply curve, we first put the individual supply curves side by side, as shown in Figure 4.2(a) and (b). We then announce a price, and for that price add the individual quantities supplied to obtain the total quantity supplied in the market. Thus, at a price of 3 cents per container, both Ushi and Yoshi wish to recycle 1,500 cans per day, so the total market supply at that price is 3,000 cans per day. Proceeding in like manner for a sequence of prices, we generate the market supply curve for recycling services shown in Figure 4.2(c). This is the same process of horizontal summation by which we generated market demand curves from individual demand curves in Chapter 3, *Demand and Elasticity*.

Alternatively, if there were many suppliers with individual supply curves identical to Ushi's, we could generate the market supply curve by simply multiplying each quantity value on the individual supply curve by the number of suppliers. For instance, Figure 4.3 shows the supply curve for a market in which there are 1,000 suppliers with individual supply curves like Ushi's.

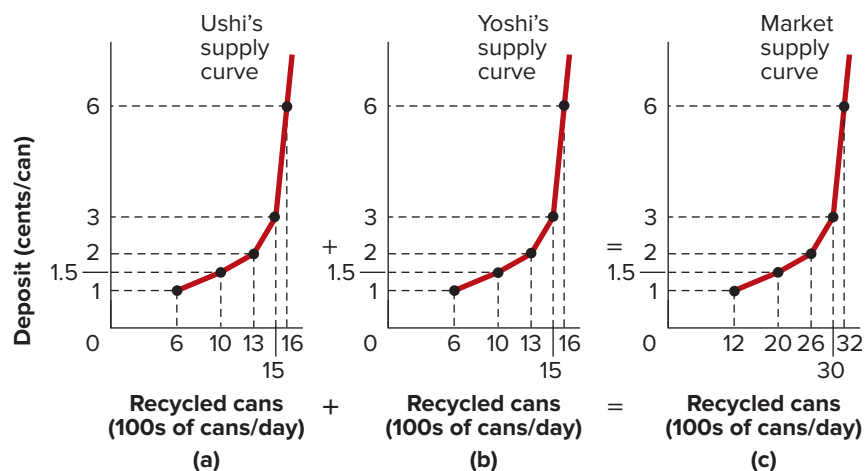
Why do individual supply curves tend to be upward-sloping? One explanation is that recyclers will tend to be more productive if they always look first for the containers that are easiest to find—such as those in plain view in readily accessible locations. As the redemption price rises, it will pay to incur the additional cost of searching farther from the beaten path.

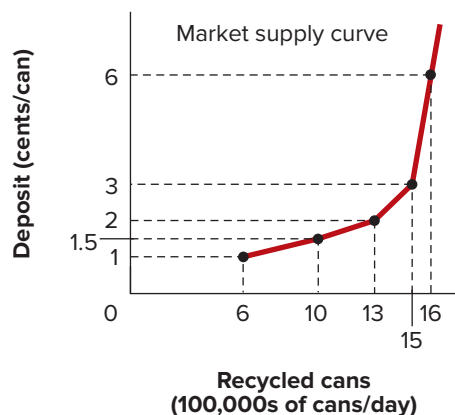
If all individuals have identical upward-sloping supply curves, the market supply curve will be upward-sloping as well. But there is an important additional reason for the positive slope of market supply curves: Individual suppliers generally differ with respect to their opportunity costs of supplying the product. Thus, whereas people facing unattractive employment opportunities in other occupations may be willing to recycle soft drink containers even when the redemption price is low, those with more attractive options will recycle only if the redemption price is relatively high.

FIGURE 4.2

The Market Supply Curve for Recycling Services.

To generate the market supply curve (c) from the individual supply curves (a) and (b), we add the individual supply curves horizontally.



**FIGURE 4.3****The Market Supply Curve with 1,000 Identical Sellers.**

To generate the market supply curve for a market with 1,000 identical sellers, we simply multiply each quantity value on the individual supply curve by 1,000.

In summary, then, the upward slope of the supply curve reflects the fact that costs tend to rise at the margin when producers expand production, partly because each individual exploits her most attractive opportunities first, but also because different potential sellers face different opportunity costs.

PROFIT-MAXIMIZING FIRMS IN PERFECTLY COMPETITIVE MARKETS

To explore the nature of the supply curve of a product more fully, we must say more about the goals of the organizations that supply the product and the kind of economic environment in which they operate. In virtually every economy, goods and services are produced by a variety of organizations that pursue a host of different motives. The Red Cross supplies blood because its organizers and donors want to help people in need; the local government fixes potholes because the mayor was elected on a promise to do so; karaoke singers perform because they like public attention; and car-wash employees are driven primarily by the hope of making enough money to pay their rent.

Profit Maximization

Notwithstanding this rich variety of motives, *most* goods and services that are offered for sale in a market economy are sold by private firms whose main reason for existing is to earn **profit** for their owners. A firm's profit is the difference between the total revenue it receives from the sale of its product and all costs it incurs in producing it.

A **profit-maximizing firm** is one whose primary goal is to maximize the amount of profit it earns. The supply curves that economists use in standard supply and demand theory are based on the assumption that goods are sold by profit-maximizing firms in **perfectly competitive markets**, which are markets in which individual firms have no influence over the market prices of the products they sell. Because of their inability to influence market price, perfectly competitive firms are often described as **price takers**.

The following four conditions are characteristic of markets that are perfectly competitive:

1. *All firms sell the same standardized product.* Although this condition is almost never literally satisfied, it holds as a rough approximation for many markets. Thus, the markets for concrete building blocks of a given size, or for apples of a given variety, may be described in this way. This condition implies that buyers are willing to switch from one seller to another if by so doing they can obtain a lower price.
2. *The market has many buyers and sellers, each of which buys or sells only a small fraction of the total quantity exchanged.* This condition implies that individual

profit the total revenue a firm receives from the sale of its product minus all costs—explicit and implicit—incurred in producing it

profit-maximizing firm a firm whose primary goal is to maximize the difference between its total revenues and total costs

perfectly competitive market a market in which no individual supplier has significant influence on the market price of the product

price taker a firm that has no influence over the price at which it sells its product

buyers and sellers will be price takers, regarding the market price of the product as a fixed number beyond their control. For example, a single farmer's decision to plant fewer acres of wheat would have no appreciable impact on the market price of wheat, just as an individual consumer's decision to become a vegetarian would have no perceptible effect on the price of beef.

3. *Productive resources are mobile.* This condition implies that if a potential seller identifies a profitable business opportunity in a market, he or she will be able to obtain the labor, capital, and other productive resources necessary to enter that market. By the same token, sellers who are dissatisfied with the opportunities they confront in a given market are free to leave that market and employ their resources elsewhere.
4. *Buyers and sellers are well informed.* This condition implies that buyers and sellers are aware of the relevant opportunities available to them. If that were not so, buyers would be unable to seek out sellers who charge the lowest prices, and sellers would have no means of deploying their resources in the markets in which they would earn the most profit.

The market for wheat closely approximates a perfectly competitive market. The market for operating systems for desktop computers, however, does not. Roughly 80 percent of desktop operating systems are sold by Microsoft, giving the company enough influence in that market to have significant control over the price it charges. For example, if it were to raise the price of its latest edition of Windows by, say, 20 percent, some consumers might switch to macOS or Linux, and others might postpone their purchase; but many—perhaps even most—would continue with their plans to buy Windows.

By contrast, if an individual wheat farmer were to charge even a few cents more than the current market price for a bushel of wheat, he wouldn't be able to sell any of his wheat at all. And since he can sell as much wheat as he wishes at the market price, he has no motive to charge less.

The Demand Curve Facing a Perfectly Competitive Firm

From the perspective of an individual firm in a perfectly competitive market, what does the demand curve for its product look like? Since it can sell as much or as little as it wishes at the prevailing market price, the demand curve for its product is perfectly elastic at the market price. Figure 4.4(a) shows the market demand and supply curves intersecting to determine a market price of P_0 . Figure 4.4(b) shows the product demand curve, D_i , as seen by any individual firm in this market, a horizontal line at the market price level P_0 .

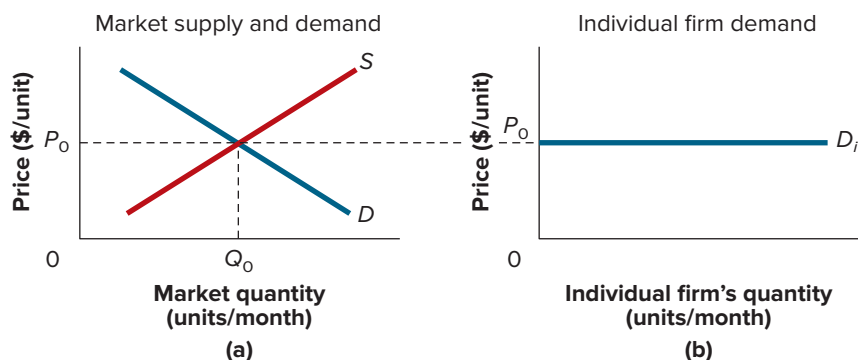
Many of the conclusions of the standard supply and demand model also hold for **imperfectly competitive firms** (or **price setters**)—those firms, like Microsoft, that

imperfectly competitive firm (or **price setter**) a firm that has at least some control over the market price of its product

FIGURE 4.4

The Demand Curve Facing a Perfectly Competitive Firm.

The market demand and supply curves intersect to determine the market price of the product (a). The individual firm's demand curve, D_i (b), is a horizontal line at the market price.



have at least some ability to vary their own prices. But certain other conclusions do not, as we shall see when we examine the behavior of such firms more closely in later chapters.

Since a perfectly competitive firm has no control over the market price of its product, it needn't worry about choosing the level at which to set that price. As we've seen, the equilibrium market price in a competitive market comes from the intersection of the industry supply and demand curves. The challenge confronting the perfectly competitive firm is to choose its output level so that it makes as much profit as it can at that price. As we investigate how the competitive firm responds to this challenge, we'll see that some costs are more important than others.

Production in the Short Run

To gain a deeper understanding of the origins of the supply curve, it is helpful to consider a perfectly competitive firm confronting the decision of how much to produce. The firm in question is a small company that makes glass bottles. To keep things simple, suppose that the silica required for making bottles is available free of charge from a nearby desert and that the only costs incurred by the firm are the wages it pays its employees and the lease payment on its bottle-making machine. The employees and the machine are the firm's only two **factors of production**—inputs used to produce goods and services. In more complex examples, factors of production also might include land, structures, entrepreneurship, and possibly others, but for the moment we consider only labor and capital.

When we refer to the **short run**, we mean a period of time during which at least some of the firm's factors of production cannot be varied. For our bottle maker, we will assume that the number of employees can be varied on short notice but that the capacity of its bottle-making machine can be altered only with significant delay. For this firm, then, the short run is simply that period of time during which the firm cannot alter the capacity of its bottle-making machine. By contrast, when we speak of the **long run**, we refer to a time period of sufficient length that all the firm's factors of production are variable.

Table 4.1 shows how the company's bottle production depends on the number of hours its employees spend on the job each day. The output–employment relationship described in Table 4.1 exhibits a pattern that is common to many such relationships. Each time we move down one row in the table, output grows by 100 bottles per day, but note in the right column that it takes larger and larger increases in the amount of labor to achieve this increase. Economists refer to this pattern as

factor of production an input used in the production of a good or service

short run a period of time sufficiently short that at least some of the firm's factors of production are fixed

long run a period of time of sufficient length that all the firm's factors of production are variable

TABLE 4.1
Employment and Output for a Glass Bottle Maker

Number of bottles per day	Number of employee-hours per day
0	0
100	1
200	2
300	4
400	7
500	11
600	16
700	22

law of diminishing returns
a property of the relationship between the amount of a good or service produced and the amount of a variable factor required to produce it; the law says that when some factors of production are fixed, increased production of the good eventually requires ever-larger increases in the variable factor

fixed factor of production
an input whose quantity cannot be altered in the short run

variable factor of production
an input whose quantity can be altered in the short run

fixed cost the sum of all payments made to the firm's fixed factors of production

the **law of diminishing returns**, and it always refers to situations in which at least some factors of production are fixed, which can be stated as follows:

Law of Diminishing Returns: When some factors of production are held fixed, increased production of the good eventually requires ever-larger increases in the variable factor.

Here, the **fixed factor** is the bottle-making machine, and the **variable factor** is labor. In the context of this example, the law of diminishing returns says simply that successive increases in the labor input eventually yield smaller and smaller increments in bottle output. (Strictly speaking, the law ought to be called the law of *eventually* diminishing returns because output may initially grow at an increasing rate with additional units of the variable factor.)

Typically, returns from additional units of the variable input eventually diminish because of some form of congestion. For instance, in an office with three secretaries and only a single desktop computer, we would not expect to get three times as many letters typed per hour as in an office with only one secretary because only one person can use a computer at a time.

Choosing Output to Maximize Profit

Suppose the lease payment for the company's bottle-making machine and the building that houses it is \$40 per day and must be paid whether the company makes any bottles or not. This payment is both a **fixed cost** (since it does not depend on the number of bottles per day the firm makes) and, for the duration of the lease, a sunk cost. For short, we'll refer to this cost as the company's *capital cost*. In the following examples, we'll explore how the company's decision about how many bottles to make depends on the price of bottles, the wage, and the cost of capital.

EXAMPLE 4.2 Maximizing Profit

If bottles sell for \$35 per hundred, and if the employee's wage is \$10 per hour, how many bottles should the company described above produce each day?

The company's goal is to maximize its profit, which is the difference between the revenue it collects from the sale of bottles and the cost of its labor and capital. Table 4.2 shows how the daily number of bottles produced (denoted *Q*) is related to the company's revenue, employment, costs, and profit.

TABLE 4.2
Output, Revenue, Costs, and Profit

(1) <i>Q</i> (bottles/day)	(2) Total revenue (\$/day)	(3) Total labor cost (\$/day)	(4) Total cost (\$/day)	(5) Profit (\$/day)
0	0	0	40	−40
100	35	10	50	−15
200	70	20	60	10
300	105	40	80	25
400	140	70	110	30
500	175	110	150	25
600	210	160	200	10
700	245	220	260	−15

To see how the entries in the table are constructed, let's examine the revenue, wage, cost, and profit values that correspond to 200 units of output (row 3 of Table 4.2). Total revenue is \$70, the company's receipts from selling 200 bottles at \$35 per hundred. To make 200 bottles, the firm's employee had to work 2 hours (see Table 4.1), and at a wage of \$10 per hour that translates into \$20 of total labor cost. When the firm's fixed capital cost of \$40 per day is added to its total labor cost, we get the **total cost** entry of \$60 per day in column 4. The firm's daily profit, finally, is total revenue – total cost = \$70 – \$60 = \$10, the entry in column 5.

From a glance at column 5 of Table 4.2, we see that the company's maximum profit, \$30 per day, occurs when it produces 400 bottles per day.

total cost the sum of all payments made to the firm's fixed and variable factors of production

EXAMPLE 4.3 Maximizing Profit: A Change in Price

If bottles sell for \$45 per hundred, and if the employee's wage is again \$10 per hour, how many bottles should the company described above produce each day?

As we see in the entries of Table 4.3, the only consequence of the change in selling price is that total revenue, and hence profit, is now higher than before at every output level. As indicated by the entries of column 5, the company now does best to produce 500 bottles per day, 100 more than when the price was only \$35 per hundred.

TABLE 4.3
Output, Revenue, Costs, and Profit

(1) Q (bottles/day)	(2) Total revenue (\$/day)	(3) Total labor cost (\$/day)	(4) Total cost (\$/day)	(5) Profit (\$/day)
0	0	0	40	–40
100	45	10	50	–5
200	90	20	60	30
300	135	40	80	55
400	180	70	110	70
500	225	110	150	75
600	270	160	200	70
700	315	220	260	55

EXAMPLE 4.4 Maximizing Profit: A Change in Hourly Wages

If bottles sell for \$35 per hundred, and if the employee's wage is now \$12 per hour, how many bottles should the company described above produce each day?

With a higher wage rate, labor costs are higher at every level of output, as shown in column 3 of Table 4.4, and maximum profit now occurs when the firm produces 300 bottles per day, or 100 fewer than when the wage rate was \$10 per hour.

TABLE 4.4
Output, Revenue, Costs, and Profit

(1) Q (bottles/day)	(2) Total revenue (\$/day)	(3) Total labor cost (\$/day)	(4) Total cost (\$/day)	(5) Profit (\$/day)
0	0	0	40	−40
100	35	12	52	−17
200	70	24	64	6
300	105	48	88	17
400	140	84	124	16
500	175	132	172	3
600	210	192	232	−22
700	245	264	304	−59

Consider one final variation in the example that follows.

EXAMPLE 4.5 Maximizing Profit: A Change in Capital Cost

If bottles sell for \$35 per hundred, and if the employee's wage is \$10 per hour, how many bottles should the company produce each day if its capital cost is now \$70 instead of \$40?

The entries in Table 4.5 are just like those in Table 4.2 except that each entry in the total cost column (4) is \$30 higher than before, with the result that each entry in the profit column (5) is \$30 lower. Note, however, that the profit-maximizing number of bottles to produce is again 400 per day, precisely the same as when capital cost was only \$40 per day. When the company produces 400 bottles daily, its daily profit is 0, but at any other output level its profit would have been negative—that is, it would have been incurring a loss.

TABLE 4.5
Output, Revenue, Costs, and Profit

(1) Q (bottles/day)	(2) Total revenue (\$/day)	(3) Total labor cost (\$/day)	(4) Total cost (\$/day)	(5) Profit (\$/day)
0	0	0	70	−70
100	35	10	80	−45
200	70	20	90	−20
300	105	40	110	−5
400	140	70	140	0
500	175	110	180	−5
600	210	160	230	−20
700	245	220	290	−45

Price Equals Marginal Cost: The Seller's Supply Rule

Marginal cost is by far the most important cost concept you will encounter in economics. As the name suggests, it is the rate at which the firm's total costs change when it expands its production.

The observation that the profit-maximizing quantity for a firm to supply does not depend on its fixed costs is not an idiosyncrasy of Example 4.5. That it holds true in general is an immediate consequence of the Cost-Benefit Principle, which says that a firm should increase its output if, and only if, the *extra* benefit exceeds the *extra* cost. If the firm expands production by 100 bottles per day, its benefit is the extra revenue it gets, which in this case is simply the price of 100 bottles. The cost of expanding production by 100 bottles is by definition the marginal cost of producing 100 bottles—the amount by which total cost increases when bottle production rises by 100 per day. The Cost-Benefit Principle thus tells us that the perfectly competitive firm should keep expanding production as long as the price of the product is greater than marginal cost.

When the law of diminishing returns applies (that is, when some factors of production are fixed), marginal cost goes up as the firm expands production. Under these circumstances, the firm's best option is to supply that level of output for which price and marginal cost are exactly equal.

Note in Example 4.5 that if the company's capital cost had been any more than \$70 per day, it would have made a loss at *every* possible level of output. As long as it still had to pay its capital cost, however, its best bet would have been to continue producing 400 bottles per day. It is better, after all, to experience a smaller loss than a larger one. If a firm in that situation expected conditions to remain the same, though, it would want to get out of the bottle business as soon as its equipment lease expired.

A Note on the Firm's Shut-Down Condition

It might seem that a firm that can sell as much output as it wishes at a constant market price would *always* do best in the short run by producing and selling the output level for which price equals marginal cost. But there are exceptions to this rule. Suppose, for example, that the market price of the firm's product falls so low that its revenue from sales is smaller than its **variable cost** when price equals marginal cost. The firm should then cease production for the time being. By shutting down, it will suffer a loss equal to its fixed costs. But by remaining open, it would suffer an even larger loss.

marginal cost the increase in total cost that results from carrying out one additional unit of an activity

variable cost the sum of all payments made to the firm's variable factors of production

SELF-TEST 4.2

In Example 4.5, suppose bottles sold not for \$35 per hundred but only \$5. Calculate the profit corresponding to each level of output, and verify that the firm's best option is to cease operations in the short run.

Graphing Marginal Cost

To plot the marginal cost curve for a specific company, we would need to know how total cost changes for every possible change in output. In the preceding examples, however, we know the firm's cost for only a small sample of production values. Even with this limited information, though, we can construct a reasonable approximation of the firm's marginal cost curve. Suppose again that the wage is \$35 per hour and that capital costs are again \$40 per day, so that we again have the production and cost relationships shown in Table 4.2, reproduced below as Table 4.6.

Note that, when the firm expands production from 100 to 200 bottles per day, its increase in cost is \$10. When we graph the marginal cost curve, what output level

TABLE 4.6
Output, Revenue, Costs, and Profit

Q (bottles/day)	Total revenue (\$/day)	Total labor cost (\$/day)	Total cost (\$/day)	Profit (\$/day)
0	0	0	40	-40
100	35	10	50	-15
200	70	20	60	10
300	105	40	80	25
400	140	70	110	30
500	175	110	150	25
600	210	160	200	10
700	245	220	260	-15

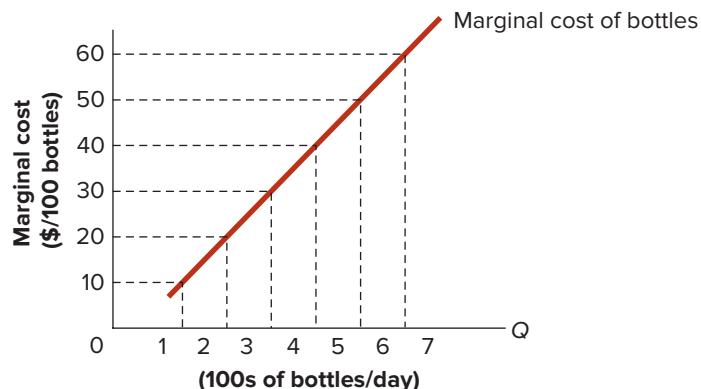
should this \$10 marginal cost correspond to? Strictly speaking, it corresponds to neither 100 nor 200, but to the movement between the two. On the graph we thus show the \$10 marginal cost value corresponding to an output level midway between 100 and 200 bottles per day—namely, 150 bottles per day, as in Figure 4.5. Similarly, when the firm expands from 200 to 300 bottles per day, its costs go up by \$20, so we plot a marginal cost of \$20 with the output level 250 in Figure 4.5. Proceeding in this fashion, we generate the marginal cost curve shown in the diagram.

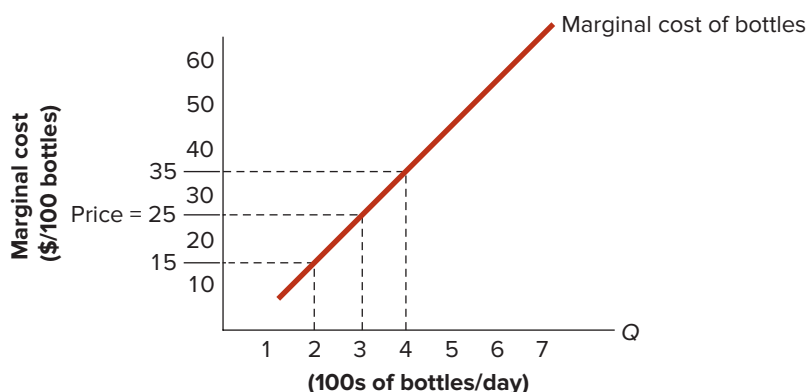
Suppose the market price facing the seller whose marginal cost curve is shown in Figure 4.5 is \$25 per hundred. If the firm's goal is to make as much profit as possible, how many bottles should it sell? It should sell the quantity for which marginal cost is equal to \$25 per hundred, and as we see in Figure 4.6, that quantity is 300 bottles per week.

To gain further confidence that 300 must be the profit-maximizing quantity when the price is \$25 per hundred, first suppose that the firm had sold some amount less than that—say, only 200 bottles per day. Its benefit from expanding output by one bottle would then be the bottle's market price, here 25 cents (since bottles sell for \$25 per hundred, each individual bottle sells for 25 cents). The cost of expanding output by one bottle is equal (by definition) to the firm's marginal cost, which at 200 bottles per day is only $\$15/100 = 15$ cents (see Figure 4.6). So by selling the 201st bottle for 25 cents and producing it for an extra cost of only 15 cents, the firm will increase its profit by $25 - 15 = 10$ cents/day. In a similar way, we can show that for *any* quantity less than the level at which price equals marginal cost, the seller can boost profit by expanding production.

FIGURE 4.5**The Firm's Marginal Cost of Production.**

The firm's cost goes up by \$10 when it expands production from 100 bottles per day to 200 bottles per day. The marginal cost of the increased output is thus \$10, and by convention we plot that value at a point midway between 100 and 200 bottles per day.




FIGURE 4.6
Price = Marginal Cost: The Perfectly Competitive Firm's Profit-Maximizing Supply Rule.

If price is greater than marginal cost, the firm can increase its profit by expanding production and sales. If price is less than marginal cost, the firm can increase its profit by producing and selling less output.

Conversely, suppose that the firm is currently selling more than 300 bottles per day—say, 400—at a price of \$25 per hundred. From Figure 4.6 we see that marginal cost at an output of 400 is $\$35/100 = 35$ cents per bottle. If the firm then contracts its output by one bottle per day, it would cut its costs by 35 cents while losing only 25 cents in revenue. As before, its profit would grow by 10 cents per day. The same argument can be made regarding any quantity larger than 300, so if the firm is currently selling an output at which price is less than marginal cost, it can always do better by producing and selling fewer bottles.

We have thus established that if the firm were selling fewer than 300 bottles per day, it could earn more profit by expanding. If it were selling more than 300, it could earn more by contracting. It follows that at a market price of \$25 per hundred, the seller does best by selling 300 units per week, the quantity for which price and marginal cost are exactly the same.

SELF-TEST 4.3

For a bottle price of \$25 per hundred, calculate the profit corresponding to each level of output, as in Table 4.6, and verify that the profit-maximizing output is 300 bottles per day.

As further confirmation of the claim that the perfectly competitive firm maximizes profit by setting price equal to marginal cost, note in Figure 4.6 that, when marginal cost is equal to a price of \$35 per hundred bottles, the corresponding quantity is 400 bottles per day. This is the same as the profit-maximizing quantity we identified for that price in Table 4.6.

The “Law” of Supply

The law of demand tells us that consumers buy less of a product when its price rises. If there were an analogous law of supply, it would say that producers offer more of a product for sale when its price rises. Is there such a law? We know that supply curves are essentially marginal cost curves and that because of the law of diminishing returns, marginal cost curves are upward-sloping in the short run. And so there is indeed a law of supply that applies as stated in the short run.

In the long run, however, the law of diminishing returns does not apply. (Recall that it holds only if at least some factors of production are fixed.) Because firms can vary the amounts of *all* factors of production they use in the long run, they can often double their production by simply doubling the amount of each input they use. In such cases, costs would be exactly proportional to output and the firm's marginal

cost curve in the long run would be horizontal, not upward-sloping. So for now we'll say only that the "law" of supply holds as stated in the short run but not necessarily in the long run. For both the long run and the short run, however, *the perfectly competitive firm's supply curve is its marginal cost curve.*¹

Every quantity of output along the market supply curve represents the summation of all the quantities individual sellers offer at the corresponding price. So the correspondence between price and marginal cost exists for the market supply curve as well as for the individual supply curves that lie behind it. That is, *for every price–quantity pair along the market supply curve, price will be equal to each seller's marginal cost of production.*

This is why we sometimes say that the supply curve represents the cost side of the market, whereas the demand curve represents the benefit side of the market. At every point along a market demand curve, price represents what buyers would be willing to pay for an additional unit of the product—and this, in turn, is how we measure the amount by which they would benefit by having an additional unit of the product. Likewise, at every point along a market supply curve, price measures what it would cost producers to expand production by one unit.



► Visit your instructor's Connect course and access your eBook to view this video.



Why are brown eggs more expensive than white ones?

Applying the Theory of Supply

Whether the activity is producing new soft drink containers or recycling used ones, or indeed any other production activity at all, the same logic governs all supply decisions in perfectly competitive markets (and in any other setting in which sellers can sell as much as they wish to at a constant price): keep expanding output until marginal cost is equal to the price of the product. This logic helps us understand why recycling efforts are more intensive for some products than others.



The Economic Naturalist 4.1

When recycling is left to private market forces, why are many more aluminum beverage containers recycled than glass ones?

In both cases, recyclers gather containers until their marginal costs are equal to the containers' respective redemption prices. When recycling is left to market forces, the redemption price for a container is based on what companies can sell it (or the materials in it) for. Aluminum containers can be easily processed into scrap aluminum, which commands a high price, and this leads profit-seeking companies to offer a high redemption price for aluminum cans. By contrast, the glass from which glass containers are made has only limited resale value, primarily because the raw materials required to make new glass containers are so cheap. This difference leads profit-seeking companies to offer much lower redemption prices for glass containers than for aluminum ones.

The high redemption prices for aluminum cans induce many people to track down these cans, whereas the low redemption prices for glass containers lead most people to ignore them. If recycling is left completely to market forces, then we would expect to see aluminum soft drink containers quickly recycled, whereas glass containers would increasingly litter the landscape. This is in fact the pattern we do see in states without recycling laws. (More on how these laws work in a moment.) This pattern is a simple consequence of the fact that the supply curves of container-recycling services are upward-sloping.



In states that don't have beverage container deposit laws, why are aluminum cans more likely to be recycled than glass bottles?

¹Again, this rule holds subject to the provision that total revenue exceed variable production cost at the output level for which price equals marginal cost.

The acquisition of valuable raw materials is only one of two important benefits from recycling. The second is that, by removing litter, recycling makes the environment more pleasant for everyone. As the next example suggests, this second benefit might easily justify the cost of recycling substantial numbers of glass containers.

EXAMPLE 4.6 Why the Optimal Amount of Pollution Isn't Zero

What is the socially optimal amount of recycling of glass containers?

Suppose that the 60,000 citizens of Burlington, Vermont, would collectively be willing to pay 6 cents for each glass container removed from their local environment. If the local market supply curve of glass container-recycling services is as shown in Figure 4.7, what is the socially optimal level of glass container recycling?

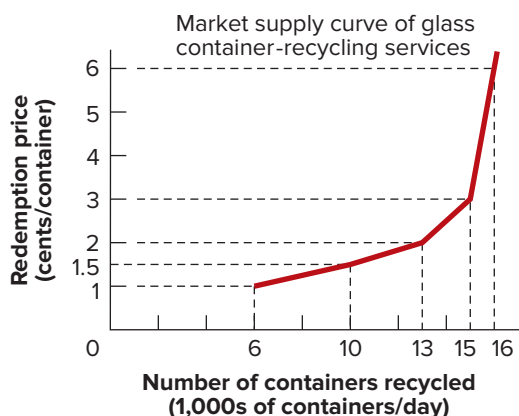


FIGURE 4.7

The Supply Curve of Container-Recycling Services for Burlington, Vermont.

Suppose the citizens of Burlington authorize their city government to collect tax money to finance litter removal. If the benefit of each glass container removed, as measured by what residents are collectively willing to pay, is 6 cents, the government should offer to pay 6 cents for each glass container recycled. To maximize the total economic surplus from recycling, we should recycle that number of containers for which the marginal cost of recycling is equal to the 6-cent marginal benefit. Given the market supply curve shown, the optimal quantity is 16,000 containers per day, and that is how many will be redeemed when the government offers 6 cents per container.

Although 16,000 containers per day will be removed from the environment in the preceding example, others will remain. After all, some are discarded in remote locations, and a redemption price of 6 cents per container is simply not high enough to induce people to track them all down.

So why not offer an even higher price and get rid of *all* glass container litter? For the example given, the reason is that the marginal cost of removing the 16,001st glass container each day is greater than the benefit of removing it. Total economic surplus is largest when we remove litter only up to the point that the marginal benefit of litter removal is equal to its marginal cost, which occurs when 16,000 containers per day are recycled. To proceed past that point is actually wasteful.

Many people become upset when they hear economists say that the socially optimal amount of litter is greater than zero. In the minds of these people, the optimal amount of litter is *exactly* zero. But this position completely ignores the Cost-Benefit Principle. Granted, there would be benefits from reducing litter further, but there also would be costs. Spending more on litter removal therefore means spending less on other useful things. No one would insist that the optimal amount of dirt in his own home is zero. (If someone does make this claim, ask him why he doesn't stay home all day vacuuming the dust that is accumulating in his absence.) If it doesn't pay to remove all the dust from



James Hardy/PhotoAlto

Is the socially optimal quantity of litter zero?

your house, it doesn't pay to remove all the bottles from the environment. Precisely the same logic applies in each case.

If 16,000 containers per day is the optimal amount of litter removal, can we expect the individual spending decisions of private citizens to result in that amount of litter removal? Unfortunately we cannot. The problem is that anyone who paid for litter removal individually would bear the full cost of those services while reaping only a tiny fraction of the benefit. In Example 4.6, the 60,000 citizens of Burlington reaped a total benefit of 6 cents per container removed, which means a benefit of only $(6/60,000) = 0.0001$ cent per person! Someone who paid 6 cents for someone else to remove a container would thus be incurring a cost 60,000 times greater than his share of the resulting benefit.

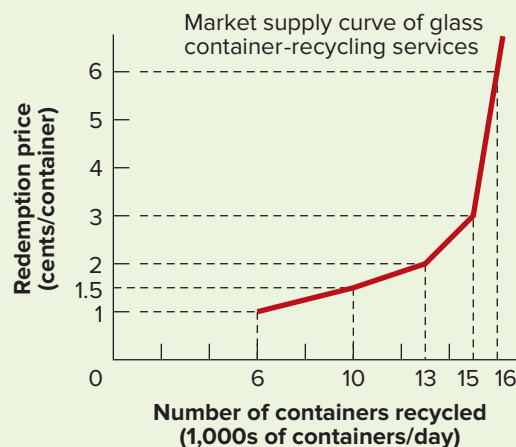
Note that the incentive problem here is similar to the one discussed in Chapter 2, *Supply and Demand*, for the person deciding whether to be vaccinated against an illness. The problem was that the incentive to be vaccinated was too weak because, even though the patient bears the full cost of the vaccination, many of the resulting benefits accrue to others. Thus, an important part of the extra benefit from any one person being vaccinated is that others also become less likely to contract the illness.

The case of glass container litter is an example in which private market forces do not produce the best attainable outcome for society as a whole. Even people who carelessly toss containers on the ground, rather than recycle them, are often offended by the unsightly landscape to which their own actions contribute. Indeed, this is why they often support laws mandating adequate redemption prices for glass containers.

People who litter do so not because they don't care about the environment, but because their private incentives make littering misleadingly attractive. Recycling requires some effort, after all, yet no individual's recycling efforts have a noticeable effect on the quality of the environment. The soft drink container deposit laws enacted by numerous states were a simple way to bring individual interests more closely into balance with the interests of society as a whole. The vast majority of container litter disappeared almost overnight in states that enacted these laws.

SELF-TEST 4.4

If the supply curve of glass container-recycling services is as shown in the diagram, and each of the city's 60,000 citizens would be willing to pay 0.00005 cent for each glass container removed from the landscape, at what level should the city government set the redemption price for glass containers, and how many will be recycled each day?



RECAP ↑**PROFIT-MAXIMIZING FIRMS IN PERFECTLY COMPETITIVE MARKETS**

The perfectly competitive firm faces a horizontal demand curve for its product, meaning that it can sell any quantity it wishes at the market price. In the short run, the firm's goal is to choose the level of output that maximizes its profits. It will accomplish this by choosing the output level for which its marginal cost is equal to the market price of its product.

DETERMINANTS OF SUPPLY REVISITED

What factors give rise to changes in supply? (Again, remember that a “change in supply” refers to a shift in the entire supply curve, as opposed to a movement along the curve, which we call a “change in the quantity supplied.”) A seller will offer more units if the benefit of selling extra output goes up relative to the cost of producing it. And since the benefit of selling output in a perfectly competitive market is a fixed market price that is beyond the seller's control, our search for factors that influence supply naturally focuses on the cost side of the calculation. The preceding examples suggest why the following factors, among others, will affect the likelihood that a product will satisfy the cost-benefit test for a given supplier.

Technology

Perhaps the most important determinant of production cost is technology. Improvements in technology make it possible to produce additional units of output at lower cost. This shifts each individual supply curve downward (or, equivalently, to the right) and hence shifts the market supply curve downward as well. Over time, the introduction of more sophisticated machinery has resulted in dramatic increases in the number of goods produced per hour of effort expended. Every such development gives rise to a rightward shift in the market supply curve.

But how do we know technological change will reduce the cost of producing goods and services? Might not new equipment be so expensive that producers who used it would have higher costs than those who relied on earlier designs? If so, then rational producers simply would not use the new equipment. The only technological changes that rational producers will adopt are those that will reduce their cost of production.

Input Prices

Whereas technological change generally (although not always) leads to gradual shifts in supply, changes in the prices of important inputs can give rise to large supply shifts literally overnight. As discussed in Chapter 3, *Demand and Elasticity*, for example, the price of crude oil, which is the most important input in the production of gasoline, often fluctuates sharply, and the resulting shifts in supply cause gasoline prices to exhibit corresponding fluctuations.

Similarly, when wage rates rise, the marginal cost of any business that employs labor also rises, shifting supply curves to the left (or, equivalently, upward). When interest rates fall, the opportunity cost of capital equipment also falls, causing supply to shift to the right.

The Number of Suppliers

Just as demand curves shift to the right when population grows, supply curves also shift to the right as the number of individual suppliers grows. For example, if container recyclers die or retire at a higher rate than new recyclers enter the industry, the

supply curve for recycling services will shift to the left. Conversely, if a rise in the unemployment rate leads more people to recycle soft drink containers (by reducing the opportunity cost of time spent recycling), the supply curve of recycling services will shift to the right.

Expectations

Expectations about future price movements can affect how much sellers choose to offer in the current market. Suppose, for example, that recyclers expect the future price of aluminum to be much higher than the current price because of growing use of aluminum components in cars. The rational recycler would then have an incentive to withhold aluminum from the market at today's lower price, thereby to have more available to sell at the higher future price. Conversely, if recyclers expected next year's price of aluminum to be lower than this year's, their incentive would be to offer more aluminum for sale in today's market.

Changes in Prices of Other Products

Apart from technological change, perhaps the most important determinant of supply is variation in the prices of other goods and services that sellers might produce. Prospectors, for example, search for those precious metals for which the surplus of benefits over costs is greatest. When the price of silver rises, many stop looking for gold and start looking for silver. Conversely, when the price of platinum falls, many platinum prospectors shift their attention to gold.

RECAP

THE DETERMINANTS OF SUPPLY

Among the relevant factors causing supply curves to shift are new technologies, changes in input prices, changes in the number of sellers, expectations of future price changes, and changes in the prices of other products that firms might produce.

price elasticity of supply the percentage change in the quantity supplied that occurs in response to a 1 percent change in the price of a good or service

THE PRICE ELASTICITY OF SUPPLY

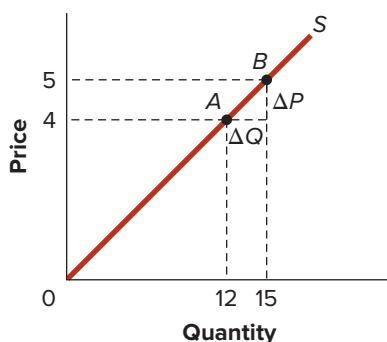
On the buyer's side of the market, we use price elasticity of demand to measure the responsiveness of quantity demanded to changes in price. On the seller's side of the market, the analogous measure is **price elasticity of supply**. It is defined as the percentage change in quantity supplied that occurs in response to a 1 percent change in price. For example, if a 1 percent increase in the price of peanuts leads to a 2 percent increase in the quantity supplied, the price elasticity of supply of peanuts would be 2.

The mathematical formula for price elasticity of supply at any point is the same as the corresponding expression for price elasticity of demand

$$\text{Price elasticity of supply} = \frac{\Delta Q/Q}{\Delta P/P}, \quad (4.2)$$

where P and Q are the price and quantity at that point, ΔP is a small change in the initial price, and ΔQ the resulting change in quantity.

As with the corresponding expression for price elasticity of demand, Equation 4.2 can be rewritten as $(P/Q) \times (\Delta Q/\Delta P)$. And since $(\Delta Q/\Delta P)$ is the reciprocal of the slope of the supply curve, the right-hand side of Equation 4.2 is equal to $(P/Q) \times (1/\text{slope})$ —the same expression we saw for price elasticity of demand. Price and quantity are always positive, as is the slope of the typical supply curve, which implies that price elasticity of supply will be a positive number at every point.

**FIGURE 4.8****Calculating the Price Elasticity of Supply Graphically.**

Price elasticity of supply is $(P/Q) \times (1/\text{slope})$, which at A is $(4/12) \times (12/4) = 1$, exactly the same as at B. The price elasticity of supply is equal to 1 at any point along a straight-line supply curve that passes through the origin.

Consider the supply curve shown in Figure 4.8. The slope of this supply curve is $1/3$, so the reciprocal of this slope is 3. Using the formula, this means that the price elasticity of supply at A is $(4/12) \times (3) = 1$. The corresponding expression at B, $(5/15) \times (3)$, yields exactly the same value. Because the ratio P/Q is the same at every point along this supply curve, price elasticity of supply will be exactly 1 at every point along this curve. Note the contrast between this result and our earlier finding that price elasticity of demand declines as we move downward along any straight-line demand curve.

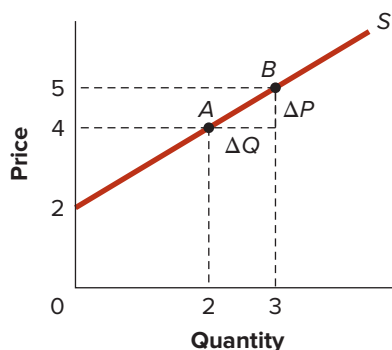
The special property that explains why price elasticity equals 1 at every point in this illustration is the fact that the supply curve was a straight line through the origin. For movements along any such line, both price and quantity always change in exactly the same proportion.

Elasticity is not constant, however, along straight-line supply curves like the one in Figure 4.9, which does not pass through the origin. Although the slope of this supply curve is equal to 1 at every point, the ratio P/Q declines as we move to the right along the curve. Elasticity at A is equal to $(4/2) \times (1) = 2$, and declines to $(5/3) \times (1) = 5/3$ at B.

SELF-TEST 4.5

For the supply curve shown in Figure 4.9, calculate the elasticity of supply when $P = 3$.

On the buyer's side of the market, two important polar cases were demand curves with infinite price elasticity and zero price elasticity. As the next two examples illustrate, analogous polar cases exist on the seller's side of the market.

**FIGURE 4.9****A Supply Curve for Which Price Elasticity Declines as Quantity Rises.**

For the supply curve shown, $(1/\text{slope})$ is the same at every point, but the ratio P/Q declines as Q increases. So elasticity $= (P/Q) \times (1/\text{slope})$ declines as quantity increases.

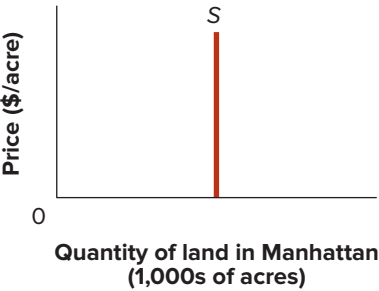
EXAMPLE 4.7 **Perfectly Inelastic**

What is the elasticity of supply of land within the borough limits of Manhattan?

perfectly inelastic supply supply is perfectly inelastic with respect to price if elasticity is zero

Land in Manhattan sells in the market for a price, just like aluminum or corn or automobiles or any other product. And the demand for land in Manhattan is a downward-sloping function of its price. For all practical purposes, however, its supply is completely fixed. No matter whether its price is high or low, the same amount of it is available in the market. The supply curve of such a good is vertical, and its price elasticity is zero at every price. Supply curves like the one shown in Figure 4.10 are said to be **perfectly inelastic**.

FIGURE 4.10
A Perfectly Inelastic Supply Curve.
Price elasticity of supply is zero at every point along a vertical supply curve.



EXAMPLE 4.8 **Perfectly Elastic Supply**

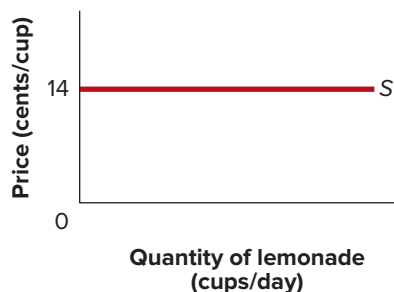
What is the elasticity of supply of lemonade?

Suppose that the ingredients required to bring a cup of lemonade to market and their respective costs are as follows.

Paper cup	2.0 cents
Lemon	3.8 cents
Sugar	2.0 cents
Water	0.2 cent
Ice	1.0 cent
Labor (30 seconds @ \$6/hour)	5.0 cents

If these proportions remain the same no matter how many cups of lemonade are made, and the inputs can be purchased in any quantities at the stated prices, draw the supply curve of lemonade and compute its price elasticity.

Since each cup of lemonade costs exactly 14 cents to make, no matter how many cups are made, the marginal cost of lemonade is constant at 14 cents per cup. And since each point on a supply curve is equal to marginal cost (see Chapter 2, *Supply and Demand*), this means that the supply curve of lemonade is not upward-sloping but is instead a horizontal line at 14 cents per cup (Figure 4.11). The price elasticity of supply of lemonade is infinite.

**FIGURE 4.11****A Perfectly Elastic Supply Curve.**

The elasticity of supply is infinite at every point along a horizontal supply curve.

Whenever additional units of a good can be produced by using the same combination of inputs, purchased at the same prices, as have been used so far, the supply curve of that good will be horizontal. Such supply curves are said to be **perfectly elastic**.

perfectly elastic supply supply is perfectly elastic with respect to price if elasticity of supply is infinite

Determinants of Supply Elasticity

The two preceding examples suggest some of the factors that govern the elasticity of supply of a good or service. The lemonade case was one whose production process was essentially like a cooking recipe. For such cases, we can exactly double our output by doubling each ingredient. If the price of each ingredient remains fixed, the marginal cost of production for such goods will be constant—and hence their horizontal supply curves.

The Manhattan land example is a contrast in the extreme. The inputs that were used to produce land in Manhattan—even if we knew what they were—could not be duplicated at any price.

The key to predicting how elastic the supply of a good will be with respect to price is to know the terms on which additional units of the inputs involved in producing that good can be acquired. In general, the more easily additional units of these inputs can be acquired, the higher price elasticity of supply will be. The following factors (among others) govern the ease with which additional inputs can be acquired by a producer.

Flexibility of Inputs

To the extent that production of a good requires inputs that are also useful for the production of other goods, it is relatively easy to lure additional inputs away from their current uses, making supply of that good relatively elastic with respect to price. Thus the fact that lemonade production requires labor with only minimal skills means that a large pool of workers could shift from other activities to lemonade production if a profitable opportunity arose. Brain surgery, by contrast, requires elaborately trained and specialized labor, which means that even a large price increase would not increase available supplies, except in the very long run.

Mobility of Inputs

If inputs can be easily transported from one site to another, an increase in the price of a product in one market will enable a producer in that market to summon inputs from other markets. For example, the supply of agricultural products is made more elastic with respect to price by the fact that thousands farmworkers are willing to migrate northward during the growing season. The supply of entertainment is similarly made more elastic by the willingness of entertainers to hit the road. Cirque

performers, lounge singers, comedians, and dancers often spend a substantial fraction of their time away from home.

For most goods, the price elasticity of supply increases each time a new highway is built, or when the telecommunications network improves, or indeed when any other development makes it easier to find and transport inputs from one place to another.

Ability to Produce Substitute Inputs

The inputs required to produce finished diamond gemstones include raw diamond crystal, skilled labor, and elaborate cutting and polishing machinery. In time, the number of people with the requisite skills can be increased, as can the amount of specialized machinery. The number of raw diamond crystals buried in the earth is probably fixed in the same way that Manhattan land is fixed, but unlike Manhattan land, rising prices will encourage miners to spend the effort required to find a larger proportion of those crystals. Still, the supply of natural gemstone diamonds tends to be relatively inelastic because of the difficulty of augmenting the number of diamond crystals.

The day is close at hand, however, when gemstone makers will be able to produce synthetic diamond crystals that are indistinguishable from real ones. Indeed, there are already synthetic crystals that fool even highly experienced jewelers. The introduction of a perfect synthetic substitute for natural diamond crystals would increase the price elasticity of supply of diamonds (or, at any rate, the price elasticity of supply of gemstones that look and feel just like diamonds).

Time

Because it takes time for producers to switch from one activity to another, and because it takes time to build new machines and factories and train additional skilled workers, the price elasticity of supply will be higher for most goods in the long run than in the short run. In the short run, a manufacturer's inability to augment existing stocks of equipment and skilled labor may make it impossible to expand output beyond a certain limit. But if a shortage of managers was the bottleneck, new MBAs can be trained in only two years. Or if a shortage of legal staff is the problem, new lawyers can be trained in three years. In the long run, firms can always buy new equipment, build new factories, and hire additional skilled workers.

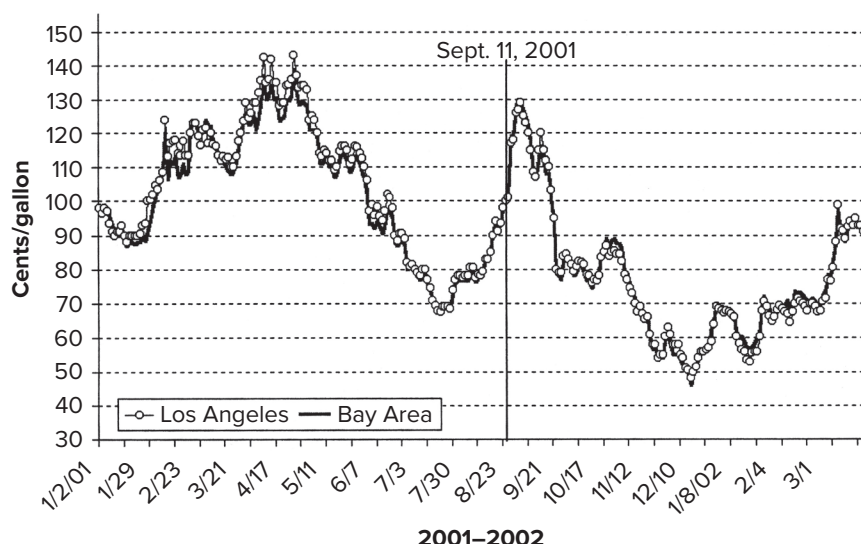
The conditions that gave rise to the perfectly elastic supply curve for lemonade in the example we discussed earlier are satisfied for many other products in the long run. If a product can be copied (in the sense that any company can acquire the design and other technological information required to produce it), and if the inputs needed for its production are used in roughly fixed proportions and are available at fixed market prices, then the long-run supply curve for that product will be horizontal. But many products do not satisfy these conditions, and their supply curves remain steeply upward-sloping, even in the very long run.



The Economic Naturalist 4.2

Why are gasoline prices so much more volatile than car prices?

Automobile price changes in the United States usually occur just once a year, when manufacturers announce an increase of only a few percentage points. In contrast, gasoline prices often fluctuate wildly from day to day. As shown in Figure 4.12, for example, the highest daily gasoline prices in California's two largest cities were three times higher than the lowest daily prices in 2001 and early 2002. Why this enormous difference in volatility?

**FIGURE 4.12**

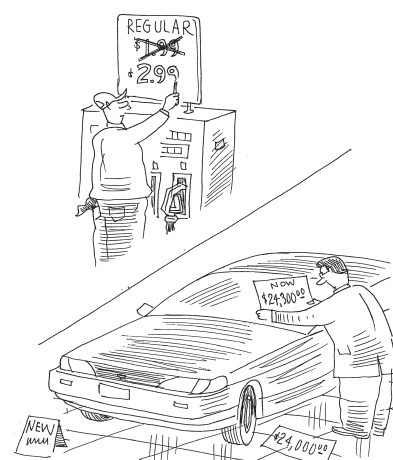
Gasoline Prices in Two California Cities.

Source: Oil Price Information Service, www.opisnet.com.

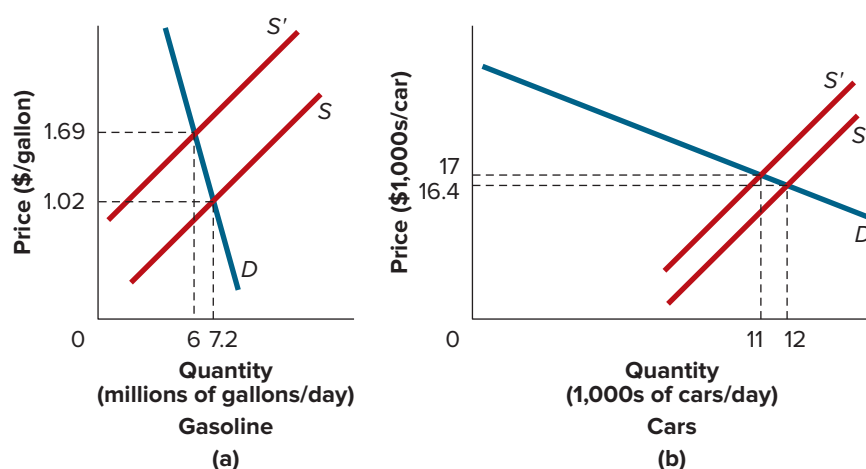
With respect to price volatility, at least two important features distinguish the gasoline market from the market for cars. One is that the short-run price elasticity of demand for gasoline is much smaller than the corresponding elasticity for cars. The other is that supply shifts are much more pronounced and frequent in the gasoline market than in the car market. (See Figure 4.13.)

Why are the two markets different in these ways? Consider first the difference in price elasticities of demand. The quantity of gasoline we demand depends largely on the kinds of cars we own and the amounts we drive them. In the short run, car ownership and commuting patterns are almost completely fixed, so even if the price of gasoline were to change sharply, the quantity we demand would not change by much. In contrast, if there were a sudden dramatic change in the price of cars, we could always postpone or accelerate our next car purchases.

To see why the supply curve in the gasoline market experiences larger and more frequent shifts than the supply curve in the car market, we need only examine the relative stability of the inputs employed by sellers in these two markets. Most of the inputs used in producing cars—steel, glass, rubber, plastics, electronic components, labor, and others—are reliably available to



Why are gasoline prices so much less stable than automobile prices?

**FIGURE 4.13**

Greater Volatility in Gasoline Prices Than in Car Prices.

Gasoline prices are more volatile because supply shifts are larger and more frequent in the gasoline market (a) than in the car market (b) and also because supply and demand are less elastic in the short run in the gasoline market.

carmakers. In contrast, the key input used in making gasoline—crude oil—is subject to profound and unpredictable supply interruptions.

This is so in part because much of the world's supply of crude oil is controlled by the Organization of the Petroleum Exporting Countries (OPEC), which has sharply curtailed its oil shipments to the United States on several previous occasions. Even in the absence of formal OPEC action, however, large supply curtailments often occur in the oil market—for example, whenever producers fear that political instability might engulf the major oil-producing countries of the Middle East.

Note in Figure 4.12 the sharp spike in gasoline prices that occurred just after the terrorist attacks on the World Trade Center and Pentagon on September 11, 2001. Because many believed that the aim of these attacks was to provoke large-scale war between Muslim societies and the West, fears of an impending oil supply interruption were perfectly rational. Similar oil price spikes occurred in the early months of 2011, when political upheaval in several Middle Eastern countries threatened to disrupt oil supplies. Such fears alone can trigger a temporary supply interruption, even if war is avoided. The prospect of war creates the expectation of oil supply cutbacks that would cause higher prices in the future, which leads producers to withdraw some of their oil from current markets (in order to sell it at higher prices later). But once the fear of war recedes, the supply curve of gasoline reverts with equal speed to its earlier position. Given the low short-run price elasticity of demand for gasoline, that's all it takes to generate the considerable price volatility we see in this market.

Price volatility is also common in markets in which demand curves fluctuate sharply and supply curves are highly inelastic. One such market was California's unregulated market for wholesale electricity during the summer of 2000. The supply of electrical generating capacity was essentially fixed in the short run. And because air-conditioning accounts for a large share of demand, several spells of unusually warm weather caused demand to shift sharply to the right. Price at one point reached more than four times its highest level from the previous summer.

Unique and Essential Inputs: The Ultimate Supply Bottleneck

Fans of professional basketball are an enthusiastic bunch. Directly through their purchases of tickets and indirectly through their support of television advertisers, they spend literally billions of dollars each year on the sport. But these dollars are not distributed evenly across all teams. A disproportionate share of all revenues and product endorsement fees accrues to the people associated with consistently winning teams, and at the top of this pyramid generally stands the National Basketball Association's championship team.

Consider the task of trying to produce a championship team in the NBA. You would need inputs such as talented players, a shrewd and dedicated coach and assistants, trainers, physicians, an arena, practice facilities, means for transporting players to away games, a marketing staff, and so on. And whereas some of these inputs can be acquired at reasonable prices in the marketplace, many others cannot. Indeed, the most important input of all—highly talented players—is in extremely limited supply. This is so because the very definition of talented player is inescapably relative—simply put, such a player is one who is better than most others.

Given the huge payoff that accrues to the NBA championship team, it is no surprise that the bidding for the most talented players has become so intense. If there were a long list of players with the potential to boost a team's winning percentage substantially, the Golden State Warriors wouldn't have agreed to pay Steph Curry a salary of more than \$40 million a year. But, of course, the supply of such players is

extremely limited. There are many hungry organizations that would like nothing better than to claim the NBA championship each year, yet no matter how much each is willing to spend, only one can succeed. The supply of NBA championship teams is perfectly inelastic with respect to price even in the very long run.

Sports champions are by no means the only important product whose supply elasticity is constrained by the inability to reproduce unique and essential inputs. In the movie industry, for example, although the supply of movies starring Dwayne Johnson is not perfectly inelastic, there are only so many films he can make each year. Because his films consistently generate huge box office revenues, scores of film producers want to sign him for their projects. But because there isn't enough of him to go around, his salary per film typically exceeds \$20 million.

In the long run, unique and essential inputs are the only truly significant supply bottleneck. If it were not for the inability to duplicate the services of such inputs, most goods and services would have extremely high price elasticities of supply in the long run.

SUMMARY

- The supply curve for a good or service is a schedule that, for any price, tells us the quantity that sellers wish to supply at that price. The prices at which goods and services are offered for sale in the market depend, in turn, on the opportunity cost of the resources required to produce them. (LO1)
- The demand curve facing a perfectly competitive firm is a horizontal line at the price for which industry supply and demand intersect. (LO2)
- Supply curves tend to be upward-sloping, at least in the short run. In general, rational producers will always take advantage of their best opportunities first, moving on to more difficult or costly opportunities only after their best ones have been exhausted. Reinforcing this tendency is the law of diminishing returns, which says that when some factors of production are held fixed, the amount of additional variable factors required to produce successive increments in output grows larger. The industry supply curve is the horizontal summation of the supply curves of individual firms in the industry. (LO2, LO3)
- In a typical production process, firms combine inputs, such as capital and labor, to produce output. The amount of a variable input can be altered in the short run, but the amount of a fixed input can be altered only in the long run. (LO3)
- For perfectly competitive markets—or, more generally, for markets in which individual sellers can sell whatever quantity they wish at a constant price—the seller's best option is to sell that quantity of output for which price equals marginal cost, provided price exceeds the minimum value of average variable cost. The supply curve for the seller thus coincides with the portion of her marginal cost curve, the curve that measures the cost of producing additional units of output. This is why we sometimes say the supply curve represents the cost side of the market (in contrast to the demand curve, which represents the benefit side of the market). (LO3)
- Among the relevant factors causing the supply curves to shift are new technologies, changes in input prices, changes in the number of sellers, expectations of future price changes, and changes in the prices of other products that firms might produce. (LO4)
- Price elasticity of supply is defined as the percentage change in quantity supplied that occurs in response to a 1 percent change in price. The mathematical formula for the price elasticity of supply at any point is $(\Delta Q/Q)/(\Delta P/P)$, where P and Q are the price and quantity at that point, ΔP is a small change in the initial price, and ΔQ is the resulting change in quantity. This formula also can be expressed as $(P/Q) \times (1/\text{slope})$ where $(1/\text{slope})$ is the reciprocal of the slope of the supply curve. (LO5)
- The price elasticity of supply of a good depends on how difficult or costly it is to acquire additional units of the inputs involved in producing that good. In general, the more easily additional units of these inputs can be acquired, the higher price elasticity of supply will be. It is easier to expand production of a product if the inputs used to produce that product are similar to inputs used to produce other products, if inputs are relatively mobile, or if an acceptable substitute for existing inputs can be developed. And like the price elasticity of demand, the price elasticity of supply is greater in the long run than in the short run. (LO5)

KEY TERMS

factor of production	marginal cost	profit-maximizing firm
fixed cost	perfectly competitive market	short run
fixed factor of production	perfectly elastic supply	total cost
imperfectly competitive firm (or price setter)	perfectly inelastic supply	variable cost
law of diminishing returns	price elasticity of supply	variable factor of production
long run	price taker	
	profit	

REVIEW QUESTIONS

1. Explain why you would expect supply curves to slope upward based on increasing opportunity cost. (LO1)
2. Explain why a perfectly competitive firm's demand curve is horizontal. (LO2)
3. Which do you think is more likely to be a fixed factor of production for an ice cream producer during the next two months: its factory building or its workers who operate the machines? Explain. (LO3)
4. True or false: The perfectly competitive firm should *always* produce the output level for which price equals marginal cost. (LO3)
5. Economists often stress that congestion helps account for the law of diminishing returns. With this in mind, explain why it would be impossible to feed all the people on Earth with food grown in a single flowerpot, even if unlimited water, labor, seed, fertilizer, sunlight, and other inputs were available. (LO3, LO4)
6. Why is supply elasticity higher in the long run than in the short run? (LO5)

PROBLEMS



1. Zoe is trying to decide how to divide her time between her job as a wedding photographer, which pays \$27 per hour for as many hours as she chooses to work, and as a fossil collector, in which her pay depends on both the price of fossils and the number of fossils she finds. Earnings aside, Zoe is indifferent between the two tasks, and the number of fossils she can find depends on the number of hours a day she searches, as shown in the table below. (LO1)
2. The supply curves for the only two firms in a competitive industry are given by, respectively, $P = 2Q_1$ and $P = 2 + Q_2$, where Q_1 is the output of firm 1 and Q_2 is the output of firm 2. What is the market supply curve for this industry? (*Hint*: Graph the two curves side by side; then add their respective quantities at a sample of different prices.) (LO2)
3. A price-taking firm makes air conditioners. The market price of one of its new air conditioners is \$120. The firm's total cost information is given in the table below.

Hours per day	Total fossils per day
1	5
2	9
3	12
4	14
5	15

- a. Using the information above, compute the lowest price that Zoe would accept per fossil in order to justify her spending more time collecting fossils and less time working as a wedding photographer.
- b. Plot these points in a graph with price on the vertical axis and quantity per day on the horizontal. What is this curve called?

Air conditioners per day	Total cost (\$ per day)
1	100
2	150
3	220
4	310
5	405
6	510
7	650
8	800

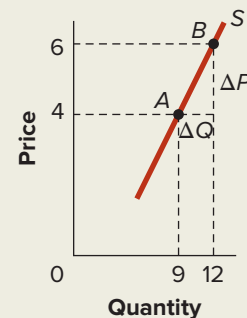
How many air conditioners should the firm produce per day if its goal is to maximize its profit? (LO3)

4. Paducah Slugger Company makes baseball bats out of lumber supplied to it by Acme Sporting Goods, which pays Paducah \$10 for each finished bat. Paducah's only factors of production are lathe operators and a small building with a lathe. The number of bats it produces per day depends on the number of employee-hours per day, as shown in the table below. (LO3, LO4)

Number of bats per day	Number of employee-hours per day
0	0
5	1
10	2
15	4
20	7
25	11
30	16
35	22

- If the wage is \$15 per hour and Paducah's daily fixed cost for the lathe and building is \$60, what is the profit-maximizing quantity of bats?
- What would be the profit-maximizing number of bats if the government imposed a tax of \$10 per day on the company? (*Hint*: Think of this tax as equivalent to a \$10 increase in fixed cost.)
- What would be the profit-maximizing number of bats if the government imposed a tax of \$2 per bat? (*Hint*: Think of this tax as equivalent to a \$2-per-bat increase in marginal cost.)
- Why do the taxes in parts b and c have such different effects?

- How would each of the following affect the U.S. market supply curve for corn? (LO3, LO4)
 - The government taxes sodas sweetened with high-fructose corn syrup.
 - The opportunity cost of farmers' time increases.
 - Scientists discover that corn consumption improves performance on standardized tests.
- The price elasticity of supply for basmati rice (an aromatic strain of rice) is likely to be which of the following? (LO5)
 - Higher in the long run than the short run, because farmers cannot easily change their decisions about how much basmati rice to plant once the current crop has been planted.
 - High, because consumers have a lot of other kinds of rice and other staple foods to choose from.
 - Low in both the long and short runs, because rice farming requires only unskilled labor.
 - High in both the long run and the short run because the inputs required to produce basmati rice can easily be duplicated.
- * What are the respective price elasticities of supply at A and B on the supply curve shown in the accompanying figure? (LO5)



*Indicates more difficult problem.

ANSWERS TO SELF-TESTS

- 4.1 Since Ushi will find 300 containers if he searches a third hour, we find his reservation price for searching a third hour by solving $p(300) = \$6$ for $p = 2$ cents. His reservation prices for additional hours of search are calculated in an analogous way. (LO1)

Fourth hour: $p(200) = \$6$, so $p = 3$ cents.
 Fifth hour: $p(100) = \$6$, so $p = 6$ cents.

- 4.2 The profit figures corresponding to a price of \$5/hundred are as shown in column 5 of the table below, where we see that the profit-maximizing output (which here means the loss-minimizing output) is 0 bottles/day. Note that the company actually loses \$40/day at that output level. But it would lose even more if it produced any other amount. If the company expects conditions to remain unchanged, it will want to go out of the

(1) Q (bottles/day)	(2) Total revenue (\$/day)	(3) Total labor cost (\$/day)	(4) Total cost (\$/day)	(5) Profit (\$/day)
0	0	0	40	-40
100	5	10	50	-45
200	10	20	60	-50
300	15	40	80	-65
400	20	70	110	-90
500	25	110	150	-125
600	30	160	200	-170
700	35	220	260	-225

bottle business as soon as its equipment lease expires. (LO3)

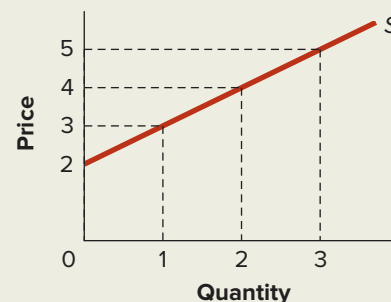
- 4.3 The profit figures corresponding to a price of \$25/hundred are as shown in column 5 of the table below, where we see that the profit-maximizing output (which here means the loss-minimizing output) is 300 bottles/day. Note that the company actually loses \$5/day at that output level. But as

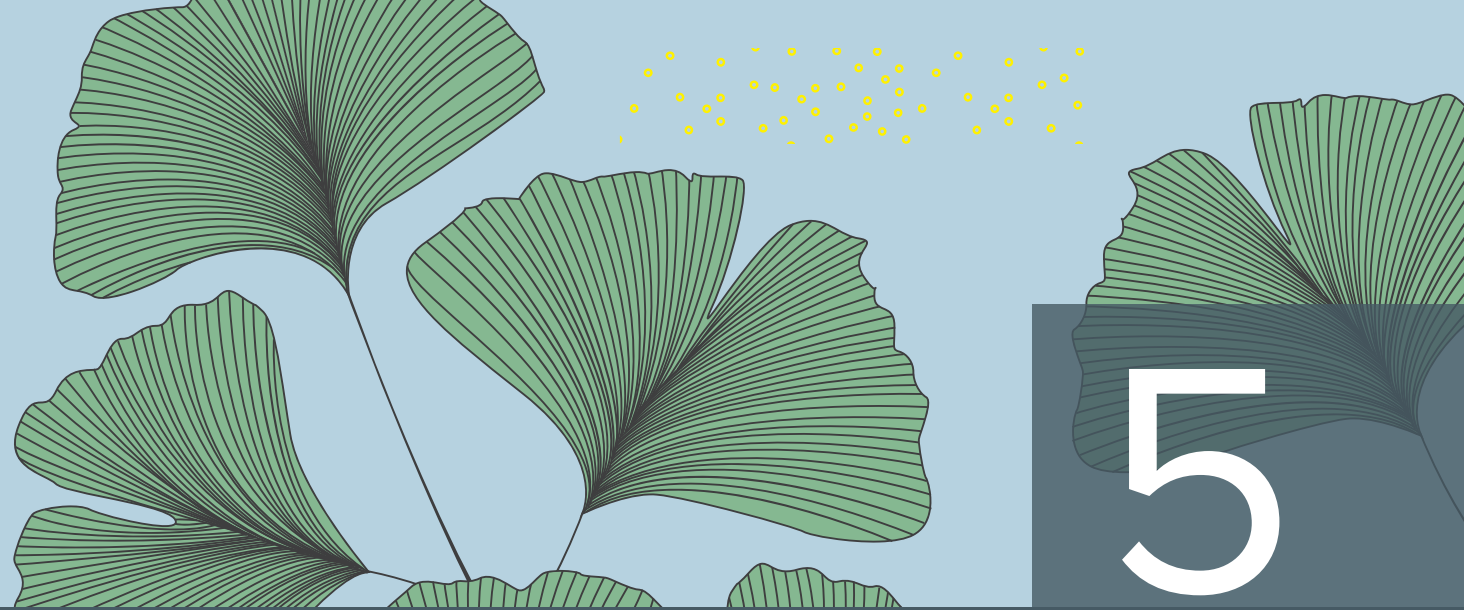
long as it remains committed to its daily lease payment of \$40, it would lose even more if it produced any other amount. If the company expects conditions to remain unchanged, it will want to go out of the bottle business as soon as its equipment lease expires. (LO3)

(1) Q (bottles/day)	(2) Total revenue (\$/day)	(3) Total labor cost (\$/day)	(4) Total cost (\$/day)	(5) Profit (\$/day)
0	0	0	40	-40
100	25	10	50	-25
200	50	20	60	-10
300	75	40	80	-5
400	100	70	110	-10
500	125	110	150	-25
600	150	160	200	-50
700	175	220	260	-85

- 4.4 The fact that each of the city's 60,000 residents is willing to pay 0.00005 cent for each bottle removed means that the collective benefit of each bottle removed is $(60,000)(0.00005) = 3$ cents. So the city should set the redemption price at 3 cents, and from the supply curve we see that 15,000 bottles per day will be recycled at that price. (LO3)

- 4.5 For the supply curve below, $Q = 1$ when $P = 3$, so elasticity of supply = $(P/Q) \times (1/\text{slope}) = (3) \times (1) = 3$. (LO5)





Efficiency, Exchange, and the Invisible Hand in Action

The restaurant market in Ithaca, New York, offered little variety in the 1970s: The city had only one Japanese, two Greek, four Italian, and three Chinese restaurants. Today, some 40 years later and with essentially the same population, Ithaca has one Sri Lankan, three Indian, one French, one Spanish, six Thai, two Korean, two Vietnamese, four Mexican, three Greek, seven Italian, two Caribbean, two Japanese, one Ethiopian, and nine Chinese restaurants. In some of the city's other markets, however, the range of available choices has narrowed. For example, several companies provided telephone answering services in 1972, but only one does so today.

Rare indeed is the marketplace in which the identities of the buyers and sellers remain static for extended periods. New businesses enter; established ones leave. There are more body-piercing studios in Ithaca now and fewer watch-repair shops; more marketing consultants and fewer intercity bus companies; and more appliances in stainless steel or black finishes, fewer in avocado or coppertone.

Driving these changes is the business owner's quest for profit. Businesses migrate to industries and locations in which profit opportunities abound and desert those whose prospects appear bleak. In perhaps the most widely quoted passage from his landmark treatise, *The Wealth of Nations*, Adam Smith wrote,

It is not from the benevolence of the butcher, the brewer, or the baker that we expect our dinner, but from their regard of their own interest. We address ourselves not to their humanity, but to their self-love, and never talk to them of our necessities, but of their advantage.¹

Smith went on to argue that although the entrepreneur "intends only his own gain," he is "led by an invisible hand to promote an end which was no part of his intention." As Smith saw it, even though self-interest is the prime mover of economic activity, the end result is an allocation of goods and services that serves society's collective interests remarkably well. If producers are offering "too much" of

LEARNING OBJECTIVES

After reading this chapter, you should be able to:

- LO1** Define and explain the differences between accounting profit, economic profit, and normal profit.
- LO2** Describe the invisible hand theory and explain why the quest for economic profit drives resource allocation and motivates firms to enter some industries and leave others.
- LO3** Explain why economic profit, unlike economic rent, tends toward zero in the long run.
- LO4** Identify conditions under which a market equilibrium may not be socially efficient and explain why no opportunities for gain remain open to individuals when a market is in equilibrium.
- LO5** Explain how resource allocation can be disrupted by policies that prevent markets from reaching equilibrium.

¹Adam Smith, *The Wealth of Nations* (New York: Everyman's Library, Book 1, 1910).



Why do most American cities now have more tattoo parlors and fewer watch-repair shops than in 1972?

one product and “not enough” of another, profit opportunities stimulate entrepreneurs into action. All the while, the system exerts relentless pressure on producers to hold the price of each good close to its cost of production, and indeed to reduce that cost in any ways possible. The invisible hand, in short, is about all the good things that can happen when people pursue their own interests in the marketplace.

Our task in this chapter is to gain deeper insight into the nature of the forces that guide the invisible hand. What exactly does “profit” mean? How is it measured, and how does the quest for it serve society’s ends? And if competition holds price close to the cost of production, why do so many entrepreneurs become fabulously wealthy?

We will also discuss cases in which misunder-

standing of Smith’s theory results in costly errors, both in everyday decision making and in the realm of government policy.

THE CENTRAL ROLE OF ECONOMIC PROFIT

The economic theory of business behavior is built on the assumption that the firm’s goal is to maximize its profit. So we must be clear at the outset about what, exactly, profit means.

Three Types of Profit

The economist’s understanding of profit is different from the accountant’s, and the distinction between the two is important for understanding how the invisible hand works. Accountants define the annual profit of a business as the difference between the revenue it takes in and its **explicit costs** for the year, which are the actual payments the firm makes to its factors of production and other suppliers. Profit thus defined is called **accounting profit**:

$$\text{Accounting profit} = \text{Total revenue} - \text{Explicit costs.}$$

Accounting profit is the most familiar profit concept in everyday discourse. It is the one that companies use, for example, when they provide statements about their profits in press releases or annual reports.²

Economists, by contrast, define profit as the difference between the firm’s total revenue and not just its explicit costs, but also its **implicit costs**, which are the opportunity costs of all the resources supplied by the firm’s owners. Profit thus defined is called **economic profit** (or **excess profit**):

$$\text{Economic profit} = \text{Total revenue} - \text{Explicit costs} - \text{Implicit costs.}$$

To illustrate the difference between accounting profit and economic profit, consider a firm with \$400,000 in total annual revenue whose only explicit costs are workers’ salaries, totaling \$250,000 per year. The owners of this firm have supplied machines and other capital equipment with a total resale value of \$1 million. This firm’s accounting profit then is \$150,000, or the difference between its total revenue of \$400,000 per year and its explicit costs of \$250,000 per year.

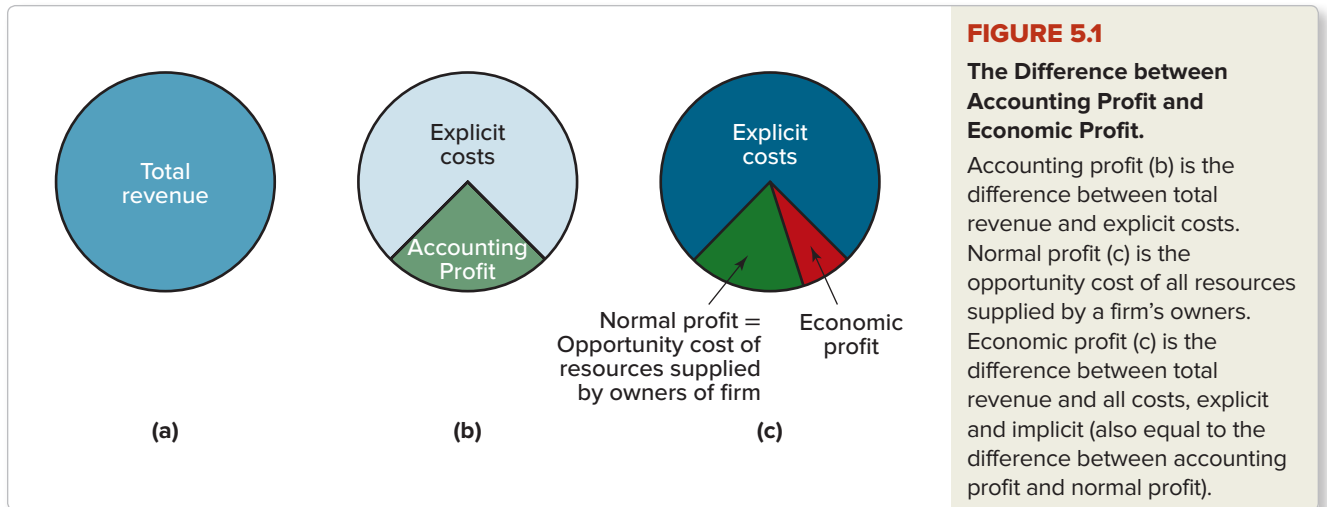
²For simplicity, this discussion ignores any costs associated with depreciation of the firm’s capital equipment. Because the buildings and machines owned by a firm tend to wear out over time, the government allows the firm to consider a fraction of their value each year as a current cost of doing business. For example, a firm that employs a \$1,000 machine with a 10-year life span might be allowed to record \$100 as a current cost of doing business each year.

explicit costs the actual payments a firm makes to its factors of production and other suppliers

accounting profit the difference between a firm’s total revenue and its explicit costs

implicit costs the opportunity costs of the resources supplied by the firm’s owners

economic profit (or **excess profit**) the difference between a firm’s total revenue and the sum of its explicit and implicit costs



To calculate the firm's economic profit, sometimes called its excess profit, we must first calculate the opportunity cost of the resources supplied by the firm's owners. Suppose the current annual interest rate on savings accounts is 10 percent. Had owners not invested in capital equipment, they could have earned an additional \$100,000 per year interest by depositing their \$1 million in a savings account. So the firm's economic profit is \$400,000 per year – \$250,000 per year – \$100,000 per year = \$50,000 per year.

Note that this economic profit is smaller than the accounting profit by exactly the amount of the firm's implicit costs—the \$100,000 per year opportunity cost of the resources supplied by the firm's owners. This difference between a business's accounting profit and its economic profit is called its **normal profit**. Normal profit is simply the opportunity cost of the resources supplied to a business by its owners.

Figure 5.1 illustrates the difference between accounting and economic profit. A firm's total revenue is represented in Figure 5.1(a), while (b) and (c) show how this revenue is apportioned among the various cost and profit categories.

The following examples illustrate why the distinction between accounting and economic profit is so important.

normal profit the opportunity cost of the resources supplied by the firm's owners, equal to accounting profit minus economic profit

EXAMPLE 5.1 Accounting versus Economic Profit, Part 1

Should Pudge Buffet stay in the farming business?

Pudge Buffet is a corn farmer who lives near Lincoln, Nebraska. His payments for land and equipment rental and for other supplies come to \$10,000 per year. The only input he supplies is his own labor, and he considers farming just as attractive as his only other employment opportunity, managing a retail store at a salary of \$11,000 per year. Apart from the matter of pay, Pudge is indifferent between farming and being a manager. Corn sells for a constant price per bushel in an international market too large to be affected by changes in one farmer's corn production. Pudge's revenue from corn sales is \$22,000 per year. What is his accounting profit? His economic profit? His normal profit? Should he remain a corn farmer?

As shown in Table 5.1, Pudge's accounting profit is \$12,000 per year, the difference between his \$22,000 annual revenue and his \$10,000 yearly payment for land, equipment, and supplies. His economic profit is that amount less the opportunity cost of his labor. Since the latter is the \$11,000 per year he could have earned as a store manager, he is making an economic profit of \$1,000 per year.

Finally, his normal profit is the \$11,000 opportunity cost of the only resource he supplies, namely, his labor. Since Pudge likes the two jobs equally well, he will be better off by \$1,000 per year if he remains in farming.

TABLE 5.1
Revenue, Cost, and Profit Summary

Total revenue (\$/year)	Explicit costs (\$/year)	Implicit costs (\$/year)	Accounting profit (= total revenue – explicit costs) (\$/year)	Economic profit (= total revenue – explicit costs – implicit costs) (\$/year)	Normal profit (= implicit costs) (\$/year)
22,000	10,000	11,000	12,000	1,000	11,000

SELF-TEST 5.1

In Example 5.1, how will Pudge's economic profit change if his annual revenue from corn production is not \$22,000, but \$20,000? Should he continue to farm?

economic loss an economic profit that is less than zero

When revenue falls from \$22,000 to \$20,000, Pudge has an economic profit of $-\$1,000$ per year. A negative economic profit is also called an **economic loss**. If Pudge expects to sustain an economic loss indefinitely, his best bet would be to abandon farming in favor of managing a retail store.

You might think that if Pudge could just save enough money to buy his own land and equipment, his best option would be to remain a farmer. But as the following example illustrates, that impression is based on a failure to perceive the difference between accounting profit and economic profit.

EXAMPLE 5.2 Accounting versus Economic Profit, Part 2

Does owning one's own land make a difference?

Let's build on Example 5.1 and Self-Test 5.1. Suppose Pudge's Uncle Warren, who owns the farmland Pudge has been renting, dies and leaves Pudge that parcel of land. If the land could be rented to some other farmer for \$6,000 per year, should Pudge remain in farming?

As shown in Table 5.2, if Pudge continues to farm his own land, his accounting profit will be \$16,000 per year, or \$6,000 more than in Self-Test 5.1. But his economic profit will still be the same as before—that is, $-\$1,000$ per year—because

TABLE 5.2
Revenue, Cost, and Profit Summary

Total revenue (\$/year)	Explicit costs (\$/year)	Implicit costs (\$/year)	Accounting profit (= total revenue – explicit costs) (\$/year)	Economic profit (= total revenue – explicit costs – implicit costs) (\$/year)	Normal profit (= implicit costs) (\$/year)
20,000	4,000	17,000	16,000	$-1,000$	17,000

Pudge must deduct the \$6,000 per year opportunity cost of farming his own land, even though he no longer must make an explicit payment to his uncle for it. The normal profit from owning and operating his farm will be \$17,000 per year—the opportunity cost of the land and labor he provides. But since Pudge earns an accounting profit of only \$16,000, he will again do better to abandon farming for the managerial job.

Pudge obviously would be wealthier as an owner than he was as a renter. But the question of whether to remain a farmer is answered the same way whether Pudge rents his farmland or owns it. He should stay in farming only if that is the option that yields the highest economic profit.

RECAP ↑

THE CENTRAL ROLE OF ECONOMIC PROFIT

A firm's accounting profit is the difference between its revenue and the sum of all explicit costs it incurs. Economic profit is the difference between the firm's revenue and *all* costs it incurs—both explicit and implicit. Normal profit is the opportunity cost of the resources supplied by the owners of the firm. When a firm's accounting profit is exactly equal to the opportunity cost of the inputs supplied by the firm's owners, the firm's economic profit is zero. For a firm to remain in business in the long run, it must earn an economic profit greater than or equal to zero.

THE INVISIBLE HAND THEORY

Two Functions of Price

In the free enterprise system, market prices serve two important and distinct functions. The first, the **rationing function of price**, is to distribute scarce goods among potential claimants, ensuring that those who get them are the ones who value them most. Thus, if three people want the only antique clock for sale at an auction, the clock goes home with the person who bids the most for it. The second function, the **allocative function of price**, is to direct productive resources to different sectors of the economy. Resources leave markets in which price cannot cover the cost of production and enter those in which price exceeds the cost of production.

Both the allocative and rationing functions of price underlie Adam Smith's celebrated **invisible hand theory** of the market. Recall that Smith thought the market system channeled the selfish interests of individual buyers and sellers so as to promote the greatest good for society. The carrot of economic profit and the stick of economic loss, he argued, were often the only forces necessary to ensure that existing supplies in any market would be allocated efficiently and that resources would be allocated across markets to produce the most efficient possible mix of goods and services.

Responses to Profits and Losses

To get a feel for how the invisible hand works, we begin by looking at how firms respond to economic profits and losses. If a firm is to remain in business in the long run, it must cover all its costs, both explicit and implicit. A firm's normal profit is just a cost of doing business. Thus, the owner of a firm that earns no more than a normal profit has managed only to recover the opportunity cost of the resources invested in the firm. By contrast, the owner of a firm that makes a positive economic profit earns more than the opportunity cost of the invested resources; she earns a normal profit and then some.

rationing function of price

changes in prices distribute scarce goods to those consumers who value them most highly

allocative function of price

changes in prices direct resources away from overcrowded markets and toward markets that are underserved

invisible hand theory

Adam Smith's theory that the actions of independent, self-interested buyers and sellers will often result in the most efficient allocation of resources

Naturally, everyone would be delighted to earn more than a normal profit, and no one wants to earn less. The result is that those markets in which firms are earning an economic profit tend to attract additional resources, whereas markets in which firms are experiencing economic losses tend to lose resources.

The following example examines how the forces of the invisible hand would respond if not just Pudge Buffet but also all other farmers in Lincoln, Nebraska, were experiencing economic losses.

EXAMPLE 5.3 The Invisible Hand Theory in Action

What would happen if all farmers in Lincoln earned less than a normal profit?

Suppose the conditions confronting Pudge Buffet in the example summarized in Table 5.2 are essentially the same as those confronting all other farmers in Lincoln, Nebraska—that is, all earn less than a normal profit. What economic changes will result?

If all farmers in Lincoln are earning a negative economic profit, some farmers will begin switching to other activities. As they abandon farming, however, the market price for farmland—and hence its opportunity cost—will begin to fall. It will continue to fall until farmers in Lincoln can once again earn a normal profit. Specifically, the price of land will fall until the yearly rental for a farm like Pudge’s is only \$5,000, for at that rent the accounting profit of someone who farmed his own land would be \$16,000 per year, exactly the same as his normal profit. His economic profit would be zero.

EXAMPLE 5.4 Incentive to Change Behavior

What would happen if all farmers earned more than a normal profit?

Suppose corn growers farm 80 acres of their own land, which sells for \$1,000 per acre. Each farm’s revenue from corn sales is \$20,000 per year. Equipment and other supplies cost \$4,000 per year, and the current annual interest rate on savings accounts is 5 percent. Farmers can earn \$11,000 per year in alternative jobs that they like equally well as farming. What is normal economic profit for these farmers? How much accounting profit will they earn? How much economic profit? Is their economic situation stable? If not, how is it likely to change?

As shown in Table 5.3, accounting profit—the difference between the \$20,000 annual revenue and the \$4,000 annual expense for equipment and supplies—is \$16,000 per year, as in the example just discussed. Normal profit is the opportunity cost of the farmer’s time and land—\$11,000 for his time and \$4,000 for his

TABLE 5.3
Revenue, Cost, and Profit Summary

Total revenue (\$/year)	Explicit costs (\$/year)	Implicit costs (\$/year)	Accounting profit (= total revenue – explicit costs) (\$/year)	Economic profit (= total revenue – explicit costs – implicit costs) (\$/year)	Normal profit (= implicit costs) (\$/year)
20,000	4,000	15,000	16,000	1,000	15,000

land (since had he sold the land for \$80,000 and put the money in the bank at 5 percent interest, he would have earned \$4,000 per year in interest)—for a total of \$15,000. Accounting profit thus exceeds normal profit by \$1,000 per year, which means that farmers are earning an economic profit of \$1,000 per year.

To see whether this situation is stable, we must ask whether people have an incentive to change their behavior. Consider the situation from the perspective of a manager who is earning \$11,000 per year. To switch to farming, he would need to borrow \$80,000 to buy land, which would mean interest payments of \$4,000 per year. With \$20,000 per year in revenue from corn sales and \$4,000 per year in expenses for supplies and equipment, in addition to \$4,000 per year in interest payments, the manager would earn an accounting profit of \$12,000 per year. And since that amount is \$1,000 per year more than the opportunity cost of the manager's time, he will want to switch to farming. Indeed, *all* managers will want to switch to farming. At current land prices, there is cash on the table in farming.

Yet any situation with cash on the table is inherently unstable. There is only so much farmland to go around, so as demand for it increases, its price will begin to rise. The price will keep rising until there is no longer any incentive for managers to switch to farming.

How much must the price of land rise to eliminate the incentive to switch? If 80 acres of land sold for \$100,000 (that is, if land sold for \$1,250 per acre), the interest on the money borrowed to buy a farm would be \$5,000 per year, an amount that would make workers indifferent between farming and being a manager. But if land sells for anything less than \$1,250 per acre, there will be excess demand for farmland.

The Effect of Market Forces on Economic Profit

A firm's normal profit is just like any other cost of doing business. Thus the owner of a firm that earns no more than a normal profit has managed only to recover the opportunity cost of the resources invested in the firm. By contrast, the owner of a firm that makes a positive economic profit earns more than the opportunity cost of the invested resources; she earns a normal profit and then some. Naturally, everyone would be delighted to earn more than a normal profit, and no one wants to earn less. The result is that those markets in which firms are earning an economic profit tend to attract additional resources, whereas markets in which firms are experiencing economic losses tend to lose resources.

In Example 5.4, we assumed that the price of corn was set in a world market too large to be influenced by the amount of corn produced in any one locality. More generally, however, we need to consider the effects of supply shifts on price.

Consider first the effect of an influx of resources in a market in which firms are currently earning an economic profit. As new firms enter the market, the supply curve will shift to the right, causing a reduction in the price of the product (see Figure 5.2).

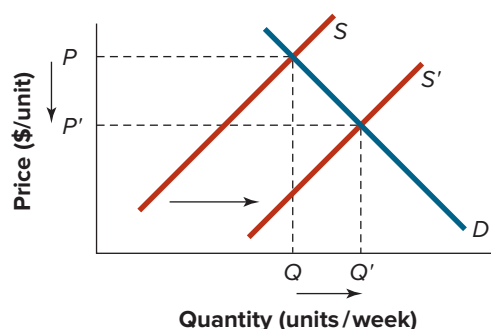


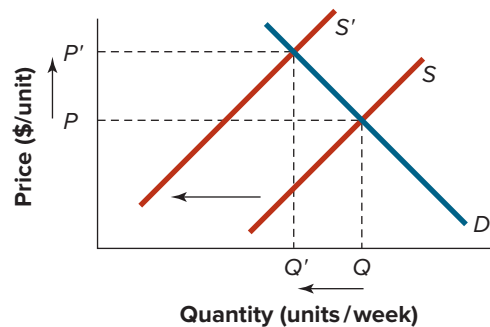
FIGURE 5.2

The Effect of Positive Economic Profit on Entry.

A market in which firms earn a positive economic profit will attract new firms from other markets. The resulting increase in supply will lead to a reduction in market price.

FIGURE 5.3**The Effect of Economic Losses on Market Exit.**

Firms tend to leave a market when they experience an economic loss. The result is a leftward shift in the supply curve and a corresponding increase in price. Firms will continue to leave the market until the price rises enough to cover all costs, including the opportunity cost of resources supplied by a firm's owners.



If firms continue to earn a positive economic profit at the new, lower price, P' , additional firms will enter, causing the market price to fall still further. The process will continue until economic profit is driven down to zero—that is, until price is just sufficient to cover all costs, including a normal profit.

Now consider the effect of resources moving out of a market in which businesses are currently experiencing an economic loss. As firms leave, the market supply curve shifts to the left, causing the price of the product to rise, as shown in Figure 5.3. Firms will continue to exit until price rises to cover all resource costs—including the opportunity cost of the resources that owners have invested in their firms. The economic loss firms have been sustaining will be eliminated.

The net result of these resource movements is that in the long run all firms will tend to earn zero economic profit. Their *goal* is not to earn zero profit. Rather, the zero-profit tendency is a consequence of the dynamics of their entry into and exit from the market. That's because when people confront an opportunity for gain, they are almost always quick to exploit it.

The Importance of Free Entry and Exit

The allocative function of price cannot operate unless firms can enter new markets and leave existing ones at will. If new firms could not enter a market in which existing firms were making a large economic profit, economic profit would not tend to fall to zero over time, and price would not tend to gravitate toward the marginal cost of production.

Forces that inhibit firms from entering new markets are called **barriers to entry**. In the book publishing market, for example, the publisher of a book enjoys copyright protection granted by the government. Copyright law forbids other publishers from producing and selling their own editions of protected works. This barrier allows the price of a popular book to remain significantly above its cost of production for an extended period, all the while generating an economic profit for its publisher. (A copyright provides no *guarantee* of a profit, and indeed most new books actually generate an economic loss for their publishers.)

Barriers to entry may result from practical as well as legal constraints. Some economists, for example, have argued that the compelling advantages of product compatibility have created barriers to entry in the computer software market. Since roughly 80 percent of new desktop computers come with Microsoft's Windows operating system already installed, rival companies have difficulty selling other operating systems that may limit users' ability to exchange files with friends and colleagues. This fact, more than any other, explains Microsoft's spectacular profit history. But the share of computers equipped with Windows has been steadily falling, leading many economists to predict that Microsoft's profits will follow a similar trajectory.

barrier to entry any force that prevents firms from entering a new market

No less important than the freedom to enter a market is the freedom to leave. When the airline industry was regulated by the federal government, air carriers were often required to serve specific markets, even though they were losing money in them. When firms discover that a market, once entered, is difficult or impossible to leave, they become reluctant to enter new markets. Barriers to exit thus become barriers to entry. Without reasonably free entry and exit, then, the implications of Adam Smith's invisible hand theory cannot be expected to hold.

All things considered, producers enjoy a high degree of freedom of entry in most U.S. markets. Because free entry is one of the defining characteristics of perfectly competitive markets, unless otherwise stated, we'll assume its existence.

The Invisible Hand in Action

To help develop your intuition about how the invisible hand works, we will examine how it helps us gain insight into patterns we observe in a wide variety of different contexts. In each case, the key idea we want you to focus on is that opportunities for private gain seldom remain unexploited for very long. Perhaps more than any other, this idea encapsulates the essence of that distinctive mindset known as "thinking like an economist."

The Invisible Hand at the Supermarket and on the Freeway

As the following example illustrates, it is not just opportunities to earn economic profits that are likely to disappear quickly, but also any other opportunities to achieve unusually desirable outcomes.

The Economic Naturalist 5.1



Why do supermarket checkout lines all tend to be roughly the same length?

Pay careful attention the next few times you go grocery shopping and you'll notice that the lines at all the checkout stations tend to be roughly the same length. Suppose you saw one line that was significantly shorter than the others as you wheeled your cart toward the checkout area. Which line would you choose? The shorter one, of course; because most shoppers would do the same, the short line seldom remains shorter for long.

McGraw Hill **connect**

► Visit your instructor's Connect course and access your eBook to view this video.



Why do supermarket checkout lines all tend to be roughly the same length?

SELF-TEST 5.2

Use the same reasoning employed in The Economic Naturalist 5.1 to explain why all lanes on a crowded, multilane freeway move at about the same speed.

The Invisible Hand and Cost-Saving Innovations

When economists speak of perfectly competitive firms, they have in mind businesses whose contribution to total market output is too small to have a perceptible impact on market price. As explained in Chapter 4, *Perfectly Competitive Supply*, such firms are often called price takers: they take the market price of their product as given and then produce that quantity of output for which marginal cost equals that price.

This characterization of the competitive firm gives the impression that the firm is essentially a passive actor in the marketplace. Yet for most firms, that is anything

but the case. As the next example illustrates, even those firms that cannot hope to influence the market prices of their products have very powerful incentives to develop and introduce cost-saving innovations.

EXAMPLE 5.5 The Impact of Cost-Saving Innovations on Economic Profit

How do cost-saving innovations affect economic profit in the short run? In the long run?

Forty merchant marine companies operate supertankers that carry oil from the Middle East to the United States. The cost per trip, including a normal profit, is \$500,000. An engineer at one of these companies develops a more efficient propeller design that results in fuel savings of \$20,000 per trip. How will this innovation affect the company's accounting and economic profits? Will these changes persist in the long run?

In the short run, the reduction in a single firm's costs will have no impact on the market price of transoceanic shipping services. The firm with the more efficient propeller will thus earn an economic profit of \$20,000 per trip (since its total revenue will be the same as before, while its total cost will now be \$20,000 per trip lower). As other firms learn about the new design, however, they will begin to adopt it, causing their individual supply curves to shift downward (since the marginal cost per trip at these firms will drop by \$20,000). The shift in these individual supply curves will cause the market supply curve to shift, which in turn will result in a lower market price for shipping and a decline in economic profit at the firm where the innovation originated. When all firms have adopted the new, efficient design, the long-run supply curve for the industry will have shifted downward by \$20,000 per trip and each company will again be earning only a normal profit. At that point, any firm that did *not* adopt the new propeller design would suffer an economic loss of \$20,000 per trip.

The incentive to come up with cost-saving innovations in order to reap economic profit is one of the most powerful forces on the economic landscape. Its beauty, in terms of the invisible hand theory, is that competition among firms ensures that the resulting cost savings will be passed along to consumers in the long run.

RECAP

THE INVISIBLE HAND THEORY

In market economies, the allocative and rationing functions of prices guide resources to their most highly valued uses. Prices influence how much of each type of good gets produced (the allocative function). Firms enter industries in which prices are sufficiently high to sustain an economic profit and leave those in which low prices result in an economic loss. Prices also direct existing supplies of goods to the buyers who value them most (the rationing function).

Industries in which firms earn a positive economic profit tend to attract new firms, shifting industry supply to the right. Firms tend to leave industries in which they sustain an economic loss, shifting supply curves to the left. In each case, the supply movements continue until economic profit reaches zero. In long-run equilibrium, the value of the last unit produced to buyers is equal to its marginal cost of production, leaving no possibility for additional mutually beneficial transactions.

ECONOMIC RENT VERSUS ECONOMIC PROFIT

Microsoft cofounder Bill Gates is one of the wealthiest people on the planet, largely because the problem of compatibility prevents rival suppliers from competing effectively in the many software markets dominated by his company. Yet numerous people have become fabulously rich even in markets with no conspicuous barriers to entry. If market forces push economic profit toward zero, how can that happen?

The answer to this question hinges on the distinction between economic profit and **economic rent**. Most people think of rent as the payment they make to a landlord or the supplier of a dorm refrigerator, but the term *economic rent* has a different meaning. Economic rent is that portion of the payment for an input that is above the supplier's reservation price for that input. Suppose, for example, that a landowner's reservation price for an acre of land is \$100 per year. That is, suppose he would be willing to lease it to a farmer as long as he received an annual payment of at least \$100, but for less than that amount he would rather leave it fallow. If a farmer gives him an annual payment not of \$100 but of \$1,000, the landowner's economic rent from that payment will be \$900 per year.

Economic profit is like economic rent in that it, too, may be seen as the difference between what someone is paid (the business owner's total revenue) and her reservation price for remaining in business (the sum of all her costs, explicit and implicit). But whereas competition pushes economic profit toward zero, it has no such effect on the economic rent for inputs that cannot be replicated easily. For example, although the lease payments for land may remain substantially above the landowner's reservation price, year in and year out, new land cannot come onto the market to reduce or eliminate the economic rent through competition. There is, after all, only so much land to be had.

As the following example illustrates, economic rent can accrue to people as well as land.

economic rent that part of the payment for a factor of production that exceeds the owner's reservation price, the price below which the owner would not supply the factor

EXAMPLE 5.6 Economic Rent

How much economic rent will a talented chef get?

A community has 100 restaurants, 99 of which employ chefs of normal ability at a salary of \$30,000 per year, the same as the amount they could earn in other occupations that are equally attractive to them. But the 100th restaurant has an unusually talented chef. Because of her reputation, diners are willing to pay 50 percent more for the meals she cooks than for those prepared by ordinary chefs. Owners of the 99 restaurants with ordinary chefs each collect \$300,000 per year in revenue, which is just enough to ensure that each earns exactly a normal profit. If the talented chef's opportunities outside the restaurant industry are the same as those of ordinary chefs, how much will she be paid by her employer at equilibrium? How much of her pay will be economic rent? How much economic profit will her employer earn?

Because diners are willing to pay 50 percent more for meals cooked by the talented chef, the owner who hires her will take in total receipts not of \$300,000 per year but of \$450,000. In the long run, competition should ensure that the talented chef's total pay each year will be \$180,000 per year, the sum of the \$30,000 that ordinary chefs get and the \$150,000 in extra revenues for which she is solely responsible. Since the talented chef's reservation price is the amount she could earn outside the restaurant industry—by assumption, \$30,000 per year, the same as for ordinary chefs—her economic rent is \$150,000 per year. The economic profit of the owner who hires her will be exactly zero.

Since the talented chef's opportunities outside the restaurant industry are no better than an ordinary chef's, why is it necessary to pay the talented chef so much? Suppose her employer were to pay her only \$60,000, which they both would consider a generous salary since it is twice what ordinary chefs earn. The employer would then earn an economic profit of \$120,000 per year since his annual revenue would be \$150,000 more than that of ordinary restaurants, but his costs would be only \$30,000 more.

But this economic profit would create an opportunity for the owner of some other restaurant to bid the talented chef away. For example, if the owner of a competing restaurant were to hire the talented chef at a salary of \$70,000, the chef would be \$10,000 per year better off and the rival owner would earn an economic profit of \$110,000 per year, rather than his current economic profit of zero. Furthermore, if the talented chef is the sole reason that a restaurant earns a positive economic profit, the bidding for that chef should continue as long as any economic profit remains. Some other owner will pay her \$80,000, still another \$90,000, and so on. Equilibrium will be reached only when the talented chef's salary has been bid up to the point that no further economic profit remains—in our example, at an annual paycheck of \$180,000.

This bidding process assumes, of course, that the reason for the chef's superior performance is that she possesses some personal talent that cannot be copied. If instead it were the result of, say, training at a culinary institute in France, then her privileged position would erode over time, as other chefs sought similar training.

RECAP

ECONOMIC RENT VERSUS ECONOMIC PROFIT

Economic rent is the amount by which the payment to a factor of production exceeds the supplier's reservation price. Unlike economic profit, which is driven toward zero by competition, economic rent may persist for extended periods, especially in the case of factors with special talents that cannot easily be duplicated.

THE DISTINCTION BETWEEN AN EQUILIBRIUM AND A SOCIAL OPTIMUM

When a market reaches equilibrium, no further opportunities for gain are available to individuals. One implication is that the market prices of resources that people own will eventually reflect their economic value. (As we will see in later chapters, the same cannot be said of resources that are not owned by anyone, such as fish in international waters.)

Note carefully, however, that these observations don't imply that there are *never* any valuable opportunities to exploit. For example, the story is told of two economists on their way to lunch when they spot what appears to be a \$100 bill lying on the sidewalk. When the younger economist stoops to pick up the bill, his older colleague restrains him, saying, "That can't be a \$100 bill." "Why not?" asks the younger colleague. "If it were, someone would have picked it up by now," the older economist replies.

The tendency for people to be quick to exploit profit opportunities doesn't mean that there *never* are any unexploited opportunities, but that there are none when the market is *in equilibrium*. Occasionally a \$100 bill does lie on the sidewalk, and the person who first spots it and picks it up gains a windfall. Likewise, when a company's earnings prospects improve, *somebody* must be the first to recognize the opportunity, and that person can make a lot of money by purchasing the stock quickly.

Still, the observation that good opportunities tend to be snatched up quickly is important. It tells us, in effect, that there are only three ways to earn a big payoff: to

work especially hard; to have some unusual skill, talent, or training; or simply to be lucky. The person who finds a big bill on the sidewalk is lucky, as are many of the investors whose stocks perform better than average. Other investors whose stocks do well achieve their gains through hard work or special talent. For example, the legendary investor Warren Buffett, whose portfolio has grown in value at almost three times the stock market average for the last 60 years, spends long hours studying annual financial reports and has a remarkably keen eye for the telling detail. Thousands of others work just as hard yet fail to beat the market averages.

It is important to stress, however, that a market being in equilibrium implies only that no additional opportunities are available *to individuals*. It does not imply that the resulting allocation is necessarily best from the point of view of society as a whole.

Smart for One, Dumb for All

Adam Smith's profound insight was that the individual pursuit of self-interest often promotes the broader interests of society. But unlike some of his modern disciples, Smith was under no illusion that this is *always* the case. Note, for example, Smith's elaboration on his description of the entrepreneur led by the invisible hand "to promote an end which was no part of his intention":

Nor is it *always* the worse for society that it was no part of it. By pursuing his own interest he *frequently* promotes that of society more effectually than when he really intends to promote it. [Emphasis added.]³

Smith was well aware that the individual pursuit of self-interest often does not coincide with society's interest. In Chapter 2, *Supply and Demand*, we cited activities that generate environmental pollution as an example of conflicting economic interests, noting that behavior in those circumstances may be described as smart for one but dumb for all. As the following example suggests, extremely high levels of investment in earnings forecasts also can be smart for one, dumb for all.

The Economic Naturalist 5.2



Are there "too many" smart people working as corporate earnings forecasters?

Stock analysts use complex mathematical models to forecast corporate earnings. The more analysts invest in the development of these models, the more accurate the models become. Thus, the analyst whose model produces a reliable forecast sooner than others can reap a windfall by buying stocks whose prices are about to rise. Given the speed with which stock prices respond to new information, however, the results of even the second-fastest forecasting model may come too late to be of much use. Individual stock analysts thus face a powerful incentive to invest more and more money in their models, in the hope of generating the fastest forecast. Does this incentive result in the socially optimal level of investment in forecast models?

Beyond some point, increased speed of forecasting is of little benefit to society as a whole, whose interests suffer little when the price of a stock moves to its proper level a few hours more slowly. If *all* stock analysts spent less money on their forecasting models, *someone's* model would still produce the winning forecast, and the resources that might otherwise be devoted to fine-tuning the models could be put to more valued uses. Yet if any one individual spends less, he can be sure the winning forecast will not be his.

The invisible hand went awry in the situation just described because the benefit of an investment to the individual who made it was larger than the benefit of that

³Adam Smith, *The Wealth of Nations* (New York: Everyman's Library, Book 1, 1910).

investment to society as a whole. In later chapters we will discuss a broad class of investments with this property. In general, the efficacy of the invisible hand depends on the extent to which the individual costs and benefits of actions taken in the marketplace coincide with the respective costs and benefits of those actions to society. These exceptions notwithstanding, some of the most powerful forces at work in competitive markets clearly promote society's interests.

Market Equilibrium and Efficiency

Private markets cannot by themselves guarantee an income distribution that most people regard as fair. Nor can they ensure clean air, uncongested highways, or safe neighborhoods for all.

In virtually all successful societies, markets are supplemented by active political coordination in at least some instances. We will almost always achieve our goals more effectively if we know what tasks private markets can do well, and then allow them to perform those tasks. Unfortunately, the discovery that markets cannot solve *every* problem seems to have led some critics to conclude that markets cannot solve *any* problems. This misperception is a dangerous one because it has prompted attempts to prevent markets from doing even those tasks for which they are ideally suited.

We will explore why many tasks are best left to the market and the conditions under which unregulated markets generate the largest possible economic surplus. We also will discuss why attempts to interfere with market outcomes often lead to unintended and undesired consequences.

As noted earlier, the mere fact that markets coordinate the production of a large and complex list of goods and services is reason enough to marvel at them. But in the absence of pollution and other externalities like the ones discussed in the preceding section, economists make an even stronger claim—namely, that markets not only produce these goods, but also produce them as efficiently as possible.

The term **efficient**, as economists use it, has a narrow technical meaning. When we say that market equilibrium is efficient, we mean simply this: *If price and quantity take anything other than their equilibrium values, a transaction that will make at least some people better off without harming others can always be found.* This conception of efficiency is also known as **Pareto efficiency**, after Vilfredo Pareto, the nineteenth-century Italian economist who introduced it.

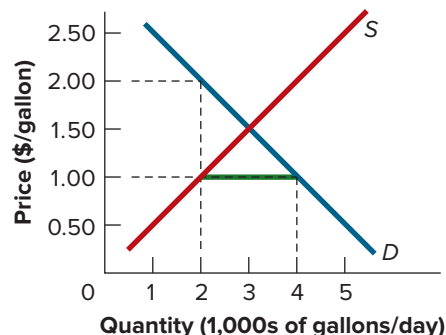
Why is market equilibrium efficient in this sense? The answer is that it is always possible to construct an exchange that helps some without harming others whenever a market is out of equilibrium. Suppose, for example, that the supply and demand curves for milk are as shown in Figure 5.4 and that the current price of milk is \$1 per gallon. At that price, sellers offer only 2,000 gallons of milk a day. At that quantity, the marginal buyer values an extra gallon of milk at \$2. This is the price that corresponds to 2,000 gallons a day on the demand curve, which represents what the marginal buyer is willing to pay for an additional gallon (another application of the vertical interpretation of the demand curve). We also know that the cost of producing an extra gallon of milk

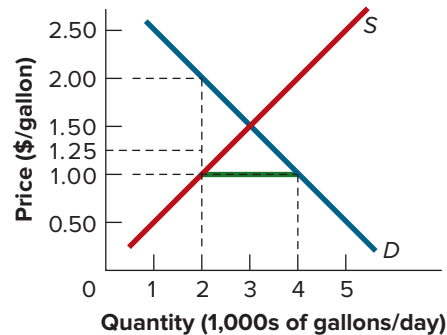
efficient (or Pareto efficient) a situation is efficient if no change is possible that will help some people without harming others

FIGURE 5.4

A Market in Which Price Is Below the Equilibrium Level.

In this market, milk is currently selling for \$1 per gallon, \$0.50 below the equilibrium price of \$1.50 per gallon.



**FIGURE 5.5****How Excess Demand Creates an Opportunity for a Surplus-Enhancing Transaction.**

At a market price of \$1 per gallon, the most intensely dissatisfied buyer is willing to pay \$2 for an additional gallon, which a seller can produce at a cost of only \$1. If this buyer pays the seller \$1.25 for the extra gallon, the buyer gains an economic surplus of \$0.75 and the seller gains an economic surplus of \$0.25.

is only \$1. This is the price that corresponds to 2,000 gallons a day on the supply curve, which equals marginal cost (another application of the vertical interpretation of the supply curve).

Furthermore, a price of \$1 per gallon leads to excess demand of 2,000 gallons per day, which means that many frustrated buyers cannot buy as much milk as they want at the going price. Now suppose a supplier sells an extra gallon of milk to the most eager of these buyers for \$1.25, as in Figure 5.5. Since the extra gallon cost only \$1 to produce, the seller is \$0.25 better off than before. And since the most eager buyer values the extra gallon at \$2, that buyer is \$0.75 better off than before. In sum, the transaction creates an extra \$1 of economic surplus out of thin air!

Note that none of the other buyers or sellers is harmed by this transaction. Thus, milk selling for only \$1 per gallon cannot be efficient. As the following Self-Test 5.3 illustrates, there was nothing special about the price of \$1 per gallon. Indeed, if milk sells for *any* price below \$1.50 per gallon (the market equilibrium price), we can design a similar transaction, which means that selling milk for any price less than \$1.50 per gallon cannot be efficient.

SELF-TEST 5.3

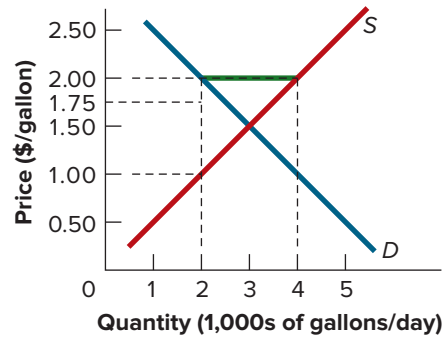
In Figure 5.4, suppose that milk initially sells for 50 cents per gallon. Describe a transaction that will create additional economic surplus for both buyer and seller without causing harm to anyone else.

Furthermore, it is always possible to describe a transaction that will create additional surplus for both buyer and seller whenever the price lies *above* the market equilibrium level. Suppose, for example, that the current price is \$2 per gallon in the milk market shown in Figure 5.6. At that price, we have excess supply of 2,000 gallons per day. Suppose the most dissatisfied producer sells a gallon of milk for \$1.75 to the buyer who values it most highly. This buyer, who would have been willing to pay \$2, will be \$0.25 better off than before. Likewise the producer, who would have been willing to sell milk for as little as \$1 per gallon (the marginal cost of production at 2,000 gallons per day), will be \$0.75 better off than before. As when the price was \$1 per gallon, the new transaction creates \$1 of additional economic surplus without harming any other buyer or seller. Since we could design a similar surplus-enhancing transaction at any price above the equilibrium level, selling milk for more than \$1.50 per gallon cannot be efficient.

The vertical interpretations of the supply and demand curves thus make it clear why only the equilibrium price in a market can be efficient. When the price is either higher or lower than the equilibrium price, the quantity exchanged in the market will

FIGURE 5.6**How Excess Supply Creates an Opportunity for a Surplus-Enhancing Transaction.**

At a market price of \$2 per gallon, dissatisfied sellers can produce an additional gallon of milk at a cost of only \$1, which is \$1 less than a buyer would be willing to pay for it. If the buyer pays the seller \$1.75 for an extra gallon, the buyer gains an economic surplus of \$0.25 and the seller gains an economic surplus of \$0.75.



always be lower than the equilibrium quantity. If the price is below equilibrium, the quantity sold will be the amount that sellers offer. If the price is above equilibrium, the quantity sold will be the amount that buyers wish to buy. In either case, the vertical value on the demand curve at the quantity exchanged, which is the value of an extra unit to buyers, must be larger than the vertical value on the supply curve, which is the marginal cost of producing that unit.

So the market equilibrium price is the *only* price at which buyers and sellers cannot design a surplus-enhancing transaction. The market equilibrium price leads, in other words, to the largest possible total economic surplus. In this specific, limited sense, free markets are said to produce and distribute goods and services efficiently.

Actually, to claim that market equilibrium is always efficient even in this limited sense is an overstatement. The claim holds only if markets are perfectly competitive and if the demand and supply curves satisfy certain other restrictions. For example, market equilibrium will not be efficient if the individual marginal cost curves that add up to the market supply curve fail to include all relevant costs of producing the product. Thus, as we saw in Chapter 2, *Supply and Demand*, the true cost of expanding output will be higher than indicated by the market supply curve if production generates pollution that harms others. The equilibrium output will then be inefficiently large and the equilibrium price inefficiently low.

Likewise, market equilibrium will not be efficient if the individual demand curves that make up the market demand curve do not capture all the relevant benefits of buying additional units of the product. For instance, if a homeowner's willingness to pay for ornamental shrubs is based only on the enjoyment she herself gains from them, and not on any benefits that may accrue to her neighbors, the market demand curve for shrubs will understate their value to the neighborhood. The equilibrium quantity of ornamental shrubs will be inefficiently small and the market price for shrubs will be inefficiently low.

We will take up such market imperfections in greater detail in later chapters. For now, we will confine our attention to perfectly competitive markets whose demand curves capture all relevant benefits and whose supply curves capture all relevant costs. For such goods, market equilibrium will always be efficient in the limited sense described earlier.

Efficiency Is Not the Only Goal

The fact that market equilibrium maximizes economic surplus is an attractive feature, to be sure. Bear in mind, however, that "efficient" does not mean the same thing as "good." For example, the market for milk may be in equilibrium at a price of \$1.50 per gallon, yet many poor families may be unable to afford milk for their children at that price. Still others may not even have a place for their children to sleep.

Efficiency is a concept that is based on predetermined attributes of buyers and sellers—their incomes, tastes, abilities, knowledge, and so on. Through the combined effects of individual cost-benefit decisions, these attributes give rise to the supply and demand curves for each good produced in an economy. If we are concerned about inequality in the distribution of attributes like income, we should not be surprised to discover that markets do not always yield outcomes we like.

Most of us could agree, for example, that the world would be a better one if all people had enough income to feed their families adequately. The claim that equilibrium in the market for milk is efficient means simply that *taking people's incomes as given*, the resulting allocation of milk cannot be altered so as to help some people without at the same time harming others.

To this a critic of the market system might respond: So what? As such critics rightly point out, imposing costs on others may be justified if doing so will help those with sufficiently important unmet demands. For example, most people would prefer to fund homeless shelters with their tax dollars rather than let homeless people freeze to death. Arguing in these terms, American policymakers responded to rapid increases in the price of oil in the late 1970s by imposing price controls on home heating oil. Many of us might agree that if the alternative had been to take no action at all, price controls might have been justified in the name of social justice.

The economist's concept of market efficiency makes clear that there *must* be a better alternative policy. Price controls on oil prevent the market from reaching equilibrium, and as we've seen, that means forgoing transactions that would benefit some people without harming others.

Why Efficiency Should Be the First Goal

Efficiency is important not because it is a desirable end in itself, but because it enables us to achieve all our other goals to the fullest possible extent. It is always possible to generate additional economic surplus when a market is out of equilibrium. To gain additional economic surplus is to gain more of the resources we need to do the things we want to do.

RECAP ↑

EQUILIBRIUM, SOCIAL OPTIMUM, AND EFFICIENCY

- A market in equilibrium is one in which no additional opportunities for gain remain available to individual buyers or sellers. Powerful economic incentives help push markets toward equilibrium. But even if all markets are in equilibrium, the resulting allocation of resources need not be socially optimal. Equilibrium will be socially optimal only if markets are perfectly competitive and if the costs and benefits to individual participants in the market are the same as those experienced by society as a whole.
- A market in equilibrium is said to be efficient, or Pareto efficient, meaning that no reallocation is possible that will benefit some people without harming others.
- When a market is not in equilibrium—because price is either above the equilibrium level or below it—the quantity exchanged is always less than the equilibrium level. At such a quantity, a transaction can always be made in which both buyer and seller benefit from the exchange of an additional unit of output.
- Total economic surplus in a market is maximized when exchange occurs at the equilibrium price. But the fact that equilibrium is “efficient” in this sense does not mean that it is “good.” All markets can be in equilibrium, yet many people may lack sufficient income to buy even basic goods and services. Still, permitting markets to reach equilibrium is important because, when economic surplus is maximized, it is possible to pursue every goal more fully.

THE COST OF PREVENTING PRICE ADJUSTMENTS

The governments of many countries employ price subsidies in the hope of protecting their citizens from rising world prices of essential goods such as gasoline. These countries have much lower incomes, on average, than countries that don't employ such subsidies. Yet despite their lower incomes, countries that subsidize fuel accounted for virtually the entire increase in worldwide oil consumption in recent years. Without this artificial demand stimulus, world oil prices would have been significantly lower.

EXAMPLE 5.7

Subsidies That Encourage Waste

How does subsidizing the price of gasoline affect people's behavior and what impact does it have on an economy's total economic surplus?

Subsidies are problematic because when the price of a good is below its cost, people use it wastefully. Suppose the world price of gasoline rose from \$2 a gallon to \$4 a gallon, but the government continued to sell gasoline for \$2 a gallon. Consider how this \$2/gallon subsidy might affect a trucker's decision about whether to accept a hauling job. A rational trucker will apply the basic cost-benefit test, which says that something is worth doing if, and only if, its benefit is at least as great as its cost. Suppose the job in question requires 1,000 gallons of fuel, available at the subsidized price of \$2 a gallon, for a total fuel outlay of \$2,000. If the cost of the trucker's time and equipment is, say, \$1,000 for the trip, his narrow interests dictate accepting the job if the shipper is willing to pay at least \$3,000. Suppose the shipper is willing to pay that amount but not more.

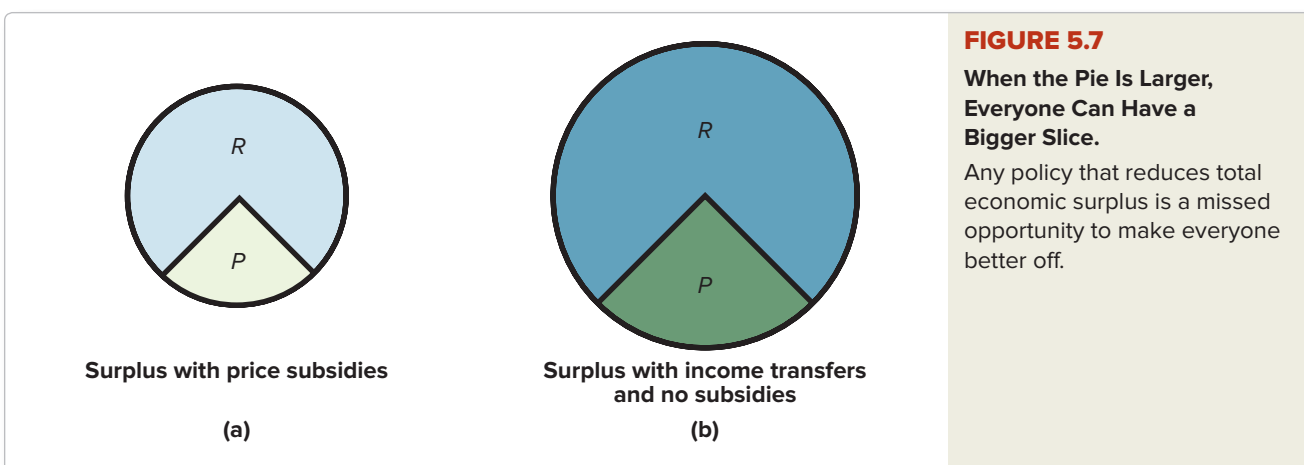
The problem is that if the trucker accepts the job at that price, the country as a whole will be worse off by more than \$2,000. Although the \$3,000 fee would cover his own costs, the government would end up paying \$2,000 in additional subsidies for the 1,000 gallons consumed. On top of that, the trip would generate additional pollution and congestion costs. So the fact that the subsidy encouraged him to accept the job means that its net effect is equivalent to throwing more than \$2,000 onto a bonfire.

The shortages that usually accompany fuel price subsidies encourage other forms of waste. For example, they prompt buyers to take costly actions to enhance their chances of being served. For example, if a service station begins selling its limited supply of subsidized fuel at 6 a.m., many buyers may arrive several hours early to ensure a place near the front of the line. Yet when all buyers show up earlier, no one gets any more gasoline than before.

To be sure, higher fuel prices produce economic suffering. The unfortunate reality, however, is that when the price of an imported resource rises in the world market, buyers must take a hit. Subsidizing fuel does nothing to reduce the inevitable suffering, and actually makes it worse.

Subsidy proponents cite the firestorm of political protest that would erupt if fuel were to sell at the international market price. That fuel subsidies are wasteful, however, implies that there must be less costly ways to keep the peace.

Consider again our trucker in Example 5.7 who accepted a job that barely covered the cost of his time, equipment, and subsidized fuel. Instead of paying \$2,000 to subsidize his fuel, the government could give him a tax cut of, say, \$1,000, and use the remaining \$1,000 to help pay for public services. Because the trucker's earnings from the hauling job were only enough to cover his costs at the subsidized fuel price, he would be \$1,000 better off with the tax cut alone than with the fuel subsidy. The additional support for public services would augment this benefit. In short, a tax cut is always a better way to keep political protest at bay because, unlike a fuel subsidy, it does not encourage shipments whose costs exceed their benefits.



Notwithstanding the fact that fuel subsidies reduce total economic surplus, their defenders might argue that they are justified because they enable at least some low-income families to buy fuel at affordable prices. Yes, but the same objective could have been accomplished in a much less costly way—namely, by giving the poor more income.

It may seem natural to wonder whether the poor, who have limited political power, can really hope to receive income transfers that would compensate for the burden of higher fuel prices. On reflection, the answer to this question would seem to be yes, if the alternative is to pay for fuel subsidies that would be even more costly than the income transfers.

This point is so important, and so often misunderstood by voters and policymakers, that we will emphasize it by putting it another way. Think of the economic surplus from a market as a pie to be divided among the various market participants. Figure 5.7(a) represents the total economic surplus available to participants in the gasoline market when the government subsidizes gasoline and limits the price of oil to \$1 per gallon. We divided this pie into two slices, labeled R and P , to denote the surpluses received by rich and poor participants.

Figure 5.7(b) represents the total economic surplus available when the price of gasoline is free to reach its equilibrium level. This pie is divided among rich and poor participants in the same proportion as the pie in the left panel.

The important point to notice is this: *Because the pie on the right side is larger, both rich and poor participants in the home heating oil market can get a bigger slice of the pie than they would have had under fuel subsidies.* Rather than tinker with the market price of gasoline, it is in everyone's interest to simply transfer additional income to the poor.

Supporters of fuel subsidies may object that income transfers to the poor might weaken people's incentive to work, and thus might prove extremely costly in the long run. Difficult issues do indeed arise in the design of programs for transferring income to the poor—issues we will consider in some detail in later chapters. But for now, suffice it to say that ways exist to transfer income without undermining work incentives significantly. One such method is the earned-income tax credit, a program that supplements the wages of low-income workers. Given such programs, transferring income to the poor will always be more efficient than trying to boost their living standard through price controls.

RECAP ↑

THE COST OF PREVENTING PRICE ADJUSTMENTS

In an effort to help low-income consumers, governments often provide price subsidies that make housing and other basic goods more affordable. Subsidies prevent markets from reaching equilibrium and typically reduce total economic surplus.

SUMMARY

- Accounting profit is the difference between a firm's revenue and its explicit expenses. It differs from economic profit, which is the difference between revenue and the sum of the firm's explicit and implicit costs. Normal profit is the difference between accounting profit and economic profit. It is the opportunity cost of the resources supplied to a business by its owners. (LO1)
- The quest for economic profit is the invisible hand that drives resource allocation in market economies. Markets in which businesses earn an economic profit tend to attract additional resources, whereas markets in which businesses experience an economic loss tend to lose resources. If new firms enter a market with economic profits, that market's supply curve shifts to the right, causing a reduction in the price of the product. Prices will continue to fall until economic profits are eliminated. By contrast, the departure of firms from markets with economic losses causes the supply curve in such markets to shift left, increasing the price of the product. Prices will continue to rise until economic losses are eliminated. In the long run, market forces drive economic profits and losses toward zero. (LO2, LO3)
- Economic rent is the portion of the payment for an input that exceeds the reservation price for that input. If a professional basketball player who is willing to play for as little as \$100,000 per year is paid \$15 million, he earns an economic rent of \$14,900,000 per year. Whereas the invisible hand drives economic profit toward zero over the long run, economic rent can persist indefinitely because replicating the services of players like Steph Curry is impossible. Talented individuals who are responsible for the superior performance of a business will tend to capture the resulting financial gains as economic rents. (LO3)
- The benefit of an investment to an individual sometimes differs from its benefit to society as a whole. Such conflicting incentives may give rise to behavior that is smart for one but dumb for all. Despite such exceptions, the invisible hand of the market works remarkably well much of the time. One of the market system's most important contributions to social well-being is the pressure it creates to adopt cost-saving innovations. Competition among firms ensures that the resulting cost savings get passed along to consumers in the long run. (LO4)
- When market supply and demand curves reflect the underlying costs and benefits to society of the production of a good or service, the quest for economic profit ensures not only that existing supplies are allocated efficiently among individual buyers, but also that resources are allocated across markets in the most efficient way possible. In any allocation other than the one generated by the market, resources could be rearranged to benefit some people without harming others. (LO4)
- When the supply and demand curves for a product capture all the relevant costs and benefits of producing that product, then market equilibrium for that product will be efficient. In such a market, if price and quantity do not equal their equilibrium values, a transaction can be found that will make at least some people better off without harming others. (LO4)
- Total economic surplus is a measure of the amount by which participants in a market benefit by participating in it. It is the sum of total consumer surplus and total producer surplus in the market. One of the attractive properties of market equilibrium is that it maximizes the value of total economic surplus. (LO5)
- Efficiency should not be equated with social justice. If we believe that the distribution of income among people is unjust, we won't like the results produced by the intersection of the supply and demand curves based on that income distribution, even though those results are efficient. (LO5)
- Even so, we should always strive for efficiency because it enables us to achieve all our other goals to the fullest possible extent. Whenever a market is out of equilibrium, the economic pie can be made larger. And with a larger pie, everyone can have a larger slice. (LO5)
- Regulations or policies that prevent markets from reaching equilibrium—such as price ceilings and price subsidies—are often defended on the grounds that they help the poor. But such schemes reduce economic surplus, meaning that we can find alternatives under which both rich and poor would be better off. The main difficulty of the poor is that they have too little income. Rather than trying to control the prices of the goods they buy, we could do better by enacting policies that raise the incomes of the poor and then letting prices seek their equilibrium levels. Those who complain that the poor lack the political power to obtain such income transfers must explain why the poor have the power to impose regulations that are far more costly than income transfers. (LO5)

KEY TERMS

accounting profit
allocative function of price
barrier to entry
economic loss

economic profit (or excess profit)
economic rent
efficient (or Pareto efficient)
explicit costs

implicit costs
invisible hand theory
normal profit
rationing function of price

REVIEW QUESTIONS

- How can a business owner who earns \$10 million per year from his business credibly claim to earn zero economic profit? (LO1)
- Why do most cities in the United States now have more radios but fewer radio-repair shops than they did in 1960? (LO2)
- Why do market forces drive economic profit but not economic rent toward zero? (LO3)
- Why do economists emphasize efficiency as an important goal of public policy? (LO4)
- You are a senator considering how to vote on a policy that would increase the economic surplus of workers by \$100 million per year but reduce the economic surplus of retirees by \$1 million per year. What additional measure might you combine with the policy to ensure that the overall result is a better outcome for everyone? (LO5)

PROBLEMS



- Explain why the following statements are true or false. (LO1, LO2, LO3, LO4)
 - The economic maxim “There’s no cash on the table” means that there are never any unexploited economic opportunities.
 - Firms in competitive environments make no accounting profit when the market is in long-run equilibrium.
 - Firms that can introduce cost-saving innovations can make an economic profit in the short run.
- Jaime owns and manages a café in Collegetown whose annual revenue is \$5,000. Annual expenses are as follows. (LO1, LO2)

Labor	\$2,000
Food and drink	500
Electricity	100
Vehicle lease	150
Rent	500
Interest on loan for equipment	1,000

 - Calculate Jaime’s annual accounting profit.
 - Jaime could earn \$1,000 per year as a recycler of aluminum cans. However, she prefers to run the café. In fact, she would be willing to pay up to \$275 per year to run the café rather than to recycle. Is the café making an economic profit? Should Jaime stay in the café business? Explain.
 - Suppose the café’s revenues and expenses remain the same, but recyclers’ earnings rise to \$1,100 per year. Is the café still making an economic profit? Explain.
 - Suppose Jaime had not gotten a \$10,000 loan at an annual interest rate of 10 percent to buy equipment, but instead had invested \$10,000 of her own money in equipment. How would your answer to parts a and b change?
 - If Jaime can earn \$1,000 a year as a recycler, and she likes recycling just as well as running the café, how much additional revenue would the café have to collect each year to earn a normal profit?
- Explain carefully why, in the absence of a patent, a technical innovation invented and pioneered in one tofu factory will cause the supply curve for the entire tofu industry to shift to the right. What will finally halt the rightward shift? (LO2)
- The city of New Orleans has 200 advertising companies, 199 of which employ designers of normal ability at a salary of \$100,000 a year. The companies that employ normal designers each collect \$500,000 in revenue a year, which is just enough to ensure that each earns exactly a normal profit. The 200th company, however, employs Janus, an unusually talented designer. Because of Janus’s talent, this company collects \$1,000,000 in revenue a year. (LO3)
 - How much will Janus earn? What proportion of her annual salary will be economic rent?
 - Why won’t the advertising company for which Janus works be able to earn an economic profit?
- Unskilled workers in a poor cotton-growing region must choose between working in a factory for \$6,000 a year and being a tenant cotton farmer. One farmer can work a 120-acre farm, which rents for \$10,000 a year. Such farms yield \$20,000 worth of

cotton each year. The total nonlabor cost of producing and marketing the cotton is \$4,000 a year. A local politician whose motto is “Working people come first” has promised that if she is elected, her administration will fund a fertilizer, irrigation, and marketing scheme that will triple cotton yields on tenant farms at no charge to tenant farmers. (LO3)

- a. If the market price of cotton would be unaffected by this policy and no new jobs would be created in the cotton-growing industry, how would the project affect the incomes of tenant farmers in the short run? In the long run?
 - b. Who would reap the benefit of the scheme in the long run? How much would they gain each year?
6. The town of Wells has 50 restaurants, 49 of which employ chefs of average ability, who earn \$200 per week, which is the usual market price for a chef. Paying this salary, each of these 49 restaurants makes normal profit and collects \$1,000 per week in total revenue. The 50th restaurant employs an unusually talented chef, and collects \$1,500 per week

in total revenue because consumers are willing to pay more for the meals he cooks. This chef’s weekly salary will be (LO3)

- a. \$500 per week.
 - b. \$200 per week.
 - c. \$700 per week.
 - d. \$1,500 per week.
 - e. None of the above.
7. Arguing that many older adults have difficulty paying their electric bills, regulators have ordered an electric utility company to charge them half the rate charged to other customers for each kilowatt-hour of power they use. The normal rates closely mirror the cost of providing power. (LO5)
- a. How will this affect their decision to leave on their porch lights at night?
 - b. What effect will the new regulation have on total economic surplus?
8. In the preceding problem, how might regulators have tried to address their policy concern in a more efficient way? (LO5)

ANSWERS TO SELF-TESTS

5.1 As shown in the table below, Pudge’s accounting profit is now \$10,000, the difference between his \$20,000 annual revenue and his \$10,000-per-year payment for land, equipment, and supplies. His economic profit is that amount minus the opportunity cost of his labor—again, the \$11,000 per year he could have earned as a store manager. So Pudge is now earning a negative economic profit, –\$1,000 per year. As before, his normal profit is

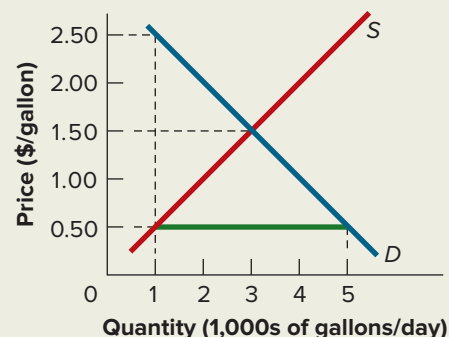
the \$11,000-per-year opportunity cost of his labor. Although an accountant would say Pudge is making an annual profit of \$10,000, that amount is less than a normal profit for his activity. An economist would therefore say that he is making an economic loss of \$1,000 per year. Since Pudge likes the two jobs equally well, he will be better off by \$1,000 per year if he leaves farming to become a manager. (LO1)

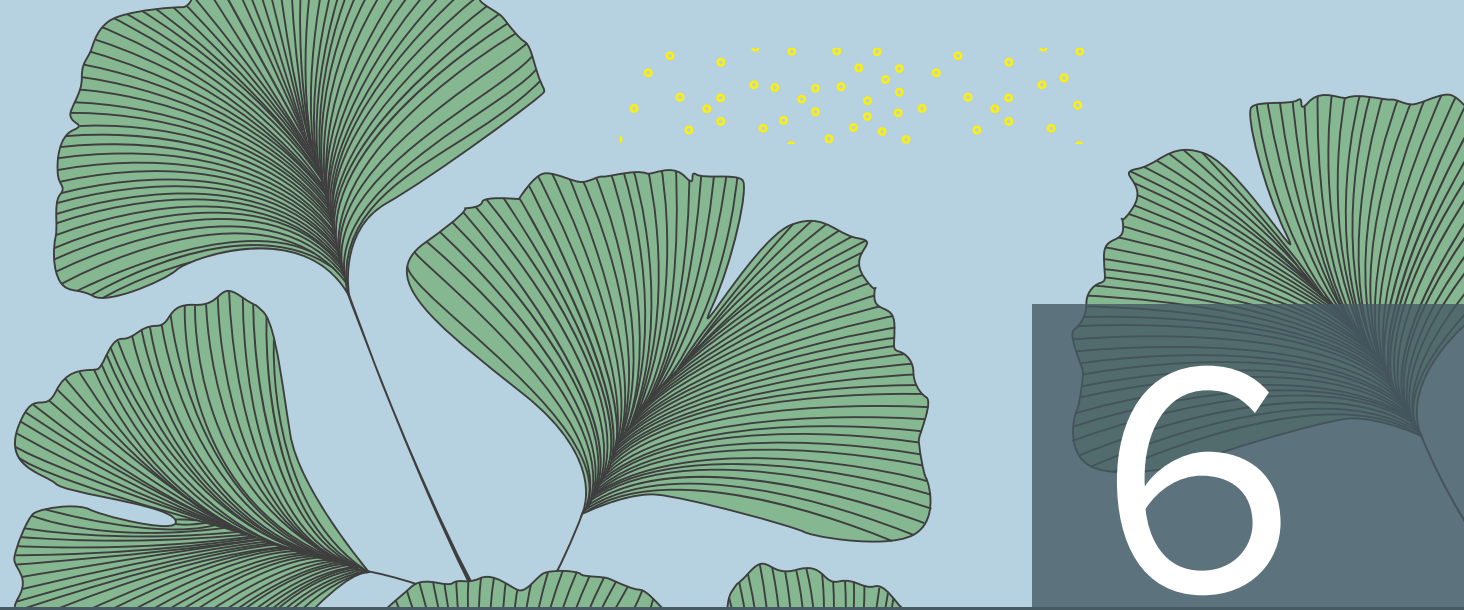
Total revenue (\$/year)	Explicit costs (\$/year)	Implicit costs (\$/year)	Accounting profit (= total revenue – explicit costs) (\$/year)	Economic profit (= total revenue – explicit costs – implicit costs) (\$/year)	Normal profit (= implicit costs) (\$/year)
20,000	10,000	11,000	10,000	–1,000	11,000

5.2 If each lane did not move at about the same pace, any driver in a slower lane could reduce his travel time by simply switching to a faster one. People will exploit these opportunities until each lane moves at about the same pace. (LO3)

5.3 At a price of 50 cents per gallon, there is excess demand of 4,000 gallons per day. Suppose a seller produces an extra gallon of milk (marginal cost = 50 cents) and sells it to the buyer who values it most (reservation price = \$2.50) for \$1.50. Both buyer and seller will gain additional economic

surplus of \$1, and no other buyers or sellers will be hurt by the transaction. (LO4)





Monopoly, Oligopoly, and Monopolistic Competition

Some years ago, fantasy game enthusiasts around the country became obsessed with the game of Magic: The Gathering. To play, you need a deck of Magic cards, available only from the creators of the game. But unlike ordinary playing cards, which can be bought in most stores for only a dollar or two, a deck of Magic cards sells for many times that. And because Magic cards cost no more to manufacture than ordinary playing cards, their producer earns an enormous economic profit.

In a perfectly competitive market, entrepreneurs would see this economic profit as cash on the table. It would entice them to offer Magic cards at slightly lower prices so that, eventually, the cards would sell for roughly their cost of production, just as ordinary playing cards do. But Magic cards have been on the market for years now, and that hasn't happened. The reason is that the cards are copyrighted, which means the government has granted the creators of the game an exclusive license to sell them.

imperfectly competitive firm (or **price setter**) a firm that has at least some control over the market price of its product

The holder of a copyright is an example of an **imperfectly competitive firm, or price setter**—that is, a firm with at least some latitude to set its own price. The competitive firm, by contrast, is a price taker, a firm with no influence over the price of its product.

Our focus in this chapter will be on the ways in which markets served by imperfectly competitive firms differ from those served by perfectly competitive firms. One salient difference is the imperfectly competitive firm's ability, under certain circumstances, to charge more than its cost of production. But if the producer of Magic cards could charge any price it wished, why does it charge only \$10? Why not \$100, or even \$1,000? We'll see that even though such a company may be the only seller of its product, its pricing freedom is far from absolute. We'll also see how some imperfectly competitive firms manage to earn an economic

LEARNING OBJECTIVES

After reading this chapter, you should be able to:

- LO1** Distinguish among three types of imperfectly competitive industries and describe how imperfect competition differs from perfect competition.
- LO2** Identify the five sources of market power.
- LO3** Describe how economies of scale are affected by how large fixed costs are in relation to marginal cost.
- LO4** Apply the concepts of marginal cost and marginal revenue to find the output level and price that maximize a monopolist's profit.
- LO5** Explain why the profit-maximizing output level for a monopolist is too small from society's perspective.
- LO6** Discuss why firms often offer discounts to buyers who are willing to jump some form of hurdle.



biglom/Getty Images

Why do Magic cards sell for 10 times as much as ordinary playing cards, even though they cost no more to produce?

profit, even in the long run, and even without government protections like copyright. And we'll explore why Adam Smith's invisible hand is less evident in a world served by imperfectly competitive firms.

PERFECT AND IMPERFECT COMPETITION

The perfectly competitive market is an ideal; the actual markets we encounter in everyday life differ from the ideal in varying degrees. Economics texts usually distinguish among three types of imperfectly competitive market structures. The classifications are somewhat arbitrary, but they are quite useful in analyzing real-world markets.

pure monopoly the only supplier of a unique product with no close substitutes

Different Forms of Imperfect Competition

Farthest from the perfectly competitive ideal is the **pure monopoly**, a market in which a single firm is the lone seller of a unique product. The producer of Magic cards is a pure monopolist, as are many providers of electric power. If the residents of Miami don't buy their electricity from Florida Power and Light Company, they simply do without. In between these two extremes are many different types of imperfect competition. We focus on two of them here: monopolistic competition and oligopoly.

monopolistic competition an industry structure in which a large number of firms produce slightly differentiated products that are reasonably close substitutes for one another

Monopolistic Competition

Recall from Chapter 4, *Perfectly Competitive Supply*, that in a perfectly competitive industry, a large number of firms typically sell products that are essentially perfect substitutes for one another. In contrast, **monopolistic competition** is an industry structure in which a large number of rival firms sell products that are close, but not quite perfect, substitutes. Rival products may be highly similar in many respects, but there are always at least some features that differentiate one product from another in the eyes of some consumers. Monopolistic competition has in common with perfect competition the feature that there are no significant barriers preventing firms from entering or leaving the market.

Local gasoline retailing is an example of a monopolistically competitive industry. The gas sold by different stations may be nearly identical in chemical terms, but a station's particular location is a feature that matters for many consumers. Convenience stores are another example. Although most of the products found on any given store's shelves are also carried by most other stores, the product lists of different stores are not identical. Some offer sim cards and lottery tickets, for example, while others offer foam coolers and flashlights. And even more so than in the case of gasoline retailing, location is an important differentiating feature of convenience stores.

Recall that if a perfectly competitive firm were to charge even just slightly more than the prevailing market price for its product, it would not sell any output at all. Things are different for the monopolistically competitive firm. The fact that its offering is not a perfect substitute for those of its rivals means that it can charge a slightly higher price than they do and not lose all its customers.

But that does not mean that monopolistically competitive firms can expect to earn positive economic profits in the long run. On the contrary, because new firms are able to enter freely, a monopolistically competitive industry is essentially the same as a perfectly competitive industry in this respect. If existing

monopolistically competitive firms were earning positive economic profits at prevailing prices, new firms would have an incentive to enter the industry. Downward pressure on prices would then result as the larger number of firms competed for a limited pool of potential customers.¹ As long as positive economic profits remained, entry would continue and prices would be driven ever lower. Conversely, if firms in a monopolistically competitive industry were initially suffering economic losses, some firms would begin leaving the industry. As long as economic losses remained, exit and the resulting upward pressure on prices would continue. So in long-run equilibrium, monopolistically competitive firms are in this respect like perfectly competitive firms: All expect to earn zero economic profit.

Although monopolistically competitive firms have some latitude to vary the prices of their product in the short run, pricing is not the most important strategic decision they confront. A far more important issue is how to differentiate their products from those of existing rivals. Should a product be made to resemble a rival's product as closely as possible? Or should the aim be to make it as different as possible? Or should the firm strive for something in between? We'll consider these questions in Chapter 7, *Games and Strategic Behavior*, where we'll focus on this type of strategic decision making.

Oligopoly

Further along the continuum between perfect competition and pure monopoly lies **oligopoly**, a structure in which the entire market is supplied by a small number of large firms. Cost advantages associated with large size are one of the primary reasons for pure monopoly, as we will discuss presently. Oligopoly is also typically a consequence of cost advantages that prevent small firms from being able to compete effectively.

In some cases, oligopolists sell undifferentiated products. In the market for wireless phone service, for example, the offerings of AT&T, Verizon, and T-Mobile are very similar. The cement industry is another example of an oligopoly selling an essentially undifferentiated product. The most important strategic decisions facing firms in such cases are more likely to involve pricing and advertising than specific features of their product. Here, too, we leave a more detailed discussion of such decisions to Chapter 7, *Games and Strategic Behavior*.

In other cases, such as the automobile and tobacco industries, oligopolists are more like monopolistic competitors than pure monopolists, in the sense that differences in their product features have significant effects on consumer demand. Many long-time Ford buyers, for example, would not even consider buying a Chevrolet, and very few smokers ever switch from Camels to Marlboros. As with oligopolists who produce undifferentiated products, pricing and advertising are important strategic decisions for firms in these industries, but so, too, are those related to specific product features.

Because cost advantages associated with large size are usually so important in oligopolies, there is no presumption that entry and exit will push economic profit to zero. Consider, for example, an oligopoly served by two firms, each of which currently earns an economic profit. Should a new firm enter this market? Possibly, but it also might be that a third firm large enough to achieve the cost advantages of the two incumbents would effectively flood the market, driving price so low that all three firms would suffer economic losses. There is no guarantee, however, that an oligopolist will earn a positive economic profit.

oligopoly an industry structure in which a small number of large firms produce products that are either close or perfect substitutes

¹See Edward Chamberlin, *The Theory of Monopolistic Competition* (Cambridge, MA: Harvard University Press, 1st ed. 1933, 8th ed. 1962); and Joan Robinson, *The Economics of Imperfect Competition* (London: Macmillan, 1st ed. 1933, 2nd ed. 1969).

As we'll see in the next section, the essential characteristic that differentiates imperfectly competitive firms from perfectly competitive firms is the same in each of the three cases. So for the duration of this chapter, we'll use the term *monopolist* to refer to any of the three types of imperfectly competitive firms. In the next chapter, we'll consider the strategic decisions confronting oligopolists and monopolistically competitive firms in greater detail.

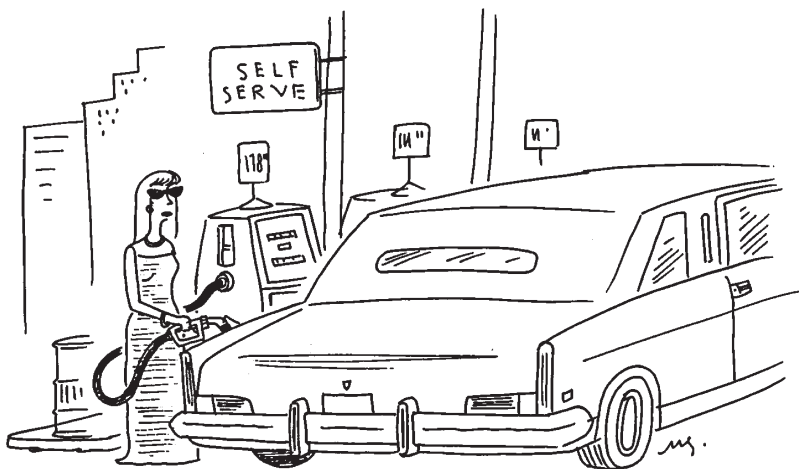
The Essential Difference between Perfectly and Imperfectly Competitive Firms

In advanced economics courses, professors generally devote much attention to the analysis of subtle differences in the behavior of different types of imperfectly competitive firms. Far more important for our purposes, however, will be to focus on the single, common feature that differentiates all imperfectly competitive firms from their perfectly competitive counterparts—namely, that *whereas the perfectly competitive firm faces a perfectly elastic demand curve for its product, the imperfectly competitive firm faces a downward-sloping demand curve.*

In the perfectly competitive industry, the supply and demand curves intersect to determine an equilibrium market price. At that price, the perfectly competitive firm can sell as many units as it wishes. It has no incentive to charge more than the market price because it won't sell anything if it does so. Nor does it have any incentive to

charge less than the market price because it can sell as many units as it wants to at the market price. The perfectly competitive firm's demand curve is thus a horizontal line at the market price, as we saw in previous chapters.

By contrast, if a local gasoline retailer—an imperfect competitor—charges a few pennies more than its rivals for a gallon of gas, some of its customers may desert it. But others will remain, perhaps because they are willing to pay a little extra to continue stopping at their most convenient location. An imperfectly competitive firm thus faces a negatively sloped demand curve. Figure 6.1 summarizes this contrast between the demand curves facing perfectly competitive and imperfectly competitive firms.



If the BP station at State and Meadow Streets raised its gasoline prices by 3 cents per gallon, would all its customers shop elsewhere?

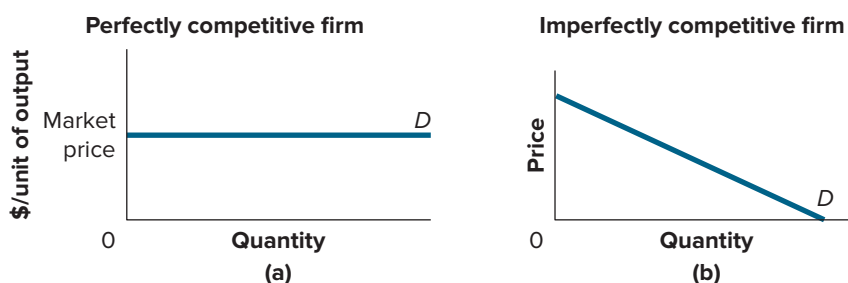


FIGURE 6.1

The Demand Curves Facing Perfectly and Imperfectly Competitive Firms.

- (a) The demand curve confronting a perfectly competitive firm is perfectly elastic at the market price.
- (b) The demand curve confronting an imperfectly competitive firm is downward-sloping.

RECAP ↑**PERFECT AND IMPERFECT COMPETITION**

Monopolistic competition is the industry structure in which a large number of small firms offer products that are similar in many respects, yet not perfect substitutes in the eyes of at least some consumers. Monopolistically competitive industries resemble perfectly competitive industries, in that entry and exit cause economic profits to tend toward zero in the long run.

Oligopoly is the industry structure in which a small number of large firms supply the entire market. Cost advantages associated with large-scale operations tend to be important. Oligopolists may produce either standardized products or differentiated products.

In contrast to each form of imperfect competition, in which all firms face downward-sloping demand curves, perfectly competitive firms face demand curves that are horizontal at the prevailing market price.

FIVE SOURCES OF MARKET POWER

Firms that confront downward-sloping demand curves are said to enjoy **market power**, a term that refers to their ability to set the prices of their products. A common misconception is that a firm with market power can sell any quantity at any price it wishes. It cannot. All it can do is pick a price–quantity combination on its demand curve. If the firm chooses to raise its price, it must settle for reduced sales.

Why do some firms have market power while others don't? Market power often carries with it the ability to charge a price above the cost of production, so such power tends to arise from factors that limit competition. In practice, the following five factors often confer such power: exclusive control over inputs, patents and copyrights, government licenses or franchises, economies of scale, and network economies.

market power a firm's ability to raise the price of a good without losing all its sales

Exclusive Control over Important Inputs

If a single firm controls an input essential to the production of a given product, that firm will have market power. For example, to the extent that some U.S. tenants are willing to pay a premium for office space in the country's tallest building, One World Trade Center, the owner of that building has market power.

Patents and Copyrights

Patents give the inventors or developers of new products the exclusive right to sell those products for a specified period of time. By insulating sellers from competition for an interval, patents enable innovators to charge higher prices to recoup their product's development costs. Pharmaceutical companies, for example, spend millions of dollars on research in the hope of discovering new drug therapies for serious illnesses. The drugs they discover are insulated from competition for an interval—currently 20 years in the United States—by government patents. For the life of the patent, only the patent holder may legally sell the drug. This protection enables the patent holder to set a price above the marginal cost of production to recoup the cost of the research on the drug. In the same way, copyrights protect the authors of movies, software, music, books, and other published works.

Government Licenses or Franchises

Yosemite Hospitality, LLC has an exclusive license from the U.S. government to run the lodging and concession operations at Yosemite National Park. One of the government's goals in granting this monopoly was to preserve the wilderness character of the area to the greatest degree possible. And indeed, the inns and cabins offered by

Yosemite Hospitality, LLC blend nicely with the valley's scenery. No garish neon signs mar the national park as they do in places where rivals compete for the tourist's dollars.

Economies of Scale and Natural Monopolies

constant returns to scale a production process is said to have constant returns to scale if, when all inputs are changed by a given proportion, output changes by the same proportion

increasing returns to scale (or **economies of scale**) a production process is said to have increasing returns to scale if, when all inputs are changed by a given proportion, output changes by more than that proportion

natural monopoly a monopoly that results from economies of scale (increasing returns to scale)

When a firm doubles all its factors of production, what happens to its output? If output exactly doubles, the firm's production process is said to exhibit **constant returns to scale**. If output more than doubles, the production process is said to exhibit **increasing returns to scale**, or **economies of scale**. When production is subject to economies of scale, the average cost of production declines as the number of units produced increases. For example, in the generation of electricity, the use of larger generators lowers the unit cost of production. The markets for such products tend to be served by a single seller, or perhaps only a few sellers, because having a large number of sellers would result in significantly higher costs. A monopoly that results from economies of scale is called a **natural monopoly**.

Network Economies

Although most of us don't care what brand of a product like dental floss others use, many other products become more valuable to us as more people use them. In the case of social media, for instance, the more people use a particular social media platform, the more valuable it becomes. After all, the goal of social media is to connect with others. With over 1 billion active monthly users, posts on Instagram, for example, have the potential to be seen by large numbers of people. If, instead, Instagram had only a handful of users, none of whom you knew, the value of posting would be much smaller. Indeed, if many of your friends stopped using Instagram, you might follow suit.

A similar network economy helps account for the dominant position of Microsoft's Windows operating system, which, as noted earlier, is currently installed in roughly 80 percent of all desktop computers. Because Microsoft's initial sales advantage gave software developers a strong incentive to write for the Windows format, the inventory of available software in the Windows format is now vastly larger than that for any competing operating system. And although general-purpose software such as word processors and spreadsheets continues to be available for multiple operating systems, specialized professional software and games usually appear first—and often only—in the Windows format. This software gap and the desire to achieve compatibility for file sharing gave people a good reason for choosing Windows, even if, as in the case of many Apple users, they believed a competing system was otherwise superior. But, again, network dominance need not be permanent, as witnessed by Apple's dramatic resurgence in recent years.

By far the most important and enduring of these sources of market power are economies of scale and network economies. Lured by economic profit, firms almost always find substitutes for exclusive inputs. If there's enough profit to be had by renting out space in this country's tallest building, some real estate developer will eventually build one taller than One World Trade Center in New York. Likewise, firms can often evade patent laws by making slight changes in the design of products. Patent protection is only temporary, in any case. Finally, governments grant very few franchises each year. But economies of scale are both widespread and enduring, even if not completely insurmountable.

Firmly entrenched network economies can be as persistent a source of natural monopoly as economies of scale. Indeed, network economies are essentially similar to economies of scale. When network economies are of value to the consumer, a product's quality increases as the number of users increases, so we can say that any given quality level can be produced at lower cost as sales volume increases. Thus network economies may be viewed as just another form of economies of scale in production, and that's how we'll treat them here.

RECAP ↑**FIVE SOURCES OF MARKET POWER**

A firm's power to raise its price without losing its entire market stems from exclusive control of important inputs, patents and copyrights, government licenses, economies of scale, or network economies. By far the most important and enduring of these are economies of scale and network economies.

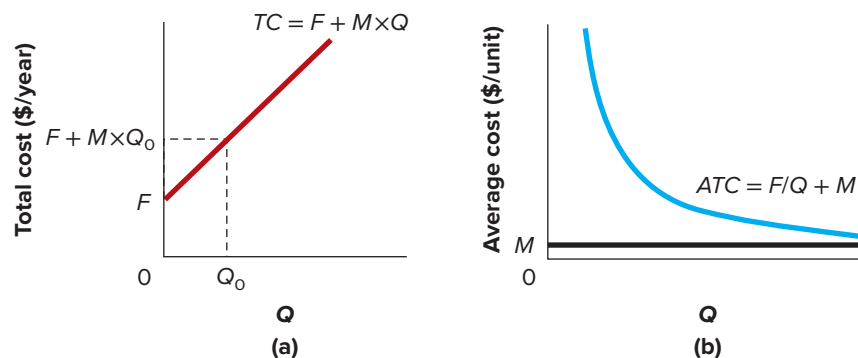
ECONOMIES OF SCALE AND THE IMPORTANCE OF START-UP COSTS

A firm's *variable costs* are those that vary with the level of output produced, while *fixed costs* are independent of output. Suppose, for example, that a firm produces output by employing one fixed input, capital, and one variable input, labor. Its payment to capital would then be a fixed cost, and its payment to labor a variable cost. Strictly speaking, there are no fixed costs in the long run because all inputs can be varied. But as a practical matter, start-up costs often loom large for the duration of a product's useful life. Most of the costs involved in the production of computer software, for example, are start-up costs of this sort, one-time costs incurred in writing and testing the software. Once those tasks are done, additional copies of the software can be produced at a very low marginal cost. A good such as software, whose production entails large fixed start-up costs and low variable costs, will be subject to significant economies of scale. Because, by definition, fixed costs don't increase as output increases, the average total cost of production for such goods will decline sharply as output increases.

To illustrate, consider a production process for which total cost is given by the equation $TC = F + M \times Q$, where F is fixed cost, M is marginal cost (assumed constant in this illustration), and Q is the level of output produced. For the production process with this simple total cost function, variable cost is simply $M \times Q$, the product of marginal cost and quantity. **Average total cost** (ATC), TC/Q , is equal to $F/Q + M$. As Q increases, average cost declines steadily because the fixed costs are spread out over more and more units of output.

average total cost a firm's total cost divided by its level of output

Figure 6.2 shows the total production cost (a) and average total cost (b) for a firm with the total cost curve $TC = F + M \times Q$ and the corresponding average total cost curve $ATC = F/Q + M$. The average total cost curve (b) shows the decline in per-unit cost as output grows. Though average total cost is always higher

**FIGURE 6.2**

Total and Average Total Costs for a Production Process with Economies of Scale.

For a firm whose total cost curve of producing Q units of output per year is $TC = F + M \times Q$, total cost (a) rises at a constant rate as output grows, while average total cost (b) declines. Average total cost is always higher than marginal cost for this firm, but the difference becomes less significant at high output levels.

than marginal cost for this firm, the difference between the two diminishes as output grows. At extremely high levels of output, average total cost becomes very close to marginal cost (M). Because the firm is spreading out its fixed cost over an extremely large volume of output, fixed cost per unit becomes almost insignificant.

As the following examples illustrate, the importance of economies of scale depends on how large fixed cost is in relation to marginal cost.

EXAMPLE 6.1 Economies of Scale—Small Fixed Cost

Two video game producers, Nintendo and Sony Computer Entertainment, each have fixed costs of \$200,000 and marginal costs of \$0.80 per game. If Nintendo produces 1 million units per year and Sony produces 1.2 million, how much lower will Sony’s average total production cost be?

Table 6.1 summarizes the relevant cost categories for the two firms. Note in the bottom row that Sony enjoys only a 3-cent average cost advantage over Nintendo. Even though Nintendo produces 20 percent fewer copies of its video game than Sony, it does not suffer a significant cost disadvantage because fixed cost is a relatively small part of total production cost.

TABLE 6.1
Costs for Two Video Game Producers (a)

	Nintendo	Sony Computer Entertainment
Annual production	1,000,000	1,200,000
Fixed cost	\$200,000	\$200,000
Variable cost	\$800,000	\$960,000
Total cost	\$1,000,000	\$1,160,000
Average total cost per game	\$1.00	\$0.97

In the next example, note how the picture changes when fixed cost looms large relative to marginal cost.

EXAMPLE 6.2 Economies of Scale—Large Fixed Cost

Two video game producers, Nintendo and Sony Computer Entertainment, each have fixed costs of \$10,000,000 and marginal costs of \$0.20 per video game. If Nintendo produces 1 million units per year and Sony produces 1.2 million, how much lower will Sony’s average total cost be?

The relevant cost categories for the two firms are now summarized in Table 6.2. The bottom row shows that Sony enjoys a \$1.67 average total cost advantage over Nintendo, substantially larger than in the previous example.

TABLE 6.2
Costs for Two Video Game Producers (b)

	Nintendo	Sony Computer Entertainment
Annual production	1,000,000	1,200,000
Fixed cost	\$10,000,000	\$10,000,000
Variable cost	\$200,000	\$240,000
Total cost	\$10,200,000	\$10,240,000
Average total cost per game	\$10.20	\$8.53

If the video games the two firms produce are essentially similar, the fact that Sony can charge significantly lower prices and still cover its costs should enable it to attract customers away from Nintendo. As more and more of the market goes to Sony, its cost advantage will become self-reinforcing. Table 6.3 shows how a shift of 500,000 units from Nintendo to Sony would cause Nintendo's average total cost to rise to \$20.20 per unit, while Sony's average total cost would fall to \$6.08 per unit. The fact that a firm cannot long survive at such a severe disadvantage explains why the video game market is served now by only a small number of firms.

TABLE 6.3
Costs for Two Video Game Producers (c)

	Nintendo	Sony Computer Entertainment
Annual production	500,000	1,700,000
Fixed cost	\$10,000,000	\$10,000,000
Variable cost	\$100,000	\$340,000
Total cost	\$10,100,000	\$10,340,000
Average total cost per game	\$20.20	\$6.08

SELF-TEST 6.1

How big will Sony's unit cost advantage be if it sells 2 million units per year, while Nintendo sells only 200,000?

An important worldwide economic trend during recent decades is that an increasing share of the value embodied in the goods and services we buy stems from fixed investment in research and development. For example, in 1984 some 80 percent of the cost of a computer was in its hardware (which has relatively high marginal cost); the remaining 20 percent was in its software. But by 1990 those proportions were reversed. Fixed cost now accounts for about 85 percent of total costs in the computer software industry, whose products are included in a growing share of ordinary manufactured goods.



The Economic Naturalist 6.1

Why does Intel sell the overwhelming majority of all microprocessors used in personal computers?

The fixed investment required to produce a new leading-edge microprocessor such as the Intel Core i9-9900KS microprocessor runs upward of several billion dollars. But once the chip has been designed and the manufacturing facility built, the marginal cost of producing each chip is only pennies. This cost pattern explains why Intel currently sells almost 80 percent of all microprocessors.

As fixed cost becomes more and more important, the perfectly competitive pattern of many small firms, each producing only a small share of its industry's total output, becomes less common. For this reason, we must develop a clear sense of how the behavior of firms with market power differs from that of the perfectly competitive firm.

RECAP

ECONOMIES OF SCALE AND THE IMPORTANCE OF START-UP COSTS

Research, design, engineering, and other fixed costs account for an increasingly large share of all costs required to bring products successfully to market. For products with large fixed costs, marginal cost is lower, often substantially, than average total cost, and average total cost declines, often sharply, as output grows. This cost pattern explains why many industries are dominated by either a single firm or a small number of firms.

PROFIT MAXIMIZATION FOR THE MONOPOLIST

Regardless of whether a firm is a price taker or a price setter, economists assume that its basic goal is to maximize its profit. In both cases, the firm expands output as long as the benefit of doing so exceeds the cost. Further, the calculation of marginal cost is also the same for the monopolist as for the perfectly competitive firm.

The profit-maximizing decision for a monopolist differs from that of a perfectly competitive firm when we look at the benefits of expanding output. For both the perfectly competitive firm and the monopolist, the marginal benefit of expanding output is the additional revenue the firm will receive if it sells one additional unit of output. In both cases, this marginal benefit is called the firm's **marginal revenue**. For the perfectly competitive firm, marginal revenue is exactly equal to the market price of the product. If that price is \$6, for example, then the marginal benefit of selling an extra unit is exactly \$6.

marginal revenue the change in a firm's total revenue that results from a one-unit change in output

Marginal Revenue for the Monopolist

The situation is different for a monopolist. *To a monopolist, the marginal benefit of selling an additional unit is strictly less than the market price.* As the following discussion will make clear, the reason is that while the perfectly competitive firm can sell as many units as it wishes at the market price, the monopolist can sell an additional unit only if it cuts the price—and it must do so not just for the additional unit but for the units it is currently selling.

Suppose, for example, that a monopolist with the demand curve shown in Figure 6.3 is currently selling 2 units of output at a price of \$6 per unit. What would be its marginal revenue from selling an additional unit?

**FIGURE 6.3****The Monopolist's Benefit from Selling an Additional Unit.**

The monopolist shown receives \$12 per week in total revenue by selling 2 units per week at a price of \$6 each. This monopolist could earn \$15 per week by selling 3 units per week at a price of \$5 each. In that case, the benefit from selling the third unit would be $\$15 - \$12 = \$3$, less than its selling price of \$5.

This monopolist's total revenue from the sale of 2 units per week is ($\$6$ per unit) (2 units per week) = \$12 per week. Its total revenue from the sale of 3 units per week would be \$15 per week. The difference—\$3 per week—is the marginal revenue from the sale of the third unit each week. Note that this amount is not only smaller than the original price (\$6) but smaller than the new price (\$5) as well.

SELF-TEST 6.2

Calculate marginal revenue for the monopolist in Figure 6.3 as it expands output from 3 to 4 units per week, and then from 4 to 5 units per week.

For the monopolist whose demand curve is shown in Figure 6.3, a sequence of increases in output—from 2 to 3, from 3 to 4, and from 4 to 5—will yield marginal revenue of \$3, \$1, and $-\$1$, respectively. We display these results in tabular form in Table 6.4.

TABLE 6.4**Marginal Revenue for a Monopolist (\$ per unit)**

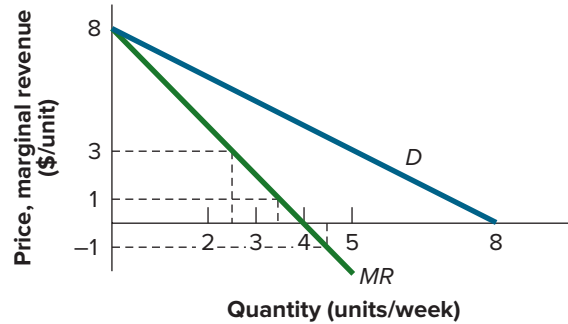
Quantity	Marginal revenue
2	3
3	1
4	−1
5	

Note in the table that the marginal revenue values are displayed between the two quantity figures to which they correspond. For example, when the firm expanded its output from 2 units per week to 3, its marginal revenue was \$3 per unit. Strictly speaking, this marginal revenue corresponds to neither quantity but to the movement between those quantities, hence its placement in the table. Likewise, in moving from 3 to 4 units per week, the firm earned marginal revenue of \$1 per unit, so that figure is placed midway between the quantities of 3 and 4, and so on.

To graph marginal revenue as a function of quantity, we would plot the marginal revenue for the movement from 2 to 3 units of output per week (\$3) at a quantity value of 2.5, because 2.5 lies midway between 2 and 3. Similarly, we would plot the marginal revenue for the movement from 3 to 4 units per week (\$1) at a quantity of 3.5 units per week, and the marginal revenue for the movement from 4 to 5 units per week ($-\$1$) at a quantity of 4.5. The resulting marginal revenue curve, *MR*, is shown in Figure 6.4.

FIGURE 6.4**Marginal Revenue in Graphical Form.**

Because a monopolist must cut price to sell an extra unit, not only for the extra unit sold but also for all existing units, marginal revenue from the sale of the extra unit is less than its selling price.



More generally, consider a monopolist with a straight-line demand curve whose vertical intercept is a and whose horizontal intercept is Q_0 , as shown in Figure 6.5. This monopolist's marginal revenue curve also will have a vertical intercept of a , and it will be twice as steep as the demand curve. Thus, its horizontal intercept will be not Q_0 , but $Q_0/2$, as shown in Figure 6.5.

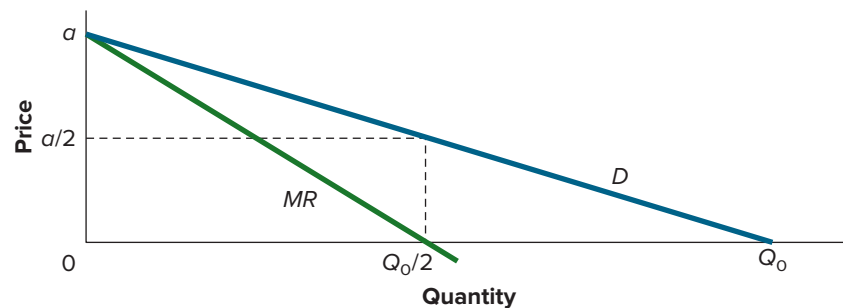
Marginal revenue curves also can be expressed algebraically. If the formula for the monopolist's demand curve is $P = a - bQ$, then the formula for its marginal revenue curve will be $MR = a - 2bQ$. If you have had calculus, this relationship is easy to derive,² but even without calculus you can verify it by working through a few numerical examples. First, translate the formula for the demand curve into a diagram, and then construct the corresponding marginal revenue curve graphically. Reading from the graph, write the formula for that marginal revenue curve.

The Monopolist's Profit-Maximizing Decision Rule

Having derived the monopolist's marginal revenue curve, we're now in a position to describe how the monopolist chooses the output level that maximizes profit. As in the case of the perfectly competitive firm, the Cost-Benefit Principle says that the monopolist should continue to expand output as long as the gain from doing so exceeds the cost. At the current level of output, the benefit from expanding output is the marginal revenue value that corresponds to that output level. The cost of expanding output is the marginal cost at that level of output. Whenever marginal revenue exceeds marginal cost, the firm should expand. Conversely, whenever marginal revenue falls short of marginal cost, the firm should reduce its output. *Profit is maximized at the level of output for which marginal revenue precisely equals marginal cost.*

FIGURE 6.5**The Marginal Revenue Curve for a Monopolist with a Straight-Line Demand Curve.**

For a monopolist with the demand curve shown, the corresponding marginal revenue curve has the same vertical intercept as the demand curve, and a horizontal intercept only half as large as that of the demand curve.



²For those who have had an introductory course in calculus, marginal revenue can be expressed as the derivative of total revenue with respect to output. If $P = a - bQ$, then total revenue will be given by $TR = PQ = aQ - bQ^2$, which means that $MR = dTR/dQ = a - 2bQ$.

When the monopolist's profit-maximizing rule is stated in this way, we can see that the perfectly competitive firm's rule is actually a special case of the monopolist's rule. When the perfectly competitive firm expands output by one unit, its marginal revenue exactly equals the product's market price (because the perfectly competitive firm can expand sales by a unit without having to cut the price of existing units). So when the perfectly competitive firm equates price with marginal cost, it is also equating marginal revenue with marginal cost. *Thus, the only significant difference between the two cases concerns the calculation of marginal revenue.*

EXAMPLE 6.3 Marginal Revenue

What is the monopolist's profit-maximizing output level?

Consider a monopolist with the demand and marginal cost curves shown in Figure 6.6. If this firm is currently producing 12 units per week, should it expand or contract production? What is the profit-maximizing level of output?

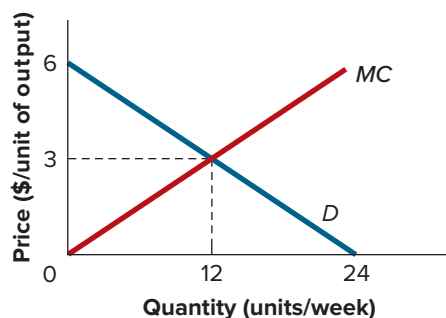


FIGURE 6.6

The Demand and Marginal Cost Curves for a Monopolist.

At the current output level of 12 units per week, price equals marginal cost. Because the monopolist's price is always greater than marginal revenue, marginal revenue must be less than marginal cost, which means this monopolist should produce less.

In Figure 6.7, we begin by constructing the marginal revenue curve that corresponds to the monopolist's demand curve. It has the same vertical intercept as the demand curve, and its horizontal intercept is half as large. Note that the monopolist's marginal revenue at 12 units per week is zero, which is clearly less than its marginal cost of \$3 per unit. This monopolist will therefore earn a higher profit by contracting production until marginal revenue equals marginal cost, which occurs at an output level of 8 units per week. At this profit-maximizing output level, the firm will charge \$4 per unit, the price that corresponds to 8 units per week on the demand curve.

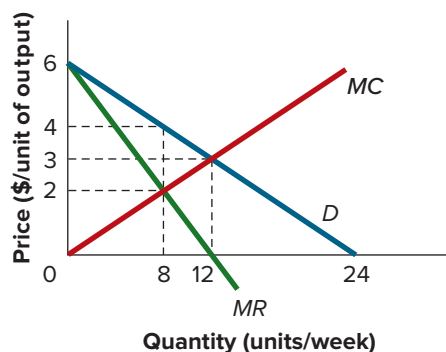


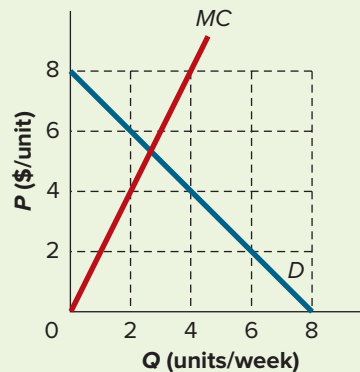
FIGURE 6.7

The Monopolist's Profit-Maximizing Output Level.

This monopolist maximizes profit by selling 8 units per week, the output level at which marginal revenue equals marginal cost. The profit-maximizing price is \$4 per unit, the price that corresponds to the profit-maximizing quantity on the demand curve.

SELF-TEST 6.3

For the monopolist with the demand and marginal cost curves shown, find the profit-maximizing price and level of output.



Being a Monopolist Doesn't Guarantee an Economic Profit

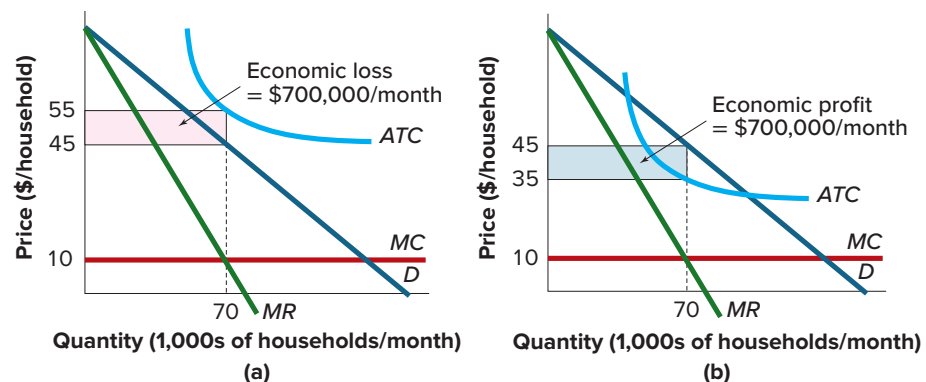
The fact that the profit-maximizing price for a monopolist will always be greater than marginal cost provides no assurance that the monopolist will earn an economic profit. Consider, for example, the local cable Internet service provider whose demand, marginal revenue, marginal cost, and average total cost curves are shown in Figure 6.8(a). This monopolist maximizes its monthly profit by providing 70 thousand households per month with cable Internet at price of \$45 per household. At that quantity, $MR = MC$, yet price is \$10 per household less than the company's average total cost of \$55 per household. As a result, the company sustains an economic loss of \$10 per household served, or a total loss of $(\$10 \text{ per household})(70,000 \text{ households per month}) = \$700,000 \text{ per month}$.

Recall that profit is the difference between a firm's total revenue ($P \times Q$) and its total cost. And because total cost is equal to average total cost times quantity ($ATC \times Q$), the firm's profit is equal to $(P \times Q) - (ATC \times Q) = (P - ATC) \times Q$. This observation suggests a convenient way to express profit graphically, as in Figure 6.8. When ATC is greater than P , as in Figure 6.8(a), the firm earns an economic loss, shown by the pink-shaded rectangle. When P is greater than ATC , as in Figure 6.8(b), it earns an economic profit, shown by the blue-shaded rectangle.

FIGURE 6.8

Even a Monopolist May Suffer an Economic Loss.

The monopolist in (a) maximizes its profit by providing cable Internet service to 70,000 households per month but suffers an economic loss of \$700,000 per month in the process. Because the profit-maximizing price of the monopolist in (b) exceeds ATC , this monopolist earns an economic profit.



The monopolist in Figure 6.8(a) suffered a loss because its profit-maximizing price was lower than its *ATC*. If the monopolist's profit-maximizing price exceeds its average total cost, however, the company will, of course, earn an economic profit. Consider, for example, the local cable Internet provider shown in Figure 6.8(b). This firm has the same demand, marginal revenue, and marginal cost curves as the firm shown in Figure 6.8(a). But because the firm in (b) has lower fixed costs, its *ATC* curve is lower at every level of output than the *ATC* curve in (a). At the profit-maximizing price of \$45 per household, the firm in Figure 6.8(b) earns an economic profit of \$10 per household, for a total economic profit of \$700,000 per month.

RECAP ↑

PROFIT MAXIMIZATION FOR THE MONOPOLIST

Both the perfectly competitive firm and the monopolist maximize profit by choosing the output level at which marginal revenue equals marginal cost. But whereas marginal revenue equals the market price for the perfectly competitive firm, it is always less than the market price for the monopolist. A monopolist will earn an economic profit only if price exceeds average total cost at the profit-maximizing level of output.

WHY THE INVISIBLE HAND BREAKS DOWN UNDER MONOPOLY

In our discussion of equilibrium in perfectly competitive markets in Chapter 5, *Efficiency, Exchange, and the Invisible Hand in Action*, we saw conditions under which the self-serving pursuits of consumers and firms were consistent with the broader interests of society as a whole. Let's explore whether the same conclusion holds true for the case of imperfectly competitive firms.

Consider the monopolist in Figures 6.6 and 6.7. Is this firm's profit-maximizing output level efficient from society's point of view? For any given level of output, the corresponding price on the demand curve indicates the amount buyers would be willing to pay for an additional unit of output. When the monopolist is producing 8 units per week, the marginal benefit to society of an additional unit of output is thus \$4 (see Figure 6.7). And since the marginal cost of an additional unit at that output level is only \$2 (again, see Figure 6.7), society would gain a net benefit of \$2 per unit if the monopolist were to expand production by 1 unit above the profit-maximizing level. Because this economic surplus is not realized, the profit-maximizing monopolist is socially inefficient.

Recall that the existence of inefficiency means that the economic pie is smaller than it might be. If that is so, why doesn't the monopolist simply expand production? The answer is that the monopolist would gladly do so, if only there were some way to maintain the price of existing units and cut the price of only the extra units. As a practical matter, however, that is not always possible.

Now, let's look at this situation from a different angle. For the market served by this monopolist, what is the socially efficient level of output?

At any output level, the cost to society of an additional unit of output is the same as the cost to the monopolist, namely, the amount shown on the monopolist's marginal cost curve. The marginal benefit to society (not to the monopolist) of an extra unit of output is simply the amount people are willing to pay for it, which is the amount shown on the monopolist's demand curve. To achieve social efficiency, the monopolist should expand production until the marginal benefit to society equals the marginal cost, which in this case occurs at a level of 12 units per week. (Again, see Figure 6.7.) Social efficiency is thus achieved at the output level at which the market demand curve intersects the monopolist's marginal cost curve.

For a monopolist, profit maximization occurs when marginal cost equals marginal revenue. Since the monopolist's marginal revenue is always less than price, the monopolist's profit-maximizing output level is always below the socially efficient level. Under perfect competition, by contrast, profit maximization occurs when marginal cost equals the market price—the same criterion that must be satisfied for social efficiency. This difference explains why the invisible hand of the market is less evident in monopoly markets than in perfectly competitive markets.

If perfect competition is socially efficient and monopoly is not, why isn't monopoly against the law? Congress has, in fact, tried to limit the extent of monopoly through antitrust laws. But even the most enthusiastic proponents of those laws recognize the limited usefulness of the legislative approach since the alternatives to monopoly often entail problems of their own.

Suppose, for example, that a monopoly results from a patent that prevents all but one firm from manufacturing some highly valued product. Would society be better off without patents? Probably not because eliminating such protection would discourage innovation. Virtually all successful industrial nations grant some form of patent protection, which gives firms a chance to recover the research and development costs without which new products would seldom reach the market.

Or suppose that the market in question is a natural monopoly—one that, because of economies of scale, is most cheaply served by a single firm. Would society do better to require this market to be served by many small firms, each with significantly higher average costs of production? Such a requirement would merely replace one form of inefficiency with another.

In short, we live in an imperfect world. Monopoly is socially inefficient, and that, needless to say, is bad. But the alternatives to monopoly aren't perfect either.

RECAP **WHY THE INVISIBLE HAND BREAKS DOWN UNDER MONOPOLY**

The monopolist maximizes profit at the output level for which marginal revenue equals marginal cost. Because its profit-maximizing price exceeds marginal revenue, and hence also marginal cost, the benefit to society of the last unit produced (the market price) must be greater than the cost of the last unit produced (the marginal cost). So the output level for an industry served by a profit-maximizing monopolist is smaller than the socially optimal level of output.

USING DISCOUNTS TO EXPAND THE MARKET

The source of inefficiency in monopoly markets is the fact that the benefit to the monopolist of expanding output is less than the corresponding benefit to society. From the monopolist's point of view, the price reduction the firm must grant existing buyers to expand output is a loss. But from the point of view of those buyers, each dollar of price reduction is a gain—one dollar more in their pockets.

Note the tension in this situation, which is similar to the tension that exists in all other situations in which the economic pie is smaller than it might otherwise be. When the economic pie grows larger, everyone can have a larger slice. To say that monopoly is inefficient means that steps could be taken to make some people better off without harming others. If people have a healthy regard for their own self-interest, why doesn't someone take those steps? Why, for example, doesn't the monopolist from the earlier examples sell 8 units of output at a price of \$4, and then once those buyers are out the door, cut the price for more price-sensitive buyers?

Price Discrimination Defined

Sometimes the monopolist does precisely that. Charging different buyers different prices for the same good or service is a practice known as **price discrimination**. Examples of price discrimination include seniors' and children's discounts on movie tickets, supersaver discounts on air travel, and rebate coupons on retail merchandise.

Attempts at price discrimination seem to work effectively in some markets, but not in others. Buyers are not stupid, after all; if the monopolist periodically offered a 50 percent discount on the \$8 list price, those who were paying \$8 might anticipate the next price cut and postpone their purchases to take advantage of it. In some markets, however, buyers may not know, or simply may not take the trouble to find out, how the price they pay compares to the prices paid by other buyers. Alternatively, the monopolist may be in a position to prevent some groups from buying at the discount prices made available to others. In such cases, the monopolist can price-discriminate effectively.

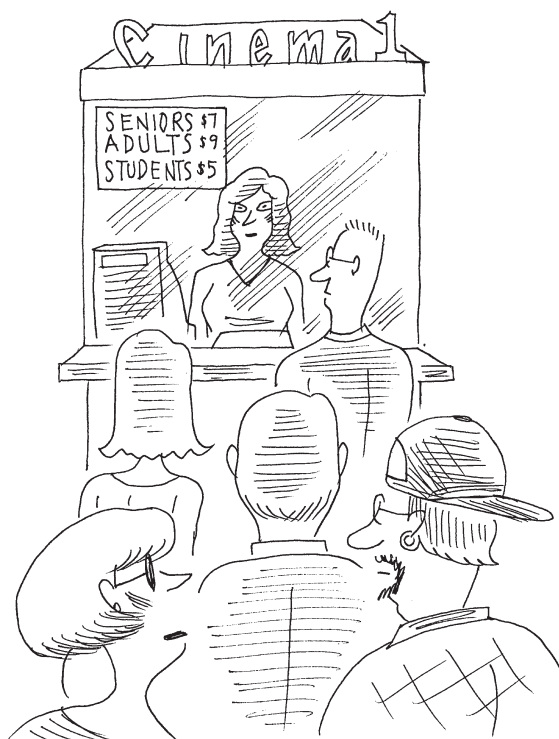
price discrimination the practice of charging different buyers different prices for essentially the same good or service

The Economic Naturalist 6.2



Why do many movie theaters offer discount tickets to students?

Whenever a firm offers a discount, the goal is to target that discount to buyers who would not purchase the product without it. People with low incomes generally have lower reservation prices for movie tickets than people with high incomes. Because students generally have lower disposable incomes than working adults, theater owners can expand their audiences by charging lower prices to students than to adults. Student discounts are one practical way of doing so. Offering student discounts also entails no risk of some people buying the product at a low price and then reselling it to others at a higher price.



Why do students pay lower ticket prices at many movie theaters?

How Price Discrimination Affects Output

In the following examples, we'll see how the ability to price-discriminate affects the monopolist's profit-maximizing level of output. First we'll consider a baseline case in which the monopolist must charge the same price to every buyer.

EXAMPLE 6.4 Profit Maximization and Opportunity Cost**How many manuscripts should Preeti edit?**

Preeti supplements her income as a teaching assistant by editing term papers for undergraduates. There are eight students per week for whom she might edit, each with a reservation price as given in the following table.

Student	Reservation price
<i>A</i>	\$40
<i>B</i>	38
<i>C</i>	36
<i>D</i>	34
<i>E</i>	32
<i>F</i>	30
<i>G</i>	28
<i>H</i>	26

Preeti is a profit maximizer. If the opportunity cost of her time to edit each paper is \$29 and she must charge the same price to each student, how many papers should she edit? How much economic profit will she make? How much accounting profit?

Table 6.5 summarizes Preeti's total and marginal revenue at various output levels. To generate the amounts in column 3, we simply multiplied the corresponding reservation price by the number of students whose reservation prices were at

TABLE 6.5
Total and Marginal Revenue from Editing

(1) Student	(2) Reservation price (\$ per paper)	(3) Total revenue (\$ per week)	(4) Marginal revenue (\$ per paper)
			40
<i>A</i>	40	40	36
<i>B</i>	38	76	32
<i>C</i>	36	108	28
<i>D</i>	34	136	24
<i>E</i>	32	160	20
<i>F</i>	30	180	16
<i>G</i>	28	196	12
<i>H</i>	26	208	

least that high. For example, to edit 4 papers per week (for students *A*, *B*, *C*, and *D*), Preeti must charge a price no higher than *D*'s reservation price (\$34). So her total revenue when she edits 4 papers per week is $(4)(\$34) = \136 per week. Preeti should keep expanding the number of students she serves as long as her marginal revenue exceeds the opportunity cost of her time. Marginal revenue, or the difference in total revenue that results from adding another student, is shown in column 4 of Table 6.5.

Note that if Preeti were editing 2 papers per week, her marginal revenue from editing a third paper would be \$32. Since that amount exceeds her \$29 opportunity cost, she should take on the third paper. But since the marginal revenue of taking on a fourth paper would be only \$28, Preeti should stop at 3 papers per week. The total opportunity cost of the time required to edit the 3 papers is $(3)(\$29) = \87 , so Preeti's economic profit is $\$108 - \$87 = \$21$ per week. Since Preeti incurs no explicit costs, her accounting profit will be \$108 per week.

EXAMPLE 6.5 Social Efficiency

What is the socially efficient number of papers for Preeti to edit?

Again, suppose that Preeti's opportunity cost of editing is \$29 per paper and that she could edit as many as 8 papers per week for students whose reservation prices are again as listed in the following table.

Student	Reservation price
<i>A</i>	\$40
<i>B</i>	38
<i>C</i>	36
<i>D</i>	34
<i>E</i>	32
<i>F</i>	30
<i>G</i>	28
<i>H</i>	26

What is the socially efficient number of papers for Preeti to edit? If she must charge the same price to each student, what will her economic and accounting profits be if she edits the socially efficient number of papers?

Students *A* to *F* are willing to pay more than Preeti's opportunity cost, so serving these students is socially efficient. But students *G* and *H* are unwilling to pay at least \$29 for Preeti's services. The socially efficient outcome, therefore, is for Preeti to edit 6 papers per week. To attract that number, she must charge a price no higher than \$30 per paper. Her total revenue will be $(6)(\$30) = \180 per week, slightly more than her total opportunity cost of $(6)(\$29) = \174 per week. Her economic profit will thus be only \$6 per week. Again, because Preeti incurs no explicit costs, her accounting profit will be the same as her total revenue, \$180 per week.

EXAMPLE 6.6 Price Discrimination**If Preeti can price-discriminate, how many papers should she edit?**

Suppose Preeti is a shrewd judge of human nature. After a moment's conversation with a student, she can discern that student's reservation price. The reservation prices of her potential customers are again as given in the following table. If Preeti confronts the same market as before, but can charge students their respective reservation prices, how many papers should she edit, and how much economic and accounting profit will she make?

Student	Reservation price
A	\$40
B	38
C	36
D	34
E	32
F	30
G	28
H	26

Preeti will edit papers for students A to F and charge each exactly his or her reservation price. Because students G and H have reservation prices below \$29, Preeti will not edit their papers. Preeti's total revenue will be $\$40 + \$38 + \$36 + \$34 + \$32 + \$30 = \$210$ per week, which is also her accounting profit. Her total opportunity cost of editing 6 papers is $(6)(\$29) = \174 per week, so her economic profit will be $\$210 - \$174 = \$36$ per week, \$30 per week more than when she edited 6 papers but was constrained to charge each customer the same price.

perfectly discriminating monopolist a firm that charges each buyer exactly his or her reservation price

A monopolist who can charge each buyer exactly his or her reservation price is called a **perfectly discriminating monopolist**. Notice that, when Preeti was discriminating among customers in this way, her profit-maximizing level of output was exactly the same as the socially efficient level of output: 6 papers per week. With a perfectly discriminating monopoly, there is no loss of efficiency. All buyers who are willing to pay a price high enough to cover marginal cost will be served.

Note that although total economic surplus is maximized by a perfectly discriminating monopolist, consumers would have little reason to celebrate if they found themselves dealing with such a firm. After all, consumer surplus is exactly zero for the perfectly discriminating monopolist. In this instance, total economic surplus and producer surplus are one and the same.

In practice, of course, perfect price discrimination can never occur because no seller knows each and every buyer's precise reservation price. But even if some sellers did know, practical difficulties would stand in the way of their charging a separate price to each buyer. For example, in many markets the seller could not prevent buyers who bought at low prices from reselling to other buyers at higher prices, capturing some of the seller's business in the process. Despite these difficulties, price discrimination is widespread. But it is generally *imperfect price discrimination*—that is, price discrimination in which at least some buyers are charged less than their reservation prices.

The Hurdle Method of Price Discrimination

The profit-maximizing seller's goal is to charge each buyer the highest price that buyer is willing to pay. Two primary obstacles prevent sellers from achieving this goal. First, sellers don't know exactly how much each buyer is willing to pay. And

second, they need some means of excluding those who are willing to pay a high price from buying at a low price. These are formidable problems, which no seller can hope to solve completely.

One common method by which sellers achieve a crude solution to both problems is to require buyers to overcome some obstacle to be eligible for a discount price. This method is called the **hurdle method of price discrimination**. For example, the seller might sell a product at a standard list price and offer a rebate to any buyer who takes the trouble to mail in a rebate coupon.

The hurdle method solves both of the seller's problems, provided that buyers with low reservation prices are more willing than others to jump the hurdle. Because a decision to jump the hurdle must satisfy the Cost-Benefit Principle, such a link seems to exist. As noted earlier, buyers with low incomes are more likely than others to have low reservation prices (at least in the case of normal goods). Because of the low opportunity cost of their time, they are more likely than others to take the trouble to send in rebate coupons. Rebate coupons thus target a discount toward those buyers whose reservation prices are low and who therefore might not buy the product otherwise.

A **perfect hurdle** is one that separates buyers precisely according to their reservation prices, and in the process imposes no cost on those who jump the hurdle. With a perfect hurdle, the highest reservation price among buyers who jump the hurdle will be lower than the lowest reservation price among buyers who choose not to jump the hurdle. In practice, perfect hurdles do not exist. Some buyers will always jump the hurdle, even though their reservation prices are high. And hurdles will always exclude at least some buyers with low reservation prices. Even so, many commonly used hurdles do a remarkably good job of targeting discounts to buyers with low reservation prices. In the example that follows, we will assume for convenience that the seller is using a perfect hurdle.

hurdle method of price discrimination the practice by which a seller offers a discount to all buyers who overcome some obstacle

perfect hurdle a threshold that completely segregates buyers whose reservation prices lie above it from others whose reservation prices lie below it, imposing no cost on those who jump the hurdle

EXAMPLE 6.7 Perfect Hurdle

How much should Preeti charge for editing if she uses a perfect hurdle?

Suppose Preeti again has the opportunity to edit as many as 8 papers per week for the students whose reservation prices are as given in the following table. This time she can offer a rebate coupon that gives a discount to any student who takes the trouble to mail it back to her. Suppose further that students whose reservation prices are at least \$36 never mail in the rebate coupons, while those whose reservation prices are below \$36 always do so.

Student	Reservation price
A	\$40
B	38
C	36
D	34
E	32
F	30
G	28
H	26

If Preeti's opportunity cost of editing each paper is again \$29, what should her list price be, and what amount should she offer as a rebate? Will her economic profit be larger or smaller than when she lacked the discount option?

The rebate coupon allows Preeti to divide her original market into two submarkets in which she can charge two different prices. The first submarket consists of students *A*, *B*, and *C*, whose reservation prices are at least \$36 and who therefore will not bother to mail in a rebate coupon. The second submarket consists of students *D* through *H*, whose lower reservation prices indicate a willingness to use rebate coupons.

In each submarket, Preeti must charge the same price to every buyer, just like an ordinary monopolist. She should therefore keep expanding output in each submarket as long as marginal revenue in that market exceeds her marginal cost. The relevant data for the two submarkets are displayed in Table 6.6.

On the basis of the entries in column 4 for the list price submarket, we see that Preeti should serve all three students (*A*, *B*, and *C*) since marginal revenue for each exceeds \$29. Her profit-maximizing price in the list price submarket is \$36, the highest price she can charge in that market and still sell her services to students *A*, *B*, and *C*. For the discount price submarket, marginal revenue exceeds \$29 only for the first two students (*D* and *E*). So the profit-maximizing price in this submarket is \$32, the highest price Preeti can charge and still sell her services to *D* and *E*. (A discount price of \$32 means that students who mail in the coupon will receive a rebate of \$4 on the \$36 list price.)

Note that the rebate offer enables Preeti to serve a total of five students per week, compared to only three without the offer. Preeti's combined total revenue for the two markets is $(3)(\$36) + 2(\$32) = \$172$ per week. Since her opportunity cost is \$29 per paper, or a total of $(5)(\$29) = \145 per week, her economic profit is $\$172$ per week $- \$145$ per week = \$27 per week, \$6 more than when she edited 3 papers and did not offer the rebate.

TABLE 6.6
Price Discrimination with a Perfect Hurdle

(1) Student	(2) Reservation price (\$ per paper)	(3) Total revenue (\$ per week)	(4) Marginal revenue (\$ per paper)
List Price Submarket			
<i>A</i>	40	40	40
<i>B</i>	38	76	36
<i>C</i>	36	108	32
Discount Price Submarket			
<i>D</i>	34	34	34
<i>E</i>	32	64	30
<i>F</i>	30	90	26
<i>G</i>	28	112	22
<i>H</i>	26	130	18

SELF-TEST 6.4

In Example 6.7, how much should Preeti charge in each submarket if she knows that only those students whose reservation prices are below \$34 will use rebate coupons?

Is Price Discrimination a Bad Thing?

We are so conditioned to think of discrimination as bad that we may be tempted to conclude that price discrimination must run counter to the public interest. In Example 6.7, however, both consumer surplus and producer surplus were actually enhanced by the monopolist's use of the hurdle method of price discrimination. To show this, let's compare consumer and producer surplus when Preeti employs the hurdle method to the corresponding values when she charges the same price to all buyers.

When Preeti had to charge the same price to every customer, she edited only the papers of students *A*, *B*, and *C*, each of whom paid a price of \$36. We can tell at a glance that the total surplus must be larger under the hurdle method because not only are students *A*, *B*, and *C* served at the same price (\$36), but also students *D* and *E* are now served at a price of \$32.

To confirm this intuition, we can calculate the exact amount of the surplus. For any student who hires Preeti to edit her paper, consumer surplus is the difference between her reservation price and the price actually paid. In both the single price and discount price examples, student *A*'s consumer surplus is thus $\$40 - \$36 = \$4$; student *B*'s consumer surplus is $\$38 - \$36 = \$2$; and student *C*'s consumer surplus is $\$36 - \$36 = 0$. Total consumer surplus in the list price submarket is thus $\$4 + \$2 = \$6$ per week, which is the same as total consumer surplus in the original situation. But now the discount price submarket generates additional consumer surplus. Specifically, student *D* receives \$2 per week of consumer surplus since this student's reservation price of \$34 is \$2 more than the discount price of \$32. So total consumer surplus is now $\$6 + \$2 = \$8$ per week, or \$2 per week more than before.

Preeti's producer surplus also increases under the hurdle method. For each paper she edits, her producer surplus is the price she charges minus her reservation price (\$29). In the single-price case, Preeti's surplus was $(3)(\$36 - \$29) = \$21$ per week. When she offers a rebate coupon, she earns the same producer surplus as before from students *A*, *B*, and *C* and an additional $(2)(\$32 - \$29) = \$6$ per week from students *D* and *E*. Total producer surplus with the discount is thus $\$21 + \$6 = \$27$ per week. Adding that amount to the total consumer surplus of \$8 per week, we get a total economic surplus of \$35 per week with the rebate coupons, \$8 per week more than without the rebate.

Note, however, that even with the rebate, the final outcome is not socially efficient because Preeti does not serve student *F*, even though this student's reservation price of \$30 exceeds her opportunity cost of \$29. Although the hurdle method is not perfectly efficient, it's still more efficient than charging a single price to all buyers.

Examples of Price Discrimination

Once you grasp the principle behind the hurdle method of price discrimination, you'll begin to see examples of it all around you. Next time you visit a grocery, hardware, or appliance store, for instance, notice how many different product promotions include cash rebates. Temporary sales are another illustration of the hurdle method. Most of the time, retailers sell their merchandise at the "regular" price but periodically offer special sales at a significant discount. The hurdle in this instance is taking the trouble to find out when and where the sales occur and then making a purchase during that period. This technique works because buyers who care most about price

(mainly, those with low reservation prices) are more likely to monitor advertisements carefully and buy only during sale periods.

To give another example, book publishers typically launch a new book in hard-cover at a price from \$20 to \$30, and a year later they bring out a paperback edition priced between \$5 and \$15. In this instance, the hurdle involves having to wait the extra year and accepting a slight reduction in the quality of the finished product. People who are strongly concerned about price end up waiting for the paperback edition, while those with high reservation prices usually spring for the hardback.

Or take the example of automobile producers, who typically offer several different models with different trim and accessories. Although GM's actual cost of producing a Cadillac may be only \$2,000 more than its cost of producing a Chevrolet, the Cadillac's selling price may be \$10,000 to \$15,000 higher than the Chevrolet's. Buyers with low reservation prices purchase the Chevrolet, while those with high reservation prices are more likely to choose the Cadillac.

Commercial air carriers have perfected the hurdle method to an extent matched by almost no other seller. Their supersaver fares are often less than half their regular coach fares. To be eligible for these discounts, travelers must purchase their tickets 7 to 21 days in advance and their journey must include a Saturday night stayover. Vacation travelers can more easily satisfy these restrictions than business travelers, whose schedules often change at the last moment and whose trips seldom involve Saturday stayovers. And—no surprise—the business traveler's reservation price tends to be much higher than the vacation traveler's.

Many sellers employ not just one hurdle but several by offering deeper discounts to buyers who jump successively more difficult hurdles. For example, movie studios initially release their major films to first-run theaters at premium prices and then, several months later, on video-on-demand platforms like iTunes. Still later they make the films available through subscription television providers like HBO and Netflix, and finally permit them to be shown on network television. Each successive hurdle involves waiting a little longer and, in the case of the televised versions, accepting lower quality. These hurdles are remarkably effective in segregating moviegoers according to their reservation prices.

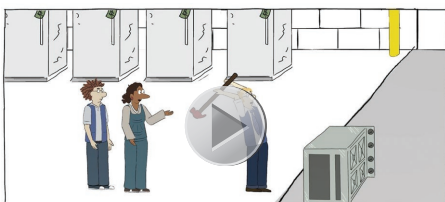
Recall that the efficiency loss from single-price monopoly occurs because, to the monopolist, the benefit of expanding output is smaller than the benefit to society as a whole. The hurdle method of price discrimination reduces this loss by giving the monopolist a practical means of cutting prices for price-sensitive buyers only. In general, the more finely the monopolist can partition a market using the hurdle method, the smaller the efficiency loss. Hurdles are not perfect, however, and some degree of efficiency will inevitably be lost.



The Economic Naturalist 6.3



► Visit your instructor's Connect course and access your eBook to view this video.



Why might an appliance retailer hammer dents into the sides of its stoves and refrigerators?

Why might an appliance retailer instruct its clerks to hammer dents into the sides of its stoves and refrigerators?

The Sears "Scratch 'n' Dent Sale" is another example of how retailers use quality differentials to segregate buyers according to their reservation prices. Many Sears stores hold an annual sale in which they display appliances with minor scratches and blemishes in the parking lot at deep discounts. People who don't care much about price are unlikely to turn out for these events, but those with very low reservation prices often get up early to be first in line. Indeed, these sales have proven so popular that it might even be in a retailer's interest to put dents in some of its sale items deliberately.

RECAP ↑**USING DISCOUNTS TO EXPAND THE MARKET**

A price-discriminating monopolist is one who charges different prices to different buyers for essentially the same good or service. A common method of price discrimination is the hurdle method, which involves granting a discount to buyers who jump over a hurdle such as mailing in a rebate coupon. An effective hurdle is one that is more easily cleared by buyers with low reservation prices than by buyers with high reservation prices. Such a hurdle enables the monopolist to expand output and thereby reduce the deadweight loss from monopoly pricing.

SUMMARY

- Our concern in this chapter was the conduct and performance of the imperfectly competitive firm, a firm that has at least some latitude to set its own price. Economists often distinguish among three different types of imperfectly competitive firms: the pure monopolist, the lone seller of a product in a given market; the oligopolist, one of only a few sellers of a given product; and the monopolistic competitor, one of a relatively large number of firms that sell similar though slightly differentiated products. (LO1)
- Although advanced courses in economics devote much attention to differences in behavior among these three types of firms, our focus was on the common feature that differentiates them from perfectly competitive firms. Whereas the perfectly competitive firm faces an infinitely elastic demand curve for its product, the imperfectly competitive firm faces a downward-sloping demand curve. For convenience, we use the term *monopolist* to refer to any of the three types of imperfectly competitive firms. (LO1)
- Monopolists are sometimes said to enjoy market power, a term that refers to their power to set the price of their product. Market power stems from exclusive control over important inputs, from economies of scale, from patents and government licenses or franchises, and from network economies. The most important and enduring of these five sources of market power are economies of scale and network economies. (LO2)
- Research, design, engineering, and other fixed costs account for an increasingly large share of all costs required to bring products successfully to market. For products with large fixed costs, marginal cost is lower, often substantially, than average total cost, and average total cost declines, often sharply, as output grows. This cost pattern explains why many industries are dominated by either a single firm or a small number of firms. (LO3)
- Unlike the perfectly competitive firm, for which marginal revenue exactly equals market price, the monopolist realizes a marginal revenue that is always less than its price. This shortfall reflects the fact that to sell more output, the monopolist must cut the price not only to additional buyers but to existing buyers as well. For the monopolist with a straight-line demand curve, the marginal revenue curve has the same vertical intercept and a horizontal intercept that is half as large as the intercept for the demand curve. (LO4)
- Whereas the perfectly competitive firm maximizes profit by producing at the level at which marginal cost equals the market price, the monopolist maximizes profit by equating marginal cost with marginal revenue. Since a monopolist's marginal revenue is less than its price, this results in an output level that is best for the monopolist but smaller than the level that would be best for society as a whole. At the profit-maximizing level of output, the benefit of an extra unit of output (the market price) is greater than its cost (the marginal cost). At the socially efficient level of output, where the monopolist's marginal cost curve intersects the demand curve, the benefit and cost of an extra unit are the same. (LO4, LO5)
- Both the monopolist and its potential customers can do better if the monopolist can grant discounts to price-sensitive buyers. The extreme example is the perfectly discriminating monopolist, who charges each buyer exactly his or her reservation price. Such producers are socially efficient because they sell to every buyer whose reservation price is at least as high as the marginal cost. (LO6)

KEY TERMS

average total cost
 constant returns to scale
 hurdle method of price discrimination
 imperfectly competitive firm (or price setter)

increasing returns to scale (or economies of scale)
 marginal revenue
 market power
 monopolistic competition
 natural monopoly

oligopoly
 perfect hurdle
 perfectly discriminating monopolist
 price discrimination
 pure monopoly

REVIEW QUESTIONS

1. What important characteristic do all three types of imperfectly competitive firms share? (LO1)
2. True or false: A firm with market power can sell whatever quantity it wishes at whatever price it chooses. (LO2)
3. Why do most successful industrial societies offer patents and copyright protection, even though these protections enable sellers to charge higher prices? (LO2)
4. Why is marginal revenue always less than price for a monopolist but equal to price for a perfectly competitive firm? (LO4)
5. True or false: Because a natural monopolist charges a price greater than marginal cost, it necessarily earns a positive economic profit. (LO4)

PROBLEMS



1. State whether the following statements are true or false, and explain why. (LO1, LO2)
 - a. In a perfectly competitive industry, the industry demand curve is horizontal, whereas for a monopoly it is downward-sloping.
 - b. Perfectly competitive firms have no control over the price they charge for their product.
 - c. For a natural monopoly, average cost declines as the number of units produced increases over the relevant output range.
2. Two car manufacturers, Nissan and Honda, have fixed costs of \$1 billion and marginal costs of \$10,000 per car. If Nissan produces 50,000 cars per year and Honda produces 200,000, calculate the average production cost for each company. On the basis of these costs, which company's market share do you think will grow in relative terms? (LO3)
3. A single-price, profit-maximizing monopolist (LO4)
 - a. causes excess demand, or shortages, by selling too few units of a good or service.
 - b. chooses the output level at which marginal revenue begins to increase.
 - c. always charges a price above the marginal cost of production.
 - d. also maximizes marginal revenue.
 - e. none of the above statements is true.
4. If a monopolist could perfectly price-discriminate (LO4, LO5, LO6)
 - a. the marginal revenue curve and the demand curve would coincide.
 - b. the marginal revenue curve and the marginal cost curve would coincide.
 - c. every consumer would pay a different price.
 - d. marginal revenue would become negative at some output level.
 - e. the resulting pattern of exchange would still be socially inefficient.
5. Explain why price discrimination and the existence of slightly different variants of the same product tend to go hand in hand. Give an example from your own experience. (LO6)
6. TotsPoses Inc., a profit-maximizing business, is the only photography business in town that specializes in portraits of small children. George, who owns

and runs TotsPoses, expects to encounter an average of eight customers per day, each with a reservation price shown in the following table. Assume George has no fixed costs, and his cost of producing each portrait is \$12. (LO4, LO5, LO6)

Customer	Reservation price (\$ per photo)
1	50
2	46
3	42
4	38
5	34
6	30
7	26
8	22

- How much should charge if he must charge a single price to all customers? At this price, how many portraits will George produce each day? What will be his economic profit?
 - How much consumer surplus is generated each day at this price?
 - What is the socially efficient number of portraits?
 - George is very experienced in the business and knows the reservation price of each of his customers. If he is allowed to charge any price he likes to any consumer, how many portraits will he produce each day and how much economic profit will he earn?
 - In this case, how much consumer surplus is generated each day?
7. Refer back to Problem 6 and answer the following questions. (LO6)
- Suppose George is permitted to charge two prices. He knows that customers with a reservation price above \$30 never bother with coupons, whereas those with a reservation price of \$30 or less always use them. At what level should George set the list price of a portrait? At what level should he set the discount price? How many photo portraits will he sell at each price?

- In this case, what is George's economic profit and how much consumer surplus is generated each day?

8. Jada is a second-grader who sells lemonade on a street corner in your neighborhood. Each cup of lemonade costs Jada 20 cents to produce; she has no fixed costs. The reservation prices for the 10 people who walk by Jada's lemonade stand each day are listed in the following table.

Person	Reservation price
A	\$1.00
B	\$0.90
C	\$0.80
D	\$0.70
E	\$0.60
F	\$0.50
G	\$0.40
H	\$0.30
I	\$0.20
J	\$0.10

Jada knows the distribution of reservation prices (that is, she knows that one person is willing to pay \$1, another \$0.90, and so on), but she does not know any specific individual's reservation price. (LO4, LO5, LO6)

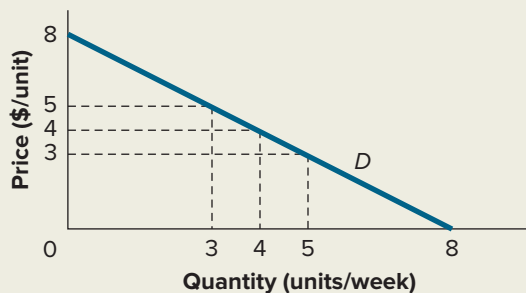
- Calculate the marginal revenue of selling an additional cup of lemonade. (Start by figuring out the price Jada would charge if she produced only one cup of lemonade, and calculate the total revenue; then find the price Jada would charge if she sold two cups of lemonade; and so on.)
- What is Jada's profit-maximizing price?
- At that price, what are Jada's economic profit and total consumer surplus?
- What price should Jada charge if she wants to maximize total economic surplus?
- Now suppose Jada can tell the reservation price of each person. What price would she charge each person if she wanted to maximize profit? Compare her profit to the total surplus calculated in part d.

ANSWERS TO SELF-TESTS

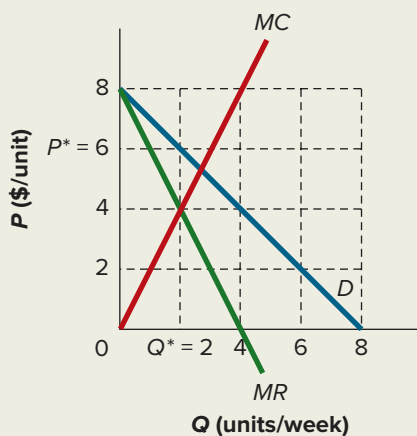
6.1 The relevant cost figures are shown in the following table, which shows that Sony's unit-cost advantage is now $\$50.20 - \$5.20 = \$45.00$. (LO3)

	Nintendo	Sony Computer Entertainment
Annual production	200,000	2,000,000
Fixed cost	\$10,000,000	\$10,000,000
Variable cost	\$40,000	\$400,000
Total cost	\$10,040,000	\$10,400,000
Average total cost per game	\$50.20	\$5.20

6.2 When the monopolist expands from 3 to 4 units per week, total revenue rises from \$15 to \$16 per week, which means that the marginal revenue from the sale of the fourth unit is only \$1 per week. When the monopolist expands from 4 to 5 units per week, total revenue drops from \$16 to \$15 per week, which means that the marginal revenue from the sale of the fifth unit is actually negative, or $-\$1$ per week. (LO4)

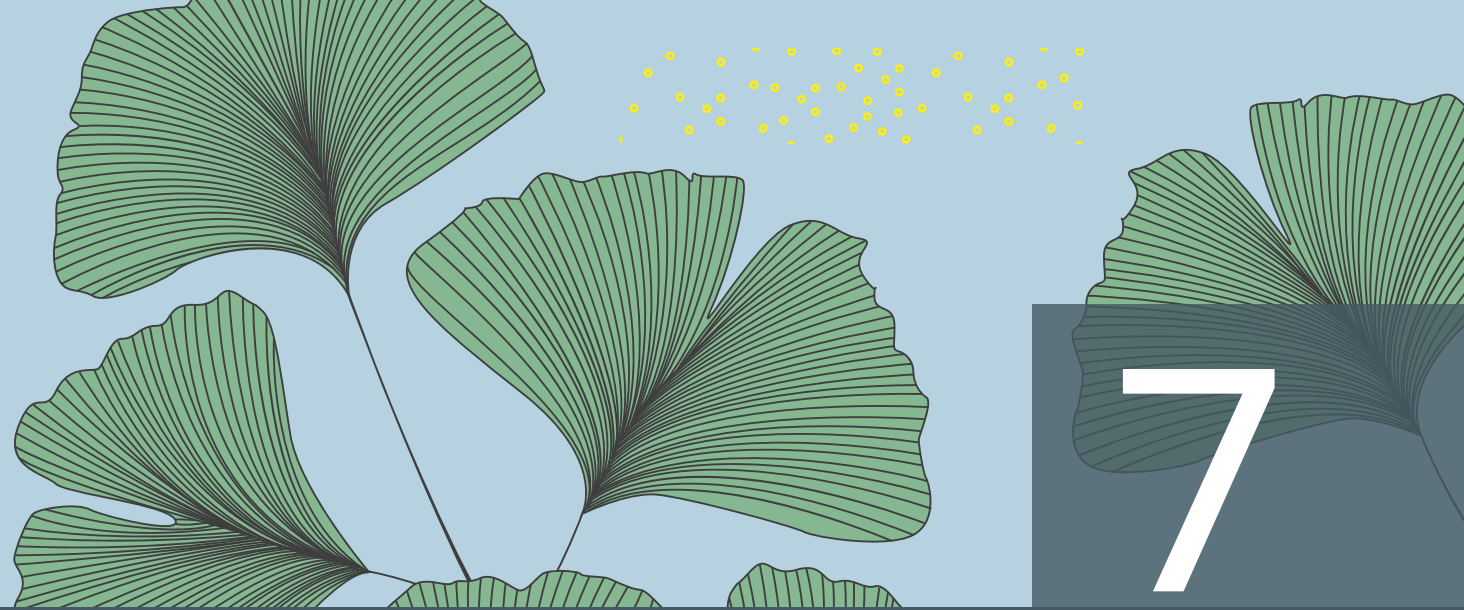


6.3 The profit-maximizing price and quantity are $P^* = \$6/\text{unit}$ and $Q^* = 2$ units/week. (LO4)



6.4 As column 4 in the following table shows, Preeti should again serve students A, B, and C in the list price submarket (at a price of \$36) and only student E in the discount submarket (at a price of \$32). (LO6)

(1) Student	(2) Reservation price (\$ per paper)	(3) Total revenue (\$ per week)	(4) Marginal revenue (\$ per paper)
List Price Submarket			
			40
A	40	40	
			36
B	38	76	
			32
C	36	108	
			28
D	34	136	
Discount Price Submarket			
			32
E	32	32	
			28
F	30	60	
			24
G	28	84	
			20
H	26	104	



Games and Strategic Behavior

At a Christmas Eve dinner party in 1997, actor Robert DeNiro pulled aside singer Tony Bennett for a moment. “Hey, Tony—there’s a film I want you in,” DeNiro said. He was referring to the project that became the 1999 Warner Brothers hit comedy *Analyze This*, in which the troubled head of a crime family, played by DeNiro, seeks the counsel of a psychotherapist, played by Billy Crystal. In the script, both the mob boss and his therapist are big fans of Bennett’s music.

Bennett heard nothing further about the project for almost a year. Then his son and financial manager, Danny Bennett, got a phone call from Warner Brothers, in which the studio offered Tony \$15,000 to sing “Got the World on a String” in the movie’s final scene. As Danny described the conversation, “They made a fatal mistake. They told me they had already shot the film. So I’m like: ‘Hey, they shot the whole film around Tony being the end gag and they’re offering me \$15,000?’”¹

Warner Brothers wound up paying \$200,000 for Bennett’s performance.

In business negotiations, as in life, timing can be everything. If executives at Warner Brothers had thought the problem through carefully, they would have negotiated with Bennett *before* shooting the movie. At that point, Bennett would have realized that the script could be rewritten if he asked too high a fee. By waiting, studio executives left themselves with no attractive option other than to pay Bennett’s price.

The payoff to many actions depends not only on the actions themselves, but also on when they’re taken and how they relate to actions taken by others. In previous chapters, economic decision makers confronted an environment that was essentially fixed. This chapter will focus on cases in which people must consider the effect of their behavior on others. For example, an imperfectly competitive firm will want to weigh the likely responses of rivals when deciding whether to cut prices or to increase its advertising budget. Interdependencies of this sort are the rule rather than the exception in economic and social life. To make sense of the world we live in, then, we must take these interdependencies into account.

¹As quoted by Geraldine Fabrikant, “Talking Money with Tony Bennett,” *The New York Times*, May 2, 1999, Money & Business, p. 1.

LEARNING OBJECTIVES

After reading this chapter, you should be able to:

- LO1** Describe the three basic elements of a game, how the possible pay-offs are summarized, and the effect of dominant and dominated strategy choices.
- LO2** Identify and explain the prisoner’s dilemma and how it applies to real-world situations.
- LO3** Explain games in which the timing of players’ choices matters.
- LO4** Discuss strategies that enable players to solve commitment problems through material or psychological incentives.

Our focus in Chapter 6, *Monopoly, Oligopoly, and Monopolistic Competition*, was on the pure monopolist. In this chapter, we'll explore how a few simple principles from the theory of games can help us better understand the behavior of oligopolists and monopolistic competitors—the two types of imperfectly competitive firms for which strategic interdependencies are most important. Along the way, we'll also see how the same principles enable us to answer a variety of interesting questions drawn from everyday social interaction.

USING GAME THEORY TO ANALYZE STRATEGIC DECISIONS

In chess, tennis, or any other game, the payoff to a given move depends on what your opponent does in response. In choosing your move, therefore, you must anticipate your opponent's responses, how you might respond, and what further moves your own response might elicit. Economists and other behavioral scientists have devised the theory of games to analyze situations in which the payoffs to different actors depend on the actions their opponents take.

The Three Elements of a Game

basic elements of a game
the players, the strategies available to each player, and the payoffs each player receives for each possible combination of strategies

A game has three **basic elements**: the players, the list of possible actions (or strategies) available to each player, and the payoffs the players receive for each possible combination of strategies. We'll use a series of examples to illustrate how these elements combine to form the basis of a theory of behavior.

The first example focuses on an important strategic decision confronting two oligopolists who produce an undifferentiated product and must decide how much to spend on advertising.

EXAMPLE 7.1 The Cost of Advertising

Should United Airlines spend more money on advertising?

Suppose that United Airlines and American Airlines are the only air carriers that serve the Chicago–St. Louis market. Each currently earns an economic profit of \$6,000 per flight on this route. If United increases its advertising spending in this market by \$1,000 per flight, and American spends no more on advertising than it does now, United's profit will rise to \$8,000 per flight and American's will fall to \$2,000. If both spend \$1,000 more on advertising, each will earn an economic profit of \$5,500 per flight. These payoffs are symmetric so that if United spends the same amount on advertising while American increases its spending by \$1,000, United's economic profit will fall to \$2,000 per flight and American's will rise to \$8,000. The payoff structure is also common knowledge—that is, each company knows what the relevant payoffs will be for both parties under each of the possible combinations of choices. If each must decide independently whether to increase spending on advertising, what should United do?

payoff matrix a table that describes the payoffs in a game for each possible combination of strategies

Think of this situation as a game. What are its three elements? The players are the two airlines. Each airline must choose one of two strategies: to raise ad spending by \$1,000 or leave it the same. The payoffs are the economic profits that correspond to the four possible scenarios resulting from their choices. One way to summarize the relevant information about this game is to display the players, strategies, and payoffs in the form of a simple table called a **payoff matrix** (see Table 7.1).

TABLE 7.1
The Payoff Matrix for an Advertising Game

		American's Choices	
		Raise ad spending	Leave ad spending the same
United's Choices	Raise ad spending	\$5,500 for United \$5,500 for American	\$8,000 for United \$2,000 for American
	Leave ad spending the same	\$2,000 for United \$8,000 for American	\$6,000 for United \$6,000 for American

Both airlines do better if both leave ad spending the same than if both raise spending. Yet if one holds spending the same, the other always does better to raise spending.

Confronted with the payoff matrix in Table 7.1, what should United Airlines do? The essence of strategic thinking is to begin by looking at the situation from the other party's point of view. Suppose United assumes that American will raise its spending on advertising (the left column in Table 7.1). In that case, United's best bet would be to follow suit (the top row in Table 7.1). Why is the top row United's best response when American chooses the left column? United's economic profits, given in the upper-left cell of Table 7.1, will be \$5,500, compared to only \$2,000 if it keeps spending the same (see the lower-left cell).

Alternatively, suppose United assumes that American will keep ad spending the same (that is, that American will choose the right column in Table 7.1). In that case, United would still do better to increase spending because it would earn \$8,000 (the upper-right cell), compared to only \$6,000 if it keeps spending the same (the lower-right cell). In this particular game, no matter which strategy American chooses, United will earn a higher economic profit by increasing its spending on advertising. And since this game is perfectly symmetric, a similar conclusion holds for American: No matter which strategy United chooses, American will do better by increasing its spending on ads.

When one player has a strategy that yields a higher payoff no matter which choice the other player makes, that player is said to have a **dominant strategy**. Not all games involve dominant strategies, but both players in this game have one, and that is to increase spending on ads. For both players, to leave ad spending the same is a **dominated strategy**—one that leads to a lower payoff than an alternative choice, regardless of the other player's choice.

Notice, however, that when each player chooses the dominant strategy, the resulting payoffs are smaller than if each had left spending unchanged. When United and American increase their spending on ads, each earns only \$5,500 in economic profits, compared to the \$6,000 each would have earned without the increase.

dominant strategy one that yields a higher payoff no matter what the other players in a game choose

dominated strategy any other strategy available to a player who has a dominant strategy

Nash equilibrium any combination of strategy choices in which each player's choice is his or her best choice, given the other players' choices

Nash Equilibrium

A game is said to be in equilibrium if each player's strategy is the best he or she can choose, given the other players' choices. This definition of equilibrium is sometimes called a **Nash equilibrium**, after the mathematician John Nash, who developed the concept in the early 1950s. Nash was awarded the Nobel Prize in Economics in 1994 for his contributions to game theory.² When a game is in equilibrium, no player has any incentive to deviate from his current strategy.

If each player in a game has a dominant strategy, as in Example 7.1, equilibrium occurs when each player follows that strategy. But even in games in which not every player has a dominant strategy, we can often identify an equilibrium outcome. Consider, for instance, the following variation on the advertising game as illustrated in Example 7.2.

EXAMPLE 7.2 Nash Equilibrium

Should American Airlines spend more money on advertising?

Suppose United Airlines and American Airlines are the only carriers that serve the Chicago–St. Louis market. Their payoff matrix for advertising decisions is shown in Table 7.2. Does United have a dominant strategy? Does American? If each firm does the best it can, given the incentives facing the other, what will be the outcome of this game?

In this game, no matter what United does, American will do better to raise its ad spending, so raising the advertising budget is a dominant strategy for American. United, however, does not have a dominant strategy. If American raises its spending, United will do better to leave its spending unchanged; but if American does not raise spending, United will do better to spend more. Even though United doesn't have a dominant strategy, we can examine the players' incentives

TABLE 7.2
Equilibrium When One Player Lacks a Dominant Strategy

		American's Choices	
		Raise ad spending	Leave ad spending the same
United's Choices	Raise ad spending	\$3,000 for United \$4,000 for American	\$8,000 for United \$3,000 for American
	Leave ad spending the same	\$4,000 for United \$5,000 for American	\$5,000 for United \$2,000 for American

In this game, United lacks a dominant strategy, but American's dominant strategy is to raise its ad spending. Because United can predict that American will choose the left column, United will do best to leave its ad spending the same. Equilibrium occurs in the lower-left cell.

²Nash was awarded the Nobel Prize in Economics in 1994 for his contributions to game theory. His life was also the subject of the Academy Award-winning film *A Beautiful Mind*.

to predict what is likely to happen in this game. United's managers are assumed to know what the payoff matrix is, so they can predict that American will spend more on ads since that is American's dominant strategy. Thus the best strategy for United, given the prediction that American will spend more on ads, is to keep its own spending unchanged. If both players do the best they can, taking account of the incentives each faces, this game will end in the lower-left cell of the payoff matrix: American will raise its spending on ads and United will not.

Note that the choices corresponding to the lower-left cell in Table 7.2 satisfy the definition of a Nash equilibrium. If United found itself in that cell, its alternative would be to raise its ad spending, a move that would reduce its payoff from \$4,000 to \$3,000. So United has no incentive to abandon the lower-left cell. Similarly, if American found itself in the lower-left cell of Table 7.2, its alternative would be to leave ad spending the same, a move that would reduce its payoff from \$5,000 to \$2,000. So American also has no incentive to abandon the lower-left cell. The lower-left cell of Table 7.2 is a Nash equilibrium—a combination of strategies for which each player's choice is the best available option, given the choice made by the other player.

SELF-TEST 7.1

What should United and American do if their payoff matrix is modified as follows?

		American	
		Raise ad spending	Leave spending the same
United	Raise ad spending	\$3,000 for United \$8,000 for American	\$4,000 for United \$5,000 for American
	Leave spending the same	\$8,000 for United \$4,000 for American	\$5,000 for United \$2,000 for American

RECAP ↑

USING GAME THEORY TO ANALYZE STRATEGIC DECISIONS

The three elements of any game are the players, the list of strategies from which they can choose, and the payoffs to each combination of strategies. This information can be summarized in a payoff matrix.

Equilibrium in a game occurs when each player's strategy choice yields the highest payoff available, given the strategies chosen by other players. Such a combination of strategies is called a Nash equilibrium.

THE PRISONER’S DILEMMA

prisoner’s dilemma a game in which each player has a dominant strategy, and when each plays it, the resulting payoffs are smaller than if each had played a dominated strategy

The first advertising example we discussed above belongs to an important class of games called the **prisoner’s dilemma**. In the prisoner’s dilemma, when each player chooses a dominant strategy, the result is unattractive to the group of players as a whole.

The Original Prisoner’s Dilemma

The next example recounts the original scenario from which the prisoner’s dilemma drew its name.

EXAMPLE 7.3 Prisoner’s Dilemma

Should the prisoners confess?

Two prisoners, Horace and Jasper, are being held in separate cells for a serious crime that they did in fact commit. The prosecutor, however, has enough hard evidence to convict them of only a minor offense, for which the penalty is a year in jail. Each prisoner is told that if one confesses while the other remains silent, the confessor will be cleared of the crime, and the other will spend 20 years in prison. If both confess, they will get an intermediate sentence of 5 years. These payoffs are summarized in Table 7.3. The two prisoners are not allowed to communicate with one another. Do they have a dominant strategy? If so, what is it?

In this game, the dominant strategy for each prisoner is to confess. No matter what Jasper does, Horace will get a lighter sentence by speaking out. If Jasper confesses, Horace will get 5 years (upper-left cell) instead of 20 (lower-left cell). If Jasper remains silent, Horace will go free (upper-right cell) instead of spending a year in jail (lower-right cell). Because the payoffs are perfectly symmetric, Jasper will also do better to confess, no matter what Horace does. The difficulty is that when each follows his dominant strategy and confesses, both will do worse than if each had shown restraint. When both confess, they each get 5 years (upper-left cell), instead of the 1 year they would have gotten by remaining silent (lower-right cell). Hence the name of this game, the prisoner’s dilemma.

TABLE 7.3
The Payoff Matrix for a Prisoner’s Dilemma

		Jasper	
		Confess	Remain silent
Horace	Confess	5 years for each	0 years for Horace 20 years for Jasper
	Remain silent	20 years for Horace 0 years for Jasper	1 year for each

The payoffs describe the lengths of prison sentences the two will receive under different combinations of choices.

SELF-TEST 7.2

GM and Chrysler must both decide whether to invest in a new process. Games 1 and 2 below show how their profits (in millions of dollars) depend on the decisions they might make. Which of these games is a prisoner's dilemma?

		Game 1 Chrysler				Game 2 Chrysler	
		Don't invest	Invest			Don't invest	Invest
GM	Don't invest	10 for each	4 for GM 12 for Chrysler	GM	Don't invest	4 for GM 12 for Chrysler	5 for each
	Invest	12 for GM 4 for Chrysler	5 for each		Invest	10 for each	12 for GM 4 for Chrysler

The prisoner's dilemma is one of the most powerful metaphors in all of human behavioral science. Countless social and economic interactions have payoff structures analogous to the one confronted by the two prisoners. Some of those interactions occur between only two players, as in the examples just discussed; many others involve larger groups. Games of the latter sort are called *multiplayer prisoner's dilemmas*. But regardless of the number of players involved, the common thread is one of conflict between the narrow self-interest of individuals and the broader interests of larger communities.

The Economics of Cartels

A **cartel** is any coalition of firms that conspires to restrict production for the purpose of earning an economic profit. As we will see in the next example, the problem confronting oligopolists who are trying to form a cartel is a classic illustration of the prisoner's dilemma.

cartel a coalition of firms that agree to restrict output for the purpose of earning an economic profit

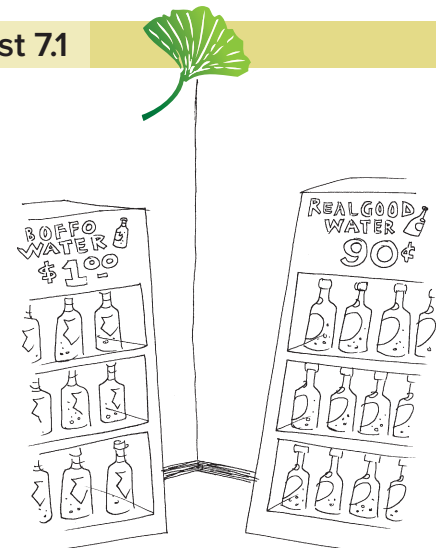
The Economic Naturalist 7.1

Why are cartel agreements notoriously unstable?

Consider a market for bottled water served by two oligopolists, Aquapure and Mountain Spring. Each firm can draw water free of charge from a mineral spring located on its own land. Customers supply their own bottles. Rather than compete with one another, the two firms decide to join together by selling water at the price a profit-maximizing pure monopolist would charge. Under their agreement (which constitutes a cartel), each firm would produce and sell half the quantity of water demanded by the market at the monopoly price (see Figure 7.1). The agreement isn't legally enforceable, however, which means that each firm has the option of charging less than the agreed price. If one firm sells water for less than the other firm, it will capture the entire quantity demanded by the market at the lower price.

Why is this agreement likely to collapse?

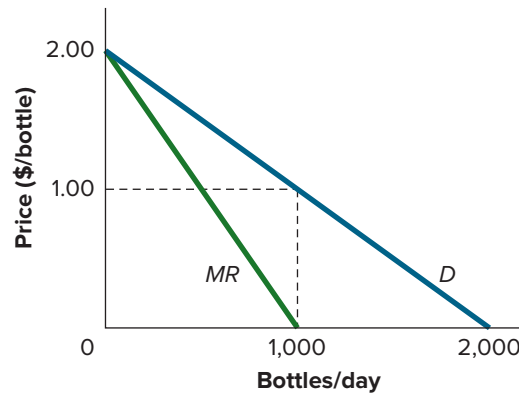
Since the marginal cost of mineral water is zero, the profit-maximizing quantity for a monopolist with the demand curve shown in Figure 7.1 is



Why is it so difficult for companies to enforce agreements against price cutting?

FIGURE 7.1**The Market Demand for Mineral Water.**

Faced with the demand curve shown, a monopolist with zero marginal cost would produce 1,000 bottles per day (the quantity at which marginal revenue equals zero) and sell them at a price of \$1 per bottle.



1,000 bottles per day, the quantity for which marginal revenue equals marginal cost. At that quantity, the monopoly price is \$1 per bottle. If the firms abide by their agreement, each will sell half the market total, or 500 bottles per day, at a price of \$1 per bottle, for an economic profit of \$500 per day.

But suppose Aquapure reduced its price to 90 cents per bottle. By underselling Mountain Spring, it would capture the entire quantity demanded by the market, which, as shown in Figure 7.2, is 1,100 bottles per day. Aquapure's economic profit would rise from \$500 per day to $(\$0.90 \text{ per bottle})(1,100 \text{ bottles per day}) = \990 per day—almost twice as much as before. In the process, Mountain Spring's economic profit would fall from \$500 per day to zero. Rather than see its economic profit disappear, Mountain Spring would match Aquapure's price cut, recapturing its original 50 percent share of the market. But when each firm charges \$0.90 per bottle and sells 550 bottles per day, each earns an economic profit of $(\$0.90 \text{ per bottle})(550 \text{ bottles per day}) = \495 per day, or \$5 less per day than before.

Suppose we view the cartel agreement as an economic game in which the two available strategies are to sell for \$1 per bottle or to sell for \$0.90 per bottle. The payoffs are the economic profits that result from these strategies. Table 7.4 shows the payoff matrix for this game. Each firm's dominant strategy is to sell at the lower price, yet in following that strategy, each earns a lower profit than if each had sold at the higher price.

The game does not end with both firms charging \$0.90 per bottle. Each firm knows that if it cuts the price a little further, it can recapture the entire market and,

FIGURE 7.2**The Temptation to Violate a Cartel Agreement.**

By cutting its price from \$1 per bottle to \$0.90 per bottle, Aquapure can sell the entire market quantity demanded at that price, 1,100 bottles per day, rather than half the monopoly quantity of 1,000 bottles per day.

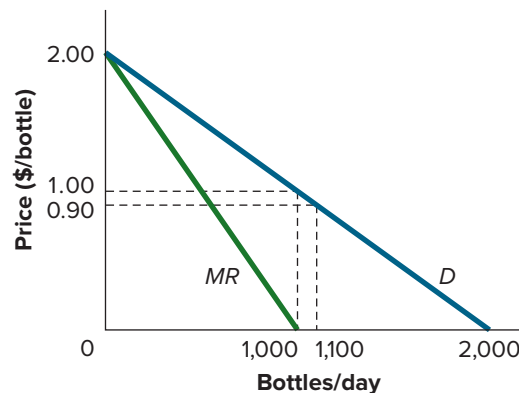


TABLE 7.4
The Payoff Matrix for a Cartel Agreement

		Mountain Spring	
		Charge \$1/bottle	Charge \$0.90/bottle
Aquapure	Charge \$1/bottle	\$500/day for each	\$0 for Aquapure \$990/day for Mt. Spring
	Charge \$0.90/bottle	\$990/day Aquapure \$0 for Mt. Spring	\$495/day for each

The dominant strategy for each firm is to charge \$0.90 per bottle, or \$0.10 per bottle less than called for by the cartel agreement. Hence the notorious instability of cartel agreements.

in the process, earn a substantially higher economic profit. At every step, the rival firm will match any price cut, until the price falls all the way to the marginal cost—in this example, zero.

Cartel agreements confront participants with the economic incentives inherent in the prisoner's dilemma, which explains why such agreements have historically been so unstable. Usually a cartel involves not just two firms, but several; an arrangement that can make retaliation against price cutters extremely difficult. In many cases, discovering which parties have broken the agreement is difficult. For example, the Organization of the Petroleum Exporting Countries (OPEC), a cartel of oil producers formed in the 1970s to restrict oil production, has no practical way to prevent member countries from secretly pumping oil offshore in the dead of night.

Tit-for-Tat and the Repeated Prisoner's Dilemma

When all players cooperate in a prisoner's dilemma, each gets a higher payoff than when all defect. So people who confront prisoner's dilemmas will be on the lookout for ways to create incentives for mutual cooperation. What they need is some way to penalize players who defect. When players interact with one another only once, this turns out to be difficult. But when they expect to interact repeatedly, new possibilities emerge.

A **repeated prisoner's dilemma** is a standard prisoner's dilemma that confronts the same players not just once but many times. Experimental research on repeated prisoner's dilemmas in the 1960s identified a simple strategy that proves remarkably effective at limiting defection. The strategy is called **tit-for-tat**, and here's how it works: The first time you interact with someone, you cooperate. In each subsequent interaction, you simply do what that person did in the previous interaction. Thus, if your partner defected on your first interaction, you'd then defect on your next interaction with her. If she then cooperates, your move next time will be to cooperate as well.

On the basis of elaborate computer simulations, University of Michigan political scientist Robert Axelrod showed that tit-for-tat was a remarkably effective strategy, even when pitted against a host of ingenious counterstrategies that had been designed for the explicit purpose of trying to exploit it. The success of tit-for-tat requires a reasonably stable set of players, each of whom can remember what other players have done in

repeated prisoner's dilemma

a standard prisoner's dilemma that confronts the same players repeatedly

tit-for-tat a strategy for the repeated prisoner's dilemma in which players cooperate on the first move and then mimic their partner's last move on each successive move

previous interactions. It also requires that players have a significant stake in what happens in the future, for it is the fear of retaliation that deters people from defecting.

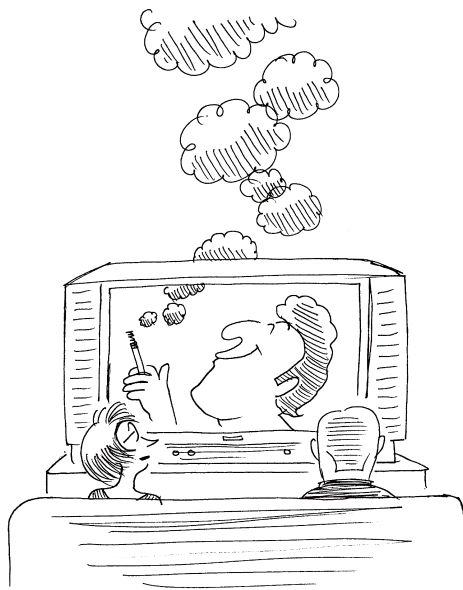
Because rival firms in the same industry interact with one another repeatedly, it might seem that the tit-for-tat strategy would ensure widespread collusion to raise prices. And yet, as noted earlier, cartel agreements are notoriously unsuccessful. One difficulty is that tit-for-tat's effectiveness depends on there being only two players in the game. In competitive and monopolistically competitive industries, there are generally many firms, and even in oligopolies there are often several. When there are more than two firms and one defects now, how do the cooperators selectively punish the defector later? By cutting price? That will penalize everyone, not just the defector. Even if there are only two firms in an industry, these firms realize that other firms may enter their industry. So the would-be cartel members have to worry not only about each other, but also about the entire list of firms that might decide to compete with them. Each firm may see this as a hopeless task and decide to defect now, hoping to reap at least some economic profit in the short run. What seems clear, in any event, is that the practical problems involved in implementing tit-for-tat have made it difficult to hold cartel agreements together for long.



The Economic Naturalist 7.2

How did Congress unwittingly solve the television advertising dilemma confronting cigarette producers?

In 1970, Congress enacted a law making cigarette advertising on television illegal after January 1, 1971. As evidenced by the steadily declining proportion of Americans who smoke, this law seems to have achieved its stated purpose of protecting citizens against a proven health hazard. But the law also had an unintended effect, which was to increase the economic profit of cigarette makers, at least in the short run. In the year before the law's passage, manufacturers spent more than \$300 million on advertising—about \$60 million more than they spent during the year after the law was enacted. Much of the saving in advertising expenditures in 1971 was reflected in higher cigarette profits at year-end. But if eliminating television advertising made companies more profitable, why didn't the manufacturers eliminate the ads on their own?



Why were cigarette manufacturers happy when Congress made it illegal for them to advertise on television?

When an imperfectly competitive firm advertises its product, its demand curve shifts rightward, for two reasons. First, people who have never used that type of product learn about it, and some buy it. Second, people who consume a different brand of the product may switch brands. The first effect boosts sales industrywide; the second merely redistributes existing sales among brands.

Although advertising produces both effects in the cigarette industry, its primary effect is brand switching. Thus, the decision of whether to advertise confronts the individual firm with a prisoner's dilemma. Table 7.5 shows the payoffs facing a pair of cigarette producers trying to decide whether to advertise. If both firms advertise on TV (upper-left cell), each earns a profit of only \$10 million per year, compared to a profit of \$20 million per year for each if neither advertises (lower-right cell). Clearly, both will benefit if neither advertises.

Yet note the powerful incentive that confronts each firm. RJR (the R.J. Reynolds Tobacco Company) sees that if Philip Morris doesn't advertise, RJR can earn higher profits by advertising (\$35 million per year) than by not advertising (\$20 million per year). RJR also sees that if Philip Morris does advertise, RJR will again earn more by advertising

TABLE 7.5
Profits from Cigarette Advertising as a Prisoner's Dilemma

		Philip Morris	
		Advertise on TV	Don't advertise on TV
RJR	Advertise on TV	\$10 million/year for each	\$35 million/year for RJR \$5 million/year for Philip Morris
	Don't advertise on TV	\$5 million/year for RJR \$35 million/year for Philip Morris	\$20 million/year for each

In many industries, the primary effect of advertising is to encourage consumers to switch brands. In such industries, the dominant strategy is to advertise heavily (upper-left cell), even though firms as a group would do better by not advertising (lower-right cell).

(\$10 million per year) than by not advertising (\$5 million per year). Thus, RJR's dominant strategy is to advertise. And because the payoffs are symmetric, Philip Morris's dominant strategy is also to advertise. So when each firm behaves rationally from its own point of view, the two together do worse than if they had both shown restraint. The congressional ad ban forced cigarette manufacturers to do what they could not have accomplished on their own.

As the following Economic Naturalist 7.3 example makes clear, understanding the prisoner's dilemma can help the economic naturalist make sense of human behavior not only in the world of business, but also in other domains of life as well.

The Economic Naturalist 7.3



Why do people shout at parties?

Whenever large numbers of people gather for conversation in a closed space, the ambient noise level rises sharply. After attending such gatherings, people often complain of sore throats and hoarse voices. If everyone spoke at a normal volume at parties, the overall noise level would be lower, and people would hear just as well. So why do people shout?

The problem involves the difference between individual incentives and group incentives. Suppose everyone starts by speaking at a normal level. But because of the crowded conditions, conversation partners have difficulty hearing one another, even when no one is shouting. The natural solution, from the point of the individual, is to simply raise one's voice a bit. But that is also the natural solution for everyone else. And when everyone speaks more loudly, the ambient noise level rises so that no one hears any better than before.

No matter what others do, the individual will do better by speaking more loudly. Doing so is a dominant strategy for everyone, in fact. Yet when everyone follows

McGraw Hill connect

► Visit your instructor's Connect course and access your eBook to view this video.



Why do people shout at parties?

the dominant strategy, the result is worse (no one can hear well) than if everyone had continued to speak normally. While shouting is wasteful, individuals acting alone have no better option. If anyone were to speak softly while others shout, that person wouldn't be heard. No one wants to go home with raw vocal cords, but people apparently prefer that cost to the alternative of not being heard at all.

RECAP ↑

THE PRISONER'S DILEMMA

The prisoner's dilemma is a game in which each player has a dominant strategy, and in which the payoff to each player when each chooses that strategy is smaller than if each had chosen a dominated strategy. Incentives analogous to those found in the prisoner's dilemma help explain a broad range of behavior in business and everyday life—among them excessive spending on advertising and cartel instability. The tit-for-tat strategy can help sustain cooperation in two-player repeated prisoner's dilemmas but tends to be ineffective in multiplayer repeated prisoner's dilemmas.

GAMES IN WHICH TIMING MATTERS

In the games discussed so far, players were assumed to choose their strategies simultaneously, and which player moved first didn't matter. For example, in the prisoner's dilemma, self-interested players would follow their dominant strategies even if they knew in advance what strategies their opponents had chosen. But in other situations, such as the negotiations between Warner Brothers and Tony Bennett described at the beginning of this chapter, timing is of the essence.

We begin with an example of a game whose outcome cannot be predicted if both players move simultaneously, but whose outcome is clear if one player has the opportunity to move before the other.

EXAMPLE 7.4

The Importance of Timing

Should Ford build a hybrid Mustang?

The Ford Mustang and the Chevrolet Camaro compete for a limited pool of domestic sports car enthusiasts. Each company knows that the other is considering whether to bring out a hybrid version of its car. If both companies bring out hybrids, each will earn \$60 million in profit. If neither brings out a hybrid, each company will earn \$50 million. If Chevrolet introduces a hybrid and Ford does not, Chevrolet will earn \$80 million and Ford will earn \$70 million. If Ford brings out a hybrid and Chevrolet does not, Ford will earn \$80 million and Chevrolet will earn \$70 million. Does either firm have a dominant strategy in this situation? What will happen in this game if Ford gets to choose first, with Chevrolet choosing after having seen Ford's choice?

When both companies must make their decisions simultaneously, the payoff matrix for the example looks like Table 7.6.

The logic of the profit figures in Table 7.6 is that although consumers generally like the idea of a hybrid sports car (hence the higher profits when both companies bring out hybrids than when neither does), the companies will have to compete more heavily with one another if both offer the same type of car

TABLE 7.6
The Advantage of Being Different

		Ford Mustang	
		Offer hybrid	Don't offer hybrid
Chevrolet Camaro	Offer hybrid	\$60 million for Chevrolet \$60 million for Ford	\$80 million for Chevrolet \$70 million for Ford
	Don't offer hybrid	\$70 million for Chevrolet \$80 million for Ford	\$50 million for Chevrolet \$50 million for Ford

Profits are higher when each company offers a different type of car than the other (upper-right and lower-left cells). Customers generally prefer hybrid cars (upper-left cell) to nonhybrids (lower-right cell).

(and hence the lower profits when both offer the same type of car than when each offers a different type).

In the payoff matrix in Table 7.6, neither company has a dominant strategy. The best outcome for Ford is to offer a hybrid Mustang while Chevrolet does not offer a hybrid Camaro (lower-left cell). The best outcome for Chevrolet is to offer a hybrid Camaro while Ford does not offer a hybrid Mustang (upper-right cell). Both the lower-left and upper-right cells are Nash equilibria of this game because if the companies found themselves in either of these cells, neither would unilaterally want to change its position. Thus, in the upper-right cell, Chevrolet wouldn't want to change (that cell is, after all, the best possible outcome for Chevrolet), and neither would Ford (since switching to a hybrid would reduce its profit from \$70 million to \$60 million). But without being told more, we simply cannot predict where the two companies will end up.

If one side can move before the other, however, the incentives for action become instantly clearer. For games in which timing matters, a **decision tree**, or **game tree**, is a more useful way of representing the payoffs than a traditional payoff matrix. This type of diagram describes the possible moves in the sequence in which they may occur, and lists the final payoffs for each possible combination of moves.

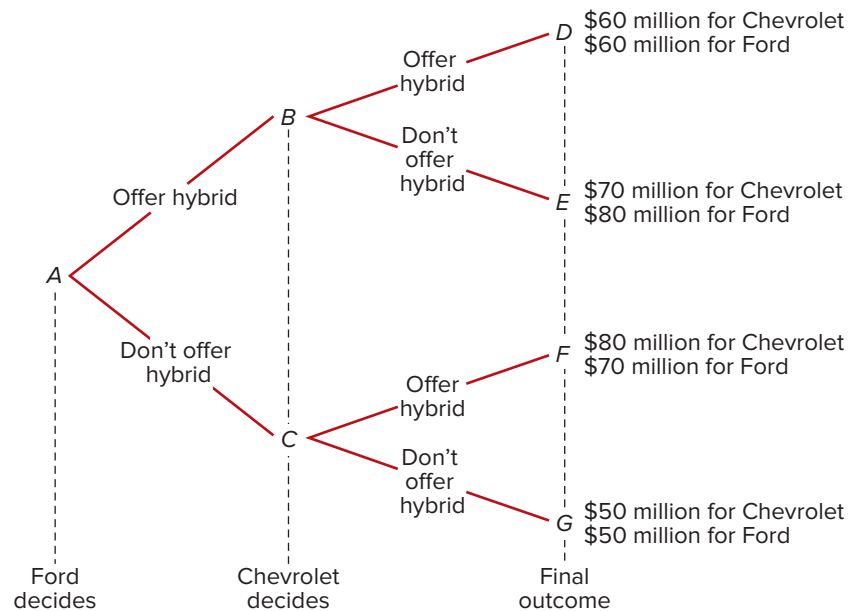
If Ford has the first move, the decision tree for the game is shown in Figure 7.3. At *A*, Ford begins the game by deciding whether to offer a hybrid. If it chooses to offer one, Chevrolet must then make its own choice at *B*. If Ford does not offer a hybrid, Chevrolet will make its choice at *C*. In either case, once Chevrolet makes its choice, the game is over.

In thinking strategically about this game, the key for Ford is to put itself in Chevrolet's shoes and imagine how Chevrolet would react to the various choices it might confront. In general, it will make sense for Ford to assume that Chevrolet will respond in a self-interested way—that is, by choosing the available option that offers the highest profit for Chevrolet. Ford knows that if it chooses to offer a hybrid, Chevy's best option at *B* will be not to offer a hybrid (since Chevy's profit is \$10 million higher at *E* than at *D*). Ford also knows that if it chooses not to offer a hybrid, Chevy's best option at *C* will be to offer one (since Chevy's profit is \$30 million higher at *F* than at *G*). Ford thus knows that if it offers a hybrid, it will end up at *E*, where it will earn \$80 million, whereas if it does not offer a hybrid, it will end

decision tree (or **game tree**)
a diagram that describes the possible moves in a game in sequence and lists the payoffs that correspond to each possible combination of moves

FIGURE 7.3**Decision Tree for Hybrid Example.**

This decision tree shows the possible moves and payoffs for the game in the hybrid example, in the sequence in which they may occur.



up at F , where it will earn only \$70 million. So when Ford has the first move in this game, its best strategy is to offer a hybrid. And Chevrolet then follows by choosing not to offer one.

Note that, in this game, the company that moves first has a considerable advantage. By moving first, Ford is able to earn \$80 million compared to Chevrolet's \$70 million. In general, being the first company to introduce a desirable product innovation is highly profitable. This undoubtedly explains the real-world battle between Ford and Chevrolet to introduce the first hybrid versions of their cars.

Credible Threats and Promises

credible threat a threat to take an action that is in the threatener's interest to carry out

Could Chevrolet have deterred Ford from offering a hybrid by threatening to offer a hybrid of its own, no matter what Ford did? The problem with this strategy is such a threat would not have been credible. In the language of game theory, a **credible threat** is one that will be in the threatener's interest to carry out when the time comes to act. People are likely to be skeptical of any threat if they know there will be no incentive to follow through when the time comes. The problem here is that Ford knows that it would not be in Chevrolet's interest to carry out its threat in the event that Ford offered a hybrid. After all, once Ford has already offered the hybrid, Chevy's best option is to offer a nonhybrid.

The concept of a credible threat figured prominently in the negotiations between Warner Brothers's managers and Tony Bennett over the matter of Mr. Bennett's fee for performing in *Analyze This*. Once most of the film had been shot, managers knew they couldn't threaten credibly to refuse Mr. Bennett's salary demand because at that point adapting the film to another singer would have been extremely costly. In contrast, a similar threat made before production of the movie had begun would have been credible.

credible promise a promise to take an action that is in the promiser's interest to keep

Just as in some games credible threats are impossible to make, in others **credible promises** are impossible. A credible promise is one that is in the interests of the promiser to keep when the time comes to act. In the following example, both players suffer because of the inability to make a credible promise.

EXAMPLE 7.5 A Credible Promise

Should the business owner open a remote office?

The owner of a thriving business wants to start up an office in a distant city. If she hires someone to manage the new office, she can afford to pay a weekly salary of \$1,000—a premium of \$500 over what the manager would otherwise be able to earn—and still earn a weekly economic profit of \$1,000 for herself. The owner's concern is that she won't be able to monitor the manager's behavior. The owner knows that by managing the remote office dishonestly, the manager can boost his take-home pay to \$1,500 while causing the owner an economic loss of \$500 per week. If the owner believes that all managers are selfish income-maximizers, will she open the new office?

The decision tree for the remote-office game is shown in Figure 7.4. At *A*, the managerial candidate promises to manage honestly, which brings the owner to *B*, where she must decide whether to open the new office. If she opens it, they reach *C*, where the manager must decide whether to manage honestly. If the manager's only goal is to make as much money as he can, he will manage dishonestly (bottom branch at *C*) because, that way, he will earn \$500 more than by managing honestly (top branch at *C*).

So if the owner opens the new office, she will end up with an economic loss of \$500. If she had not opened the office (bottom branch at *B*), she would have realized an economic profit of zero. Because zero is better than $-\$500$, the owner will choose not to open the remote office. In the end, the opportunity cost of the manager's inability to make a credible promise is \$1,500: the manager's forgone \$500 salary premium and the owner's forgone \$1,000 return.

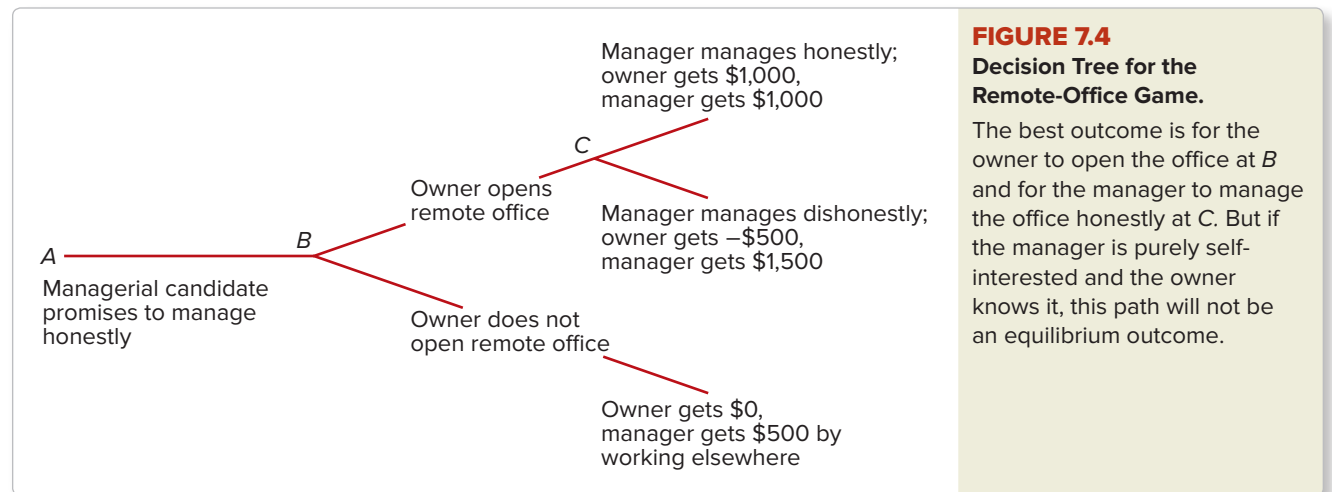


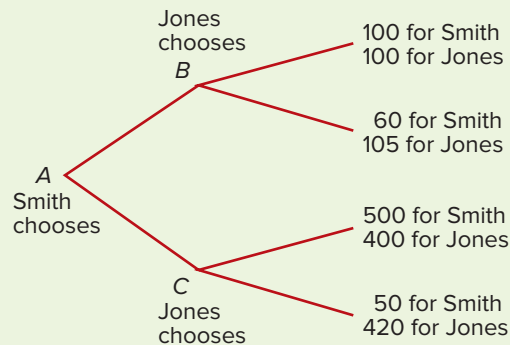
FIGURE 7.4
Decision Tree for the Remote-Office Game.

The best outcome is for the owner to open the office at *B* and for the manager to manage the office honestly at *C*. But if the manager is purely self-interested and the owner knows it, this path will not be an equilibrium outcome.

SELF-TEST 7.3

Smith and Jones are playing a game in which Smith has the first move at *A* in the decision tree shown below. Once Smith has chosen either the top or bottom branch at *A*, Jones, who can see what Smith has chosen, must choose the top or bottom branch at *B* or *C*. If the payoffs at the end of each branch are as

shown, what is the equilibrium outcome of this game? If before Smith chose, Jones could make a credible commitment to choose either the top or bottom branch when his turn came, what would he do?



Monopolistic Competition When Location Matters

In many sequential games, the player who gets to move first enjoys a strategic advantage. That was the case, for instance, in the decision of whether to produce a hybrid sports car in Example 7.4. In that example, the first mover did better because he was able to exploit the knowledge that both firms do better if each one's product is different from the other's rather than similar to it. But that won't always be true. When the feature that differentiates one seller's product from another's is temporal or spatial location, the firm with the last move in a game sometimes enjoys the upper hand, as The Economic Naturalist 7.4 illustrates.



The Economic Naturalist 7.4

Why do we often see convenience stores located on adjacent street corners?



► Visit your instructor's Connect course and access your eBook to view this video.

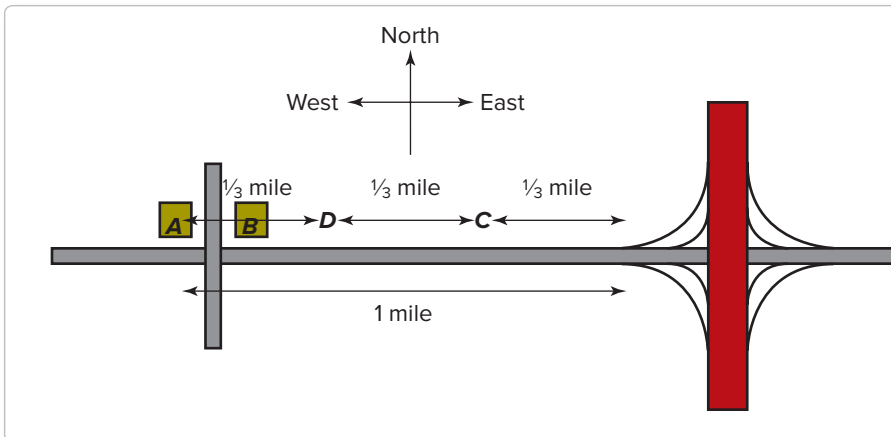


Why do we often see convenience stores located on adjacent street corners?

In many cities, it's common to see convenience stores located in clusters, followed by long stretches with no stores at all. If the stores were more spread out, almost all consumers would enjoy a shorter walk to the nearest convenience store. Why do stores tend to cluster in this fashion?

In Figure 7.5, suppose that when the convenience store located at A first opened, it was the closest store for the 1,200 shoppers who live in identical apartment houses evenly distributed along the road between A and the freeway one mile to the east.³ Those who live to the east of the freeway shop elsewhere because they cannot cross the freeway. Those who live to the west of the store at A shop either at A or at some other store still further to the west, whichever is closer. In this setting, why might a profit-maximizing entrepreneur planning to open a new store

³"Evenly distributed" means that the number of shoppers who live on any segment of the road between A and the freeway is exactly proportional to the length of that segment. For example, the number who live along a segment one-tenth of a mile in length would be $1/10 \times 1,200 = 120$.

**FIGURE 7.5****The Curious Tendency of Monopolistic Competitors to Cluster.**

As a group, consumers would enjoy a shorter walk if the store at *B* were instead located at *C*, or even at *D*. But a second store will attract more customers by locating at *B*.

between *A* and the freeway choose to locate at *B* rather than at some intermediate location such as *C*?

It turns out that a store located at *C* would, in fact, minimize the distance that shoppers living between *A* and the freeway would have to walk to reach the nearest store. If there were a store at *C*, no shopper on this stretch of road would have to walk more than $\frac{1}{3}$ of a mile to reach the nearest store. The 800 people who live between point *D* (which is halfway between *A* and *C*) and the freeway would shop at *C*, while the 400 who live between *D* and *A* would shop at *A*.

Despite the fact that *C* is the most attractive location for a new store from the perspective of consumers, it is not the most advantageous for the store's owner. The reason is that the owner's profit depends on how many people choose to shop at his store, not on how far they have to walk to get there. Given that consumers shop at the store closest to where they live, the best option from the entrepreneur's perspective is to locate his store at *B*, on the street corner just east of *A*. That way, his store will be closer to all 1,200 people who live between *A* and the freeway. It is this logic that often helps explain the clustering of convenience stores, gas stations, and other monopolistically competitive firms whose most important differentiating feature is geographic location.

The insight that helped answer the question posed in The Economic Naturalist 7.4 comes from economist Harold Hotelling.⁴ Hotelling employed this insight to explain why two hot dog vendors on a stretch of beach almost invariably locate next to one another midway between the endpoints of the beach.

For many oligopolistic or monopolistically competitive firms, an important dimension of product differentiation is location in time rather than in physical space. The timing of flight departures for different airlines in the New York–Los Angeles market is one example. The timing of film showings by different local movie theaters is another. In these cases, too, we often see product clustering. Thus, in the New York–Los Angeles market, both United and American have flights throughout the afternoon departing exactly on the hour. And in many local movie markets, the first evening showing starts at 7:15 p.m. in dozens of different theaters.

In other examples, the differentiating features that matter most might be said to describe the product's location in a more abstract "product space." With soft drinks,

⁴Harold Hotelling, "Stability and Competition," *Economic Journal* 39, no. 1 (1929), pp. 41–57.

for example, we might array different products according to their degrees of sweetness or carbonation. Here, too, it is common to see rival products that lie very close to one another, such as Coca-Cola and Pepsi. Clustering occurs in these cases for the reasons analogous to those discussed by Hotelling in his classic paper.

RECAP

GAMES IN WHICH TIMING MATTERS

The outcomes in many games depend on the timing of each player's move. For such games, the payoffs are best summarized by a decision tree rather than a payoff matrix. Sometimes the second mover does best to offer a product that differs markedly from existing products. Other times, the second mover does best to mimic existing products closely.

commitment problem a situation in which people cannot achieve their goals because of an inability to make credible threats or promises

commitment device a way of changing incentives so as to make otherwise empty threats or promises credible

COMMITMENT PROBLEMS

Games like the one in Self-Test 7.3, as well as the prisoner's dilemma, the cartel game, and the remote-office game, confront players with a **commitment problem**—a situation in which they have difficulty achieving the desired outcome because they cannot make credible threats or promises. If both players in the original prisoner's dilemma could make a binding promise to remain silent, both would be assured of a shorter sentence, hence the logic of the underworld code of *Omerta*, under which the family of anyone who provides evidence against a fellow mob member is killed. A similar logic explains the adoption of military arms control agreements, in which opponents sign an enforceable pledge to curtail weapons spending.

The commitment problem in the remote-office game could be solved if the managerial candidate could find some way of committing himself to manage honestly if hired. The candidate needs a **commitment device**—something that provides the candidate with an incentive to keep his promise.

Business owners are well aware of commitment problems in the workplace and have adopted a variety of commitment devices to solve them. Consider, for example, the problem confronting the owner of a restaurant. She wants her table staff to provide good service so that customers will enjoy their meals and come back in the future. Because good service is valuable to her, she would be willing to pay servers extra for it. For their part, servers would be willing to provide good service in return for the extra pay. The problem is that the owner cannot always monitor whether the servers do provide good service. Her concern is that having been paid extra for it, the servers may slack off when she isn't looking. Unless the owner can find some way to solve this problem, she will not pay extra; the servers will not provide good service; and she, they, and the diners will suffer. A better outcome for all concerned would be for the servers to find some way to commit themselves to good service.

Restaurateurs in many countries have tried to solve this commitment problem by encouraging diners to leave tips at the end of their meals. The attraction of this solution is that the diner is *always* in a good position to monitor service quality. The diner should be happy to reward good service with a generous tip because doing so will help ensure good service in the future. And the server has a strong incentive to provide good service because he knows that the size of his tip may depend on it.

The various commitment devices just discussed—the underworld code of *Omerta*, military arms control agreements, the tip for the server—all work because they change the incentives facing the decision makers. But as the next example illustrates, changing incentives in precisely the desired way is not always practical.

EXAMPLE 7.6 Changing Incentives

Will Sylvester leave a tip when dining on the road?

Sylvester has just finished a \$100 steak dinner at a restaurant that is 500 miles from where he lives. The server provided good service. If Sylvester cares only about himself, will he leave a tip?

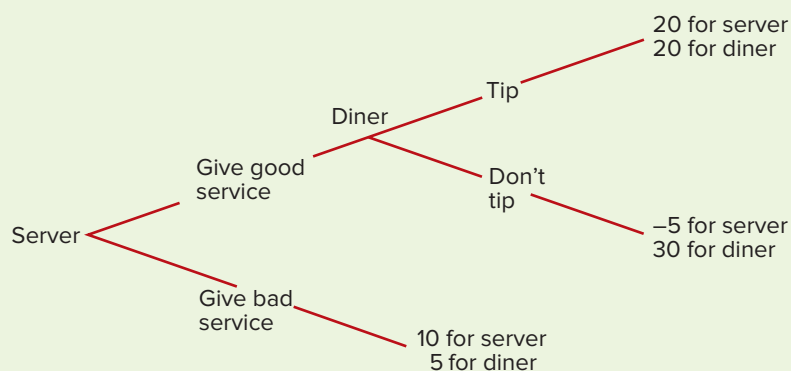
Once the server has provided good service, there is no way for her to take it back if the diner fails to leave a tip. In restaurants patronized by local diners, failure to tip is not a problem because the server can simply provide poor service the next time a nontipper comes in. But the server lacks that leverage with out-of-town diners. Having already received good service, Sylvester must choose between paying \$100 and paying \$120 for his meal. If he is an essentially selfish person, the former choice may be a compelling one.



Will leaving a tip at an out-of-town restaurant affect the quality of service you receive?

SELF-TEST 7.4

A traveler dines at a restaurant far from home. Both he and the server are rational and self-interested in the narrow sense. The server must first choose between providing good service and bad service, whereupon the diner must choose whether or not to leave a tip. The payoffs for their interaction are as summarized on the accompanying game tree. What is the most the diner would be willing to pay for the right to make a binding commitment (visible to the server) to leave a tip at the end of the meal in the event of having received good service?



Solving Commitment Problems with Psychological Incentives

In all the games we have discussed so far, players were assumed to care only about obtaining the best possible outcome for themselves. Thus, each player's goal was to get the highest monetary payoff, the shortest jail sentence, the best chance to be heard, and so on. The irony, in most of these games, is that players do not attain the best outcomes. Better outcomes can sometimes be achieved by altering the material incentives selfish players face, but not always.

If altering the relevant material incentives is not possible, commitment problems can sometimes be solved by altering people's psychological incentives. As the next example illustrates, in a society in which people are strongly conditioned to develop moral sentiments—feelings of guilt when they harm others, feelings of sympathy for their trading partners, feelings of outrage when they are treated unjustly—commitment problems arise less often than in more narrowly self-interested societies.

EXAMPLE 7.7

The Impact of Moral Sentiments

In a moral society, will the business owner open a remote office?

Consider again the owner of the thriving business who is trying to decide whether to open an office in a distant city. Suppose the society in which she lives is one in which all citizens have been strongly conditioned to behave honestly. Will she open the remote office?

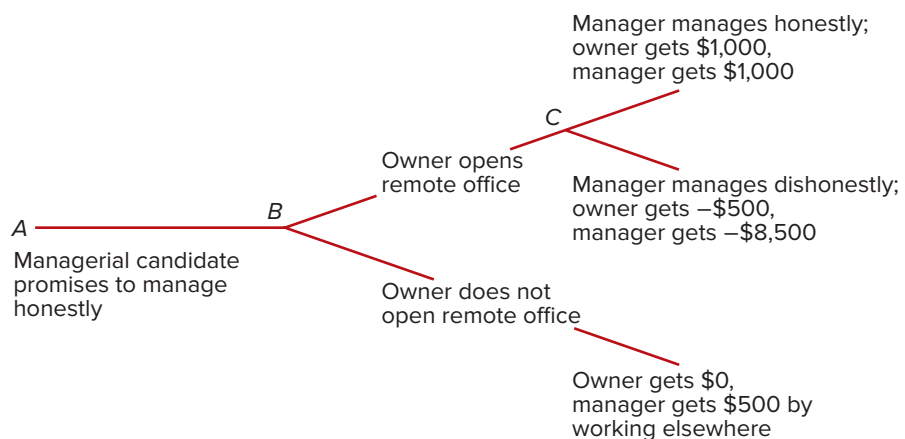
Suppose, for instance, that the managerial candidate would suffer guilt pangs if he embezzled money from the owner. Most people would be reluctant to assign a monetary value to guilty feelings. But for the sake of discussion, let's suppose that those feelings are so unpleasant that the manager would be willing to pay at least \$10,000 to avoid them. On this assumption, the manager's payoff if he manages dishonestly will be not \$1,500, but $\$1,500 - \$10,000 = -\$8,500$. The new decision tree is shown in Figure 7.6.

In this case, the best choice for the owner at *B* will be to open the remote office because she knows that at *C* the manager's best choice will be to manage honestly. The irony, of course, is that the honest manager in this example ends up richer than the selfish manager in the previous example, who earned only a normal salary.

FIGURE 7.6

The Remote-Office Game with an Honest Manager.

If the owner can identify a managerial candidate who would choose to manage honestly at *C*, she will hire that candidate at *B* and open the remote office.



Are People Fundamentally Selfish?

As Example 7.7 suggests, the assumption that people are self-interested in the narrow sense of the term does not always capture the full range of motives that govern choice in strategic settings. Think, for example, about the last time you had a meal at an out-of-town restaurant. Did you leave a tip? If so, your behavior was quite normal. Researchers have found that tipping rates in restaurants patronized mostly by out-of-town diners are essentially the same as in restaurants patronized mostly by local diners.

Indeed, there are many exceptions to the outcomes predicted on the basis of the assumption that people are self-interested in the narrowest sense of the term. People who have been treated unjustly often seek revenge even at ruinous cost to themselves. Every day, people walk away from profitable transactions whose terms they believe to be “unfair.” In these and countless other ways, people do not seem to be pursuing self-interest narrowly defined. And if motives beyond narrow self-interest are significant, we must take them into account in attempting to predict and explain human behavior.

Preferences as Solutions to Commitment Problems

Economists tend to view preferences as ends in themselves. Taking them as given, they calculate what actions will best serve those preferences. This approach to the study of behavior is widely used by other social scientists, and by game theorists, military strategists, philosophers, and others. In its standard form, it assumes purely self-interested preferences for present and future consumption goods of various sorts, leisure pursuits, and so on. Concerns about fairness, guilt, honor, sympathy, and the like typically play no role.

Yet such concerns clearly affect the choices people make in strategic interactions. Sympathy for one’s trading partner can make a businessperson trustworthy even when material incentives favor cheating. A sense of justice can prompt a person to incur the costs of retaliation, even when incurring those costs will not undo the original injury.

Preferences can clearly shape behavior in these ways; however, this alone does not solve commitment problems. The solution to such problems requires not only that a person *have* certain preferences, but also that others have some way of *discerning* them. Unless the business owner can identify the trustworthy employee, that employee cannot land a job whose pay is predicated on trust. And unless the predator can identify a potential victim whose character will motivate retaliation, that person is likely to become a victim.

From among those with whom we might engage in ventures requiring trust, can we identify reliable partners? If people could make *perfectly* accurate character judgments, they could always steer clear of dishonest persons. That people continue to be victimized at least occasionally by dishonest persons suggests that perfectly reliable character judgments are either impossible to make or prohibitively expensive.

Vigilance in the choice of trading partners is an essential element in solving (or avoiding) commitment problems, for if there is an advantage in being honest and being perceived as such, there is an even greater advantage in only *appearing* to be honest. After all, a liar who appears trustworthy will have better opportunities than one who glances about furtively, sweats profusely, and has difficulty making eye contact. Indeed, he will have the same opportunities as an honest person but will get higher payoffs because he will exploit them to the fullest.

In the end, the question of whether people can make reasonably accurate character judgments is an empirical one. Experimental studies have shown that even on the basis of brief encounters involving strangers, subjects are adept at predicting who will cooperate and who will defect in prisoner’s dilemma games. For example, in one experiment in which only 26 percent of subjects defected, the accuracy rate of predicted defections was more than 56 percent. One might expect that predictions regarding those we know well would be even more accurate.

Do you know someone who would return an envelope containing \$1,000 in cash to you if you lost it at a crowded concert? If so, then you accept the claim that personal character can help people solve commitment problems. As long as honest individuals can identify at least some others who are honest, and can interact selectively with them, honest individuals can prosper in a competitive environment.

RECAP **COMMITMENT PROBLEMS AND THE EFFECTS OF PSYCHOLOGICAL INCENTIVES**

- Commitment problems arise when the inability to make credible threats and promises prevents people from achieving desired outcomes. Such problems can sometimes be solved by employing commitment devices—ways of changing incentives to facilitate making credible threats or promises.
- Most applications of the theory of games assume that players are self-interested in the narrow sense of the term. In practice, however, many choices—such as leaving tips in out-of-town restaurants—appear inconsistent with this assumption.
- The fact that people seem driven by a more complex range of motives makes behavior more difficult to predict, but also creates new ways of solving commitment problems. Psychological incentives often can serve as commitment devices when changing players' material incentives is impractical. For example, people who are able to identify honest trading partners, and interact selectively with them, are able to solve commitment problems that arise from lack of trust.

SUMMARY

- Economists use the theory of games to analyze situations in which the payoffs of one's actions depend on the actions taken by others. Games have three basic elements: the players; the list of possible actions, or strategies, from which each player can choose; and the payoffs the players receive for those strategies. The payoff matrix is the most useful way to summarize this information in games in which the timing of the players' moves is not decisive. In games in which timing matters, a decision tree provides a much more useful summary of the information. (LO1, LO3)
- Equilibrium in a game occurs when each player's strategy choice yields the highest payoff available, given the strategies chosen by the other. (LO1)
- A dominant strategy is one that yields a higher payoff regardless of the strategy chosen by the other player. In some games such as the prisoner's dilemma, each player has a dominant strategy. Equilibrium occurs in such games when each player chooses his or her dominant strategy. In other games, not all players have a dominant strategy. (LO1, LO2)
- Equilibrium outcomes are often unattractive from the perspective of players as a group. The prisoner's dilemma has this feature because it is each prisoner's dominant strategy to confess, yet each spends more time in jail if both confess than if both remain silent. The incentive structure of this game helps explain such disparate social dilemmas as excessive advertising, cartels, and failure to reap the potential benefits of interactions requiring trust. (LO2, LO3)
- Individuals often can resolve these dilemmas if they can make binding commitments to behave in certain ways. Some commitments—such as those involved in military arms control agreements—are achieved by altering the material incentives confronting the players. Other commitments can be achieved by relying on psychological incentives to counteract material payoffs. Moral sentiments such as guilt, sympathy, and a sense of justice often foster better outcomes than can be achieved by narrowly self-interested players. For this type of commitment to work, the relevant moral sentiments must be discernible by one's potential trading partners. (LO4)

KEY TERMS

basic elements of a game
cartel
commitment device
commitment problem
credible promise

credible threat
decision tree (or game tree)
dominant strategy
dominated strategy
Nash equilibrium

payoff matrix
prisoner's dilemma
repeated prisoner's dilemma
tit-for-tat

REVIEW QUESTIONS

1. Identify the three basic elements of a game. (LO1)
2. How is your incentive to defect in a prisoner's dilemma altered if you learn that you will play the game not just once but rather indefinitely many times with the same partner? (LO2)
3. Explain why a cartel is an example of a prisoner's dilemma. (LO2)
4. Why did Warner Brothers make a mistake by waiting until the filming of *Analyze This* was almost finished before negotiating with Tony Bennett to perform in the final scene? (LO3)
5. Suppose General Motors is trying to hire a small firm to manufacture the door handles for Buick sedans. The task requires an investment in expensive capital equipment that cannot be used for any other purpose. Why might the president of the small firm refuse to undertake this venture without a long-term contract fixing the price of the door handles? (LO3)
6. Describe the commitment problem that narrowly self-interested diners and servers would confront at restaurants located on interstate highways. Given that in such restaurants tipping does seem to ensure reasonably good service, do you think people are always selfish in the narrowest sense? (LO4)

PROBLEMS



1. Consider the following game, called matching pennies, which you are playing with a friend. Each of you has a penny hidden in your hand, facing either heads up or tails up (you know which way the one in your hand is facing). On the count of "three," you simultaneously show your pennies to each other. If the face-up side of your coin matches the face-up side of your friend's coin, you get to keep the two pennies. If the faces do not match, your friend gets to keep the pennies. (LO1)
 - a. Who are the players in this game? What are each player's strategies? Construct a payoff matrix for the game.
 - b. Does either player have a dominant strategy? If so, what?
 - c. Is there an equilibrium? If so, what?
2. Consider the following game. Harry has four quarters. He can offer Sally from one to four of them. If she accepts his offer, she keeps the quarters Harry offered her and Harry keeps the others. If Sally declines Harry's offer, they both get nothing (\$0). They play the game only once, and each cares only about the amount of money he or she ends up with. (LO1, LO3)
 - a. Who are the players? What are each player's strategies? Construct a decision tree for this game.
 - b. Given their goal, what is the optimal choice for each player?
3. Blackadder and Baldrick are rational, self-interested criminals imprisoned in separate cells in a dark medieval dungeon. They face the prisoner's dilemma displayed in the matrix.

Assume that Blackadder is willing to pay \$1,000 for each year by which he can reduce his sentence below 20 years. A corrupt jailer tells Blackadder that before he decides whether to confess or deny the crime, she can tell him Baldrick's decision. How much is this information worth to Blackadder? (LO2)

		Blackadder	
		Confess	Deny
Baldrick	Confess	5 years for each	0 for Baldrick 20 years for Blackadder
	Deny	20 years for Baldrick 0 for Blackadder	1 year for each

4. In studying for your economics final, you are concerned about only two things: your grade and the amount of time you spend studying. A good grade will give you a benefit of 20; an average grade, a benefit of 5; and a poor grade, a benefit of 0. By studying a lot, you will incur a cost of 10; by studying a little, a cost of 6. Moreover, if you study a lot and all other students study a little, you will get a good grade and they will get poor ones. But if they study a lot and you study a little, they will get good grades and you will get a poor one. Finally, if you and all other students study the same amount of time, everyone will get average grades. Other students share your preferences regarding grades and study time. (LO2)

- Model this situation as a two-person prisoner's dilemma in which the strategies are to study a little and to study a lot, and the players are you and all other students. Construct a payoff matrix in which the payoffs account for both the cost and benefit of studying.
 - What is the equilibrium outcome in this game? Which outcome would everyone (both the other students and you) prefer?
5. Newfoundland's fishing industry has recently declined sharply due to overfishing, even though fishing companies were supposedly bound by a quota agreement. If all companies had abided by the agreement, yields could have been maintained at high levels. (LO2)
- Model this situation as a prisoner's dilemma in which the players are Company A and Company B, and the strategies are to keep the quota and break the quota. Suppose that if both companies keep the quota, then each receives a payoff of \$100, and if both break the quota, then each receives a payoff of \$0. On the other hand, if one company breaks the quota and the other keeps the quota, then the company that breaks the quota receives a payoff of \$150 and the company that keeps the quota receives a payoff of -\$50. Construct the corresponding payoff matrix, and explain why overfishing is inevitable in the absence of effective enforcement of the quota agreement.
 - Provide another environmental example of a prisoner's dilemma.
 - In many potential prisoner's dilemmas, a way out of the dilemma for a would-be cooperator is to make reliable character judgments about the trustworthiness of potential partners. Explain why this solution is not available in many situations involving degradation of the environment.
6. Two airplane manufacturers are considering the production of a new product, a 150-passenger jet. Both are deciding whether to enter the market and

produce the new planes. The payoff matrix is as follows (payoff values are in millions of dollars).

		Airbus	
		Produce	Don't produce
Boeing	Produce	-5 for each	100 for Boeing 0 for Airbus
	Don't produce	0 for Boeing 100 for Airbus	0 for each

The implication of these payoffs is that the market demand is large enough to support only one manufacturer. If both firms enter, both will sustain a loss. (LO2)

- Identify two possible equilibrium outcomes in this game.
 - Consider the effect of a subsidy. Suppose the European Union decides to subsidize the European producer, Airbus, with a check for \$25 million if it enters the market. Revise the payoff matrix to account for this subsidy. What is the new equilibrium outcome?
 - Compare the two outcomes (pre- and post-subsidy). What qualitative effect does the subsidy have?
7. Jill and Jack both have two pails that can be used to carry water down a hill. Each makes only one trip down the hill, and each pail of water can be sold for \$5. Carrying the pails of water down requires considerable effort. Both Jill and Jack would be willing to pay \$2 each to avoid carrying one pail down the hill and an additional \$3 to avoid carrying a second pail down the hill. (LO2)
- Given market prices, how many pails of water will each child fetch from the top of the hill?
 - Jill and Jack's parents are worried that the two children don't cooperate enough with one another. Suppose they make Jill and Jack share equally their revenues from selling the water. Given that both are self-interested, construct the payoff matrix for the decisions Jill and Jack face regarding the number of pails of water each should carry. What is the equilibrium outcome?
8. The owner of a thriving business wants to open a new office in a distant city. If he can hire someone who will manage the new office honestly, he can

afford to pay that person a weekly salary of \$2,000 (\$1,000 more than the manager would be able to earn elsewhere) and still earn an economic profit of \$800. The owner's concern is that he will not be able to monitor the manager's behavior and that the manager would therefore be in a position to embezzle money from the business. The owner knows that if the remote office is managed dishonestly, the manager can earn \$3,100, which results in an economic loss of \$600 per week. (LO3)

- a. If the owner believes that all managers are narrowly self-interested income maximizers, will he open the new office?
 - b. Suppose the owner knows that a managerial candidate is a genuinely honest person who condemns dishonest behavior and who would be willing to pay up to \$15,000 to avoid the guilt she would feel if she were dishonest. Will the owner open the remote office?
9. Consider the following "dating game," which has two players, A and B, and two strategies, to buy a movie ticket or a baseball ticket. The payoffs, given in points, are as shown in the matrix below. Note that the highest payoffs occur when both A and B attend the same event.

		B	
		Buy movie ticket	Buy baseball ticket
A	Buy movie ticket	2 for A 3 for B	0 for A 0 for B
	Buy baseball ticket	1 for A 1 for B	3 for A 2 for B

Assume that players A and B buy their tickets separately and simultaneously. Each must decide what to do knowing the available choices and payoffs but not

what the other has actually chosen. Each player believes the other to be rational and self-interested. (LO1, LO2, LO3)

- a. Does either player have a dominant strategy?
 - b. How many potential equilibria are there? (*Hint:* To see whether a given combination of strategies is an equilibrium, ask whether either player could get a higher payoff by changing his or her strategy.)
 - c. Is this game a prisoner's dilemma? Explain.
 - d. Suppose player A gets to buy his or her ticket first. Player B does not observe A's choice but knows that A chose first. Player A knows that player B knows he or she chose first. What is the equilibrium outcome?
 - e. Suppose the situation is similar to part d, except that player B chooses first. What is the equilibrium outcome?
10. Imagine yourself sitting in your car in a campus parking lot that is currently full, waiting for someone to pull out so that you can park your car. Somebody pulls out, but at the same moment a driver who has just arrived overtakes you in an obvious attempt to park in the vacated spot before you can. Suppose this driver would be willing to pay up to \$10 to park in that spot and up to \$30 to avoid getting into an argument with you. (That is, the benefit of parking is \$10 and the cost of an argument is \$30.) At the same time he guesses, accurately, that you too would be willing to pay up to \$30 to avoid a confrontation and up to \$10 to park in the vacant spot. (LO3, LO4)
- a. Model this situation as a two-stage decision tree in which his bid to take the space is the opening move and your strategies are (1) to protest and (2) not to protest. If you protest (initiate an argument), the rules of the game specify that he has to let you take the space. Show the payoffs at the end of each branch of the tree.
 - b. What is the equilibrium outcome?
 - c. What would be the advantage of being able to communicate credibly to the other driver that your *failure* to protest would be a significant psychological cost to you (for example, maybe a cost of \$25)?

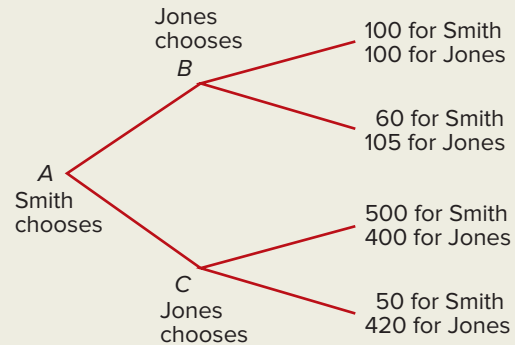
ANSWERS TO SELF-TESTS

- 7.1 No matter what American does, United will do better to leave ad spending the same. No matter what United does, American will do better to raise ad spending. So each player will play its dominant strategy: American will raise its ad spending and United will leave its ad spending the same. (LO1)

		American's Choice	
		Raise ad spending	Leave ad spending the same
United's Choice	Raise ad spending	United gets \$3,000 American gets \$8,000	United gets \$4,000 American gets \$5,000
	Leave ad spending the same	United gets \$8,000 American gets \$4,000	United gets \$5,000 American gets \$2,000

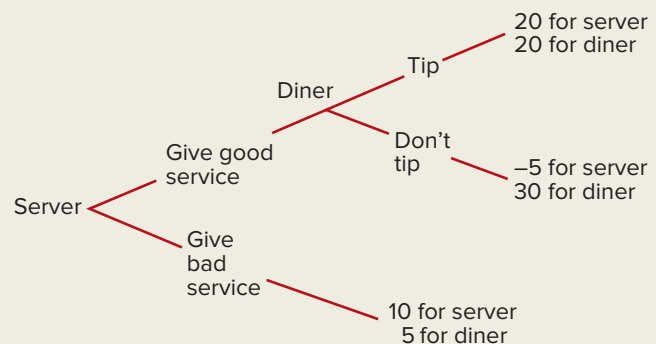
- 7.2 In game 1, no matter what Chrysler does, GM will do better to invest, and no matter what GM does, Chrysler will do better to invest. Each has a dominant strategy, but in following it, each does worse than if it had not invested. So game 1 is a prisoner's dilemma. In game 2, no matter what Chrysler does, GM again will do better to invest; but no matter what GM does, Chrysler will do better *not* to invest. Each has a dominant strategy, and in following it, each gets a payoff of 10—which is 5 more than if each had played its dominated strategy. So game 2 is not a prisoner's dilemma. (LO2)
- 7.3 Smith assumes that Jones will choose the branch that maximizes his payoff, which is the bottom branch at either B or C. So Jones will choose the bottom branch when his turn comes, no matter what Smith chooses. Since Smith will do better (60) on the bottom branch at B than on the bottom branch at C (50), Smith will choose the top

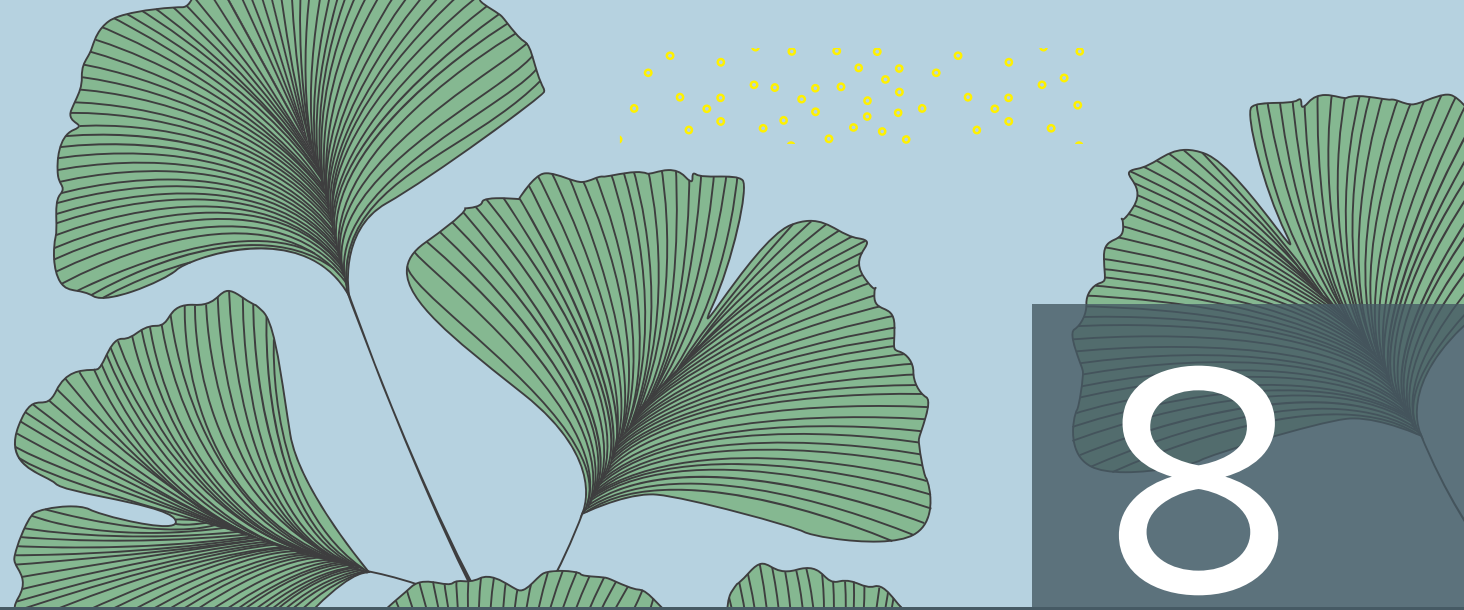
branch at A. So equilibrium in this game is for Smith to choose the top branch at A and Jones to choose the bottom branch at B. Smith gets 60 and Jones gets 105.



If Jones could make a credible commitment to choose the top branch no matter what, both would do better. Smith would choose the bottom branch at A and Jones would choose the top branch at C, giving Smith 500 and Jones 400. (LO3)

- 7.4 The equilibrium of this game in the absence of a commitment to tip is that the server will give bad service because if she provides good service, she knows that the diner's best option will be not to tip, which leaves the server worse off than if she had provided good service. Since the diner gets an outcome of 20 if he can commit to leaving a tip (15 more than he would get in the absence of such a commitment), he would be willing to pay up to 15 for the right to commit. (LO4)





8

An Introduction to Behavioral Economics

Until the 1980s, traditional economic models almost invariably assumed a decision maker who was narrowly self-interested, well informed, highly disciplined, and possessed of sufficient cognitive capacity to solve relatively simple optimization problems. This mythical individual was often called **homo economicus** for short.

During the intervening years, theoretical and empirical developments in economics and psychology have challenged each of these core assumptions. The challenges fall into three broad categories: (1) we often make systematic cognitive errors that prevent us from discovering which choices will best promote our interests; (2) even when we can discern which choices would be best, we often have difficulty summoning the willpower to execute them; and (3) we often pursue goals that appear inconsistent with self-interest, narrowly understood. In this chapter, we will consider each of these challenges in turn.

The late Nobel laureate Herbert Simon (1984) was the first to impress upon economists that human beings are incapable of behaving like the rational beings portrayed in standard rational choice models. Simon, a pioneer in the field of artificial intelligence, stumbled upon this realization in the process of trying to instruct a computer to “reason” about a problem. He discovered that when we ourselves confront a puzzle, we rarely reach a solution in a neat, linear fashion. Rather, we search in a haphazard way for potentially relevant facts and information and usually quit once our understanding reaches a certain threshold. Our conclusions are often inconsistent, sometimes even flatly incorrect. But much of the time, we come up with serviceable, if imperfect, solutions. In Simon’s terms, we are “**satisficers**,” not maximizers. We move on once we feel we’ve got a solution that is “good enough.”

Subsequent economists have taken Simon’s lead and developed a very sophisticated literature on decision making under incomplete information. We now realize that when information is costly to gather and cognitive processing ability is limited, it is not even rational to make fully informed choices of the sort portrayed in traditional models. Paradoxically, it is irrational to be completely well informed! When information is costly, the benefit from gathering more of it—from being able to make a very

LEARNING OBJECTIVES

After reading this chapter, you should be able to:

- LO1** Describe and illustrate numerous rules of thumb people employ when making decisions using incomplete information and explain how reliance on these rules sometimes leads to poor choices.
- LO2** Explain how misinterpretation of contextual clues can lead to poor choices.
- LO3** Describe situations in which impulse-control problems prevent people from executing rationally conceived economic plans.
- LO4** Describe the phenomenon of loss aversion and explain how it often creates a powerful bias in favor of the status quo.
- LO5** Describe how concerns about relative position often create conflicts between the interests of individuals and those of the broader community.
- LO6** List examples of laws and regulations that appear motivated by the behavioral tendencies described in this chapter.

homo economicus the narrowly self-interested, well-informed, highly disciplined, and cognitively formidable actor often assumed in traditional economic models

satisficing a decision-making strategy that aims for adequate results because optimal results may necessitate excessive expenditure of resources

fungibility the property of an entity whose individual units are interchangeable, as money in separate accounts



Carsten Rehder/AFP/Getty Images

Richard Thaler, who was awarded the Nobel Prize in Economics in 2017 for his pioneering work in behavioral economics.

judgmental and decision heuristics rules of thumb that reduce computation costs

availability heuristic a rule of thumb that estimates the frequency of an event by the ease with which it is possible to summon examples from memory

good decision rather than just a good one—may simply not justify the added cost. Thus, the literature on decision making under incomplete information, far from being a challenge to the traditional model, has actually bolstered our confidence in it.

But there is another offshoot of Simon's work, one that is less friendly to traditional models. This research, which has been strongly influenced by the economist Richard Thaler and the cognitive psychologists Daniel Kahneman and the late Amos Tversky, demonstrates that even with transparently simple problems, people often violate the most fundamental axioms of rational choice. One of the most cherished tenets of the rational choice model, for example, is that wealth is fungible, or freely movable across categories. **Fungibility** implies, among other things, that our total wealth, not the amount we have in any particular account, determines what we buy. But Tversky and Kahneman provided a vivid experimental demonstration to the contrary.¹ They asked one group of people to imagine that, having earlier purchased tickets for \$10, they arrive at the theater to discover they have lost them. Members of a second group were asked to picture themselves arriving just before the performance to buy their tickets when they find that they have each lost \$10 from their wallets. People in both groups were then asked whether they would continue with their plans to attend the performance. In the rational choice model, the forces governing this decision are exactly the same for both groups. Losing a \$10 ticket should have precisely the same effect as losing a \$10 bill. And yet, in repeated trials, most people in the lost-ticket group said they would not attend the performance, while an overwhelming majority—88 percent—in the lost-bill group say they would.

Richard Thaler developed a theory of mental accounting to explain this anomaly.² He began by noting that people apparently organize their spending into separate “mental accounts” for food, housing, entertainment, general expenses, and so on. People who lose their tickets act as if they debit \$10 from their mental entertainment accounts, while those who lose a \$10 bill debit that amount from their general expense account. For people in the former group, the loss makes the apparent cost of seeing the show rise from \$10 to \$20, whereas for those in the second it remains \$10. The rational choice model makes clear that the second group's assessment is the correct one. And on reflection, most people do, in fact, agree that losing a ticket is no better reason not to see the performance than losing a \$10 bill.

Working with numerous collaborators, Thaler, Kahneman, and Tversky identified a large catalog of systematic departures from rational choice, many of which stem from the application of **judgmental and decision heuristics**, or rules of thumb. Next, we describe some of the heuristics they identified.

JUDGMENTAL HEURISTICS OR RULES OF THUMB

Availability

We often estimate the frequency of an event, or class of events, by the ease with which we can summon relevant examples from memory.³ The more easily we can recall examples, the more likely we judge an event to be. This rule of thumb is called the **availability heuristic**. On balance, it is an effective strategy since it is generally easier to recall examples of things that happen frequently. The problem is that frequency is not the only factor that influences ease of recall. If people are asked, for example, whether there are more murders than suicides in New York State each year, most answer confidently in the affirmative, yet there are always more suicides than murders. Murders are easier to recall not because they are more frequent, but because they are more salient.

¹Amos Tversky and Daniel Kahneman, “Judgment under Uncertainty: Heuristics and Biases,” *Science* 185 (1974), pp. 1124–1131.

²Richard Thaler, “Mental Accounting and Consumer Choice,” *Marketing Science* 4, no. 3 (Summer 1985), pp. 199–214.

³Amos Tversky and Daniel Kahneman, “The Framing of Decisions and the Psychology of Choice,” *Science* 211 (1981), pp. 453–458.

EXAMPLE 8.1 **Availability Heuristic**

Which category of words is more common in the English language: those that begin with “r” or those that have “r” as their third letter?

Using the availability heuristic, most people react by trying to summon examples in each category. And since most people find it easier to think of examples of words starting with “r,” the availability heuristic leads them to answer that such words occur more frequently. Yet English words with “r” in the third slot are actually far more numerous. The availability heuristic fails here because frequency isn’t the only thing that governs ease of recall. We store words in memory in multiple ways—by their meanings, by the sounds they make, by the images they evoke, by their first letters, and by numerous other features. But virtually no one stores words in memory by the identity of their third letter, which is why words with “r” in that slot are harder to recall.

Representativeness

Another common rule of thumb is to assume that the likelihood of something belonging to a given category is directly related to the degree to which it shares characteristics thought to be representative of membership in that category.⁴ For example, because librarians are stereotypically viewed as introverted while salespersons are viewed as gregarious, the **representativeness heuristic** suggests that a given shy person is more likely to be a librarian than a salesperson. But this reasoning leads us astray when we fail to account for other factors that influence the relevant probabilities. Suppose the popular stereotypes are accurate—that, say, 90 percent of all librarians are shy, as compared with only 20 percent of salespeople. Would it then be safe to conclude that a given shy person is more likely to be a librarian than a salesperson? Not necessarily. Suppose, conservatively, that there are 90 salespeople in the population for every 10 librarians (the true ratio is more than 1,000 to 1). Under the assumed proportions of shy persons in each group, there would then be 9 shy librarians in the population for every 18 shy salespeople. And in that case, any given shy person would actually be twice as likely to be a salesperson than to be a librarian!

representativeness heuristic a rule of thumb according to which the likelihood of something belonging to a given category increases with the extent to which it shares characteristics with stereotypical members of that category

EXAMPLE 8.2 **Representativeness Heuristic**

Only two moving companies, United and North American, provide local delivery service in a small western city. United operates 80 percent of the vans, North American the remaining 20 percent. On a dark and rainy night, a pedestrian is run over and killed by a moving van. The lone witness to the incident testifies that the van was owned by North American. The law states that a company shall be held liable for damages only if it can be shown that the probability that the company was the guilty party is at least half. An independent laboratory hired by the court finds that under dark and rainy conditions, the witness is able to identify the owner of a moving van with 60 percent accuracy. What is the probability that a North American van ran over the pedestrian?

Out of every 100 moving vans in the area, 80 are United’s and 20 are North American’s. Of the 80 United vans, the witness incorrectly identifies 32 (or 40 percent of them) as belonging to North American. Of the 20 North American trucks, the witness correctly identifies 12 as belonging to North American. Thus the probability that a van identified as North American’s actually is North American’s is $12/(32 + 12) = 3/11$. So North American will not be held liable.

⁴Daniel Kahneman and Amos Tversky, “Subjective Probability: A Judgment of Representativeness,” *Cognitive Psychology* 3, no. 3 (1972), pp. 430–454.

SELF-TEST 8.1

A witness testifies that the taxicab that struck and injured Smith in a dark alley was green. On investigation, the attorney for Green Taxi Company discovers that the witness identifies the correct color of a taxi in a dark alley 80 percent of the time. There are two taxi companies in town, Green and Blue. Green operates 15 percent of all local taxis. The law says that Green Taxi Company is liable for Smith's injuries if and only if the probability that it caused them is greater than 0.5. Is Green liable?

regression to the mean
the phenomenon that unusual events are likely to be followed by more nearly normal ones

Regression to the Mean

Another common judgment error is to ignore a phenomenon known to statisticians as **regression to the mean**. The idea is that if our first measurement of something is far from its average value, a second measurement will tend to be closer to its average value. Suppose, for example, that you draw a ball from an urn containing balls numbered from 1 to 100 and get a ball numbered 85. Would you then be surprised that if you drew a second ball from that same urn, the number on it would be smaller than 85? Since the urn contains 84 balls with smaller numbers and only 15 with higher ones, it is in fact very likely that your second ball will have a smaller number. Errors often occur when people fail to take this phenomenon into account.

Suppose, for example, that you are a manager choosing whether to adopt a supportive stance toward your employees—by praising them when they perform well—or a critical stance—by criticizing them when they perform poorly. To help you decide, you do an experiment in which you adopt a critical stance toward employees who have performed unusually poorly and a supportive stance toward those who have performed unusually well. On the basis of this experiment, you decide that the critical stance works better. Why might this inference have been misleading?

When you praise an employee, it will often be after that employee has turned in an unusually good performance. Similarly, when you are critical of an employee, it will often be after that employee has performed unusually poorly. Quite apart from any direct effects of praise or blame, subsequent performances in both cases are likely to be more nearly normal. This could lead you to conclude that praise caused worse performance and that criticism caused better performance. Independent experimental work suggests that this conclusion would be erroneous. Supportive management styles, it appears, are more likely than critical styles to bring out the best in employees.

**The Economic Naturalist 8.1**

Why did the American Olympic swimmer Shirley Babashoff, who set one world record and six national records at the 1976 Olympics, refuse to appear on the cover of *Sports Illustrated*?



AP Images

Invitations to appear on the cover of the magazine typically come, as in Babashoff's case, in the wake of outstanding athletic performances. Many athletes believed that appearing on the cover of *Sports Illustrated* made them subject to the *Sports Illustrated* jinx, which doomed them to subpar performance during the ensuing months. But this perception was a statistical illusion. As the psychologist Tom Gilovich explained, "It does not take much statistical sophistication to see how regression effects may be responsible for the belief in the *Sports Illustrated* jinx. Athletes' performances at different times are imperfectly correlated.

Thus, due to regression alone, we can expect an extraordinary good performance to be followed, on the average, by a somewhat less extraordinary performance. Athletes appear on the cover of *Sports Illustrated* when they are newsworthy—i.e., when their performance is extraordinary. Thus, an athlete's superior performance in the weeks preceding a cover story is very likely to be followed by somewhat poorer performance in the weeks after. Those who believe in the jinx . . . are mistaken, not in what they observe, but in how they interpret what they see. Many athletes do suffer deterioration in their performance after being pictured on the cover of *Sports Illustrated*, and the mistake lies in citing a jinx, rather than citing regression as the proper interpretation of this phenomenon.”⁵

SELF-TEST 8.2

Following an unusual increase in burglaries in a New York City neighborhood, the chief of police assigned additional officers to patrol the neighborhood. In the following month, the number of burglaries declined significantly. Does this mean the increased patrols were effective?

Anchoring and Adjustment

Another common judgmental heuristic is called **anchoring and adjustment**. It holds that when we try to estimate something, we often begin with a tentative initial estimate, called the anchor, which we then adjust in the light of whatever additional information is available. A common bias pattern observed in anchoring and adjustment is that the anchors are sometimes of questionable relevance, and the adjustments people make are often insufficient.

Suppose, for example, that a professor who stands 6'4" asks his students to estimate his height. They may often anchor on the average height for men—about 5'10" in the United States—and then adjust that value upward because their visual assessment indicates that their professor is significantly taller than average. Typically, however, their adjustments will fall short. In most cases, the average of their estimates of the professor's height in this example will be two or three inches shorter than his actual height.

A particularly vivid example of this potential bias comes from an experiment in which American students were asked to estimate the percentage of African countries belonging to the United Nations. Most had no idea what the true percentage was. Before even hearing the question, subjects in two groups were asked to perform the ostensibly unrelated task of spinning a wheel that could stop on any number from 1 to 100. In the first group, the wheel was rigged to stop on 65, while in the second it was rigged to stop on 10. Students in both groups surely knew that the number on which the spinning wheel stopped bore no relation to the true percentage of African nations in the UN. Yet the average estimate of the group whose wheel stopped on 65 was 45 percent, while the average estimate of the group whose wheel stopped on 10 was only 25.⁶

anchoring and adjustment

an estimation technique that begins with an initial approximation (the anchor), which is then modified in accordance with additional available information (the adjustment)

⁵Thomas Gilovich, *How We Know What Isn't So: The Fallibility of Human Reason in Everyday Life* (New York: Free Press, 1991).

⁶Amos Tversky and Daniel Kahneman, "Judgment under Uncertainty: Heuristics and Biases," *Science* 185 (1974), pp. 1124–1131.

RECAP ↑

JUDGMENTAL HEURISTICS AND RULES OF THUMB

According to the *availability heuristic*, we often estimate the frequency of an event by the ease with which we can recall relevant examples. Errors often result because frequency is only imperfectly correlated with ease of recall.

The *representative heuristic* holds that the likelihood of something belonging to a given category is directly related to the degree to which it resembles stereotypical members of that category. Here, too, errors often result because other factors also influence the relevant probabilities.

The tendency to ignore *regression to the mean*—the idea that extreme events are likely to be followed by more nearly normal ones—is yet another source of judgmental errors.

Anchoring and adjustment holds that when we try to estimate something, we often begin with a tentative anchor, which we then adjust in the light of additional information. The anchors that people employ are often of questionable relevance, and the corresponding adjustments are often insufficient.

MISINTERPRETATION OF CONTEXTUAL CLUES

Another important class of judgment errors stems from misinterpretation of common contextual cues.

The Psychophysics of Perception

As psychologists have long understood, our assessments are almost always heavily shaped by the local environmental context. Consider a couple driving with their 10-year-old daughter to visit her grandparents. They are 10 miles from their destination when she asks, “Are we almost there yet?” If those 10 miles remain on a 120-mile journey, her parents will reassure her that they’re nearly there. But if the same 10 miles remain on a 12-mile journey they will answer differently: “No, honey, we’ve still got quite a way to go.” Similarly, in the year 1920, if your car could reach a speed of 60 miles per hour eventually, it would have seemed blazingly fast. But unless a car could reach 60 mph in under 6 seconds today, many motorists would think it sluggish.

Weber-Fechner law the relationship according to which the perceived change in any stimulus varies according to the size of the change measured as a proportion of the original stimulus

Every assessment of this sort rests on an explicit or implicit frame of reference. According to a relationship known as the **Weber-Fechner law** of psychophysics, our perception of the change in any stimulus varies according to the size of the change measured as a proportion of the original stimulus. In some contexts, there is nothing problematic about reckoning importance in proportional terms. But in other contexts, it can lead us astray. Real estate developers, for example, often exploit people’s tendency to think in percentage terms by bundling additional features into the prices of their houses for sale. Buyers might be reluctant to spend \$5,000 for a jacuzzi for a house they already owned, but might see the same expense as far less daunting if it meant paying \$250,000 for a new house instead of \$245,000.⁷

The Difficulty of Actually Deciding

In the tale of Buridan’s ass, a starving donkey stands equidistant from two identical bales of hay. But because his attraction to each bale is exactly equal, he is unable to decide which bale to approach and therefore dies of hunger. Although it is difficult

⁷For more examples along similar lines, see Richard Thaler, “Toward a Positive Theory of Consumer Choice,” *Journal of Economic Behavior and Organization* 1 (1980), pp. 39–60.

to imagine such indecision causing a human to die of hunger, the fact remains that many people experience significant anxiety when forced to decide between two options that are roughly equally attractive. If the options were, in fact, equally attractive, it would of course not matter which one was chosen. But the anxiety some people experience in such situations leads them to violate basic predictions of traditional economic models.

Suppose you were faced with a choice between the two apartments labeled *A* and *B* in Figure 8.1. The attraction of *A* is that it's close to campus, but the downside is that its rent is fairly high. In contrast, *B*'s rent is much more reasonable, but unfortunately it happens to be much farther from campus. By suitable manipulation of the rent/distance values, it should be possible to describe pairs of apartments like *A* and *B* such that a given individual would be essentially indifferent between them. The interesting thing, however, is that this doesn't mean that she would find it easy to choose. On the contrary, many people experience significant anxiety when confronted with such choices.

Now suppose we gather a large group of subjects and manipulate the distance/rent values of *A* and *B* so that when forced to choose, half the subjects choose *A* and the remaining half choose *B*. What do you think would have happened if we had asked that same group of subjects to choose among three apartments—the same *A* and *B* as before plus a new option at *C*, as shown in Figure 8.2. According to traditional rational choice theory, the apartment at *C* is an irrelevant option, since it is worse along both dimensions than *B*. And, in fact, when *C* is added to the set of options, no one chooses it.

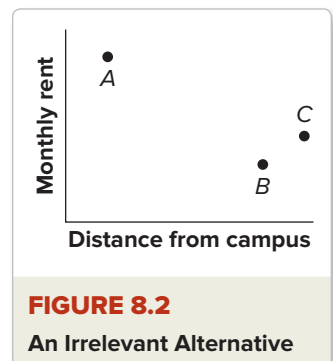
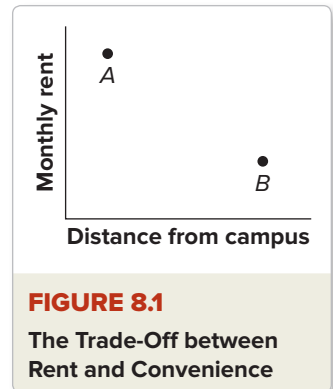
Yet the presence of the option at *C* has a profound effect on the observed pattern of choices between *A* and *B*. This time many more subjects choose *B*, and many fewer choose *A*.⁸ The apparent explanation of the shift is that subjects find it easy to choose *B* over *C*, and that this invests option *B* with a halo that leads subjects to favor it when confronted with the anxiety-inducing choice between *A* and *B*.

Experienced salespersons may exploit this pattern when dealing with people who have difficulty choosing among hard-to-compare alternatives. To close the deal, it may be enough to expose the client to a new option that is worse along every dimension than one of the original options.

As the preceding examples illustrate, cognitive errors often lead to departures from the predictions of standard rational choice models. We call such examples “departures from rational choice with regret,” the idea being that once people are made aware of the cognitive errors they have made, they seem motivated to avoid those errors in the future.

Many proponents of traditional economic models appear reluctant to amend those models to take account of systematic cognitive errors. Some argue, for example, that in competitive environments, the penalties associated with such errors will reduce their frequency, enabling us to safely neglect them. Yet the mere fact that people make systematic cognitive errors does not imply that that reliance on heuristics is maladaptive in any global sense. The important question is whether following some alternative strategy would lead to better results on average. Rules like the availability heuristic are extremely easy to apply and work well much of the time. The costs of an occasional wrong decision must be weighed against the obvious advantages of reliance on simple decision and judgment rules.

In any event, there is clear evidence that systematic cognitive errors exist and are widespread. Absent clear evidence that reliance on heuristics is a losing strategy on balance, there is a strong methodological case for taking explicit account of cognitive errors in descriptive models of individual behavior. A plausible case can also be made that by becoming more aware of the circumstances that are likely to elicit cognitive errors, we can learn to commit those errors less frequently.



McGraw Hill connect

► Visit your instructor's Connect course and access your eBook to view this video.



Why do real estate agents often show clients two nearly identical houses, even though one is both cheaper and in better condition than the other?

⁸Amos Tversky and Itamar Simonson, “Context-Dependent Preferences,” *Management Science* 39, no. 10 (October 1993), pp. 1179–1189.

RECAP ↑**MISINTERPRETATION OF CONTEXTUAL CLUES**

Almost every human judgment requires an explicit or implicit frame of reference. The *Weber-Fechner law* holds that perceptions of the change in any stimulus vary according to the size of the change as a proportion of the original stimulus. This reckoning leads many to say, for example, that they would drive across town to save \$10 on a \$20 radio but not to save \$10 on a \$1,000 computer. Yet if the benefit of driving across town exceeds the cost of driving across town for the radio, the same must be true for the computer (since the benefit is \$10 in both cases and the cost of driving is the same in both cases).

Other choice anomalies arise because people often have difficulty choosing between two alternatives whose characteristics are difficult to compare. When confronted by such choices, the addition of a third alternative that is worse than the first of the original alternatives on every dimension often has the effect of tilting choice heavily in favor of the first alternative. According to traditional models, the addition of such an alternative should have no effect on the choice between the other two.

IMPULSE-CONTROL PROBLEMS

Another troubling feature of traditional economic models is that they appear to rule out the possibility that people might regret having chosen behaviors whose consequences were perfectly predictable at the outset. Yet such expressions of regret are common. Many people wake up wishing they had drunk less the night before, but few wake up wishing they had drunk more.

It is likewise a puzzle within traditional economic models that people often incur great expense and inconvenience to prevent behaviors they would otherwise freely choose. For example, the model has difficulty explaining why some people would pay thousands of dollars to attend weight-loss camps that will feed them only 1,500 calories per day.

Welfare analysis based on the traditional models assigns considerable importance, as it should, to people's own judgments about what makes them better off. In this framework, it is not clear how one could ever conclude that a risk freely chosen by a well-informed person was a wrong choice. By itself, the fact that the choice led to a bad outcome is clearly not decisive. If someone is unlucky enough to be killed crossing the street, for example, we would not conclude that crossing streets is a generally bad idea.

Compelling evidence from psychology suggests, however, that expressions of regret may often be genuine, that people might sensibly impose constraints on their own behavior, and that risks freely chosen by well-informed people may not be optimal from their own point of view. Much of this evidence concerns behavior in the realm of intertemporal choice.

The rational choice model says that people will discount future costs and benefits in such a way that the choice between two rewards that are separated by a fixed time interval will be the same no matter when the choice is made.⁹ Consider, for example, the pair of choices A and B:

A: \$100 tomorrow vs. \$110 a week from tomorrow.

B: \$100 after 52 weeks vs. \$110 after 53 weeks.

⁹This is known as "exponential discounting."

Both pairs of rewards are separated by exactly one week. If future receipts are discounted as assumed in traditional models, people will always make the same choice under alternative A as they do under alternative B. Since the larger payoff comes a week later in each case, the cost of choosing the smaller, but earlier, payoff is the same in each case—namely, the \$10 forgone by doing so. When people confront such choices in practice, however, there is a pronounced tendency to choose differently under the two scenarios: Most pick the \$100 option in A, whereas most choose the \$110 option in B.

Researchers who have studied intertemporal choices experimentally have amassed substantial evidence that perceived future costs and benefits fall much more sharply with delay than assumed in traditional models.¹⁰ One consequence is that preference reversals of the kind just observed occur frequently. The classic reversal involves choosing the larger, later reward when both alternatives occur with substantial delay, then switching to the smaller, earlier reward when its availability becomes imminent. This reversal shows up in *The Economic Naturalist* 8.2.

The Economic Naturalist 8.2



Why would people pay thousands of dollars to attend a weight-loss camp that will feed them only 1,500 calories per day?

Viewed within the framework of conventional rational choice models, it would be puzzling indeed that people would spend large sums to be prevented from eating food that they were free to refrain from eating on their own. But if people tend to discount future costs and benefits sharply, the answer is clear: They really want to eat less, but they know that if tempting foods are readily available, they will lack the willpower to do so. The weight-loss camp solves this problem by making tempting foods less available.

The tendency to discount future costs and benefits excessively gives rise to a variety of familiar impulse-control problems and, in turn, to a variety of strategies for solving them. Anticipating their temptation to overeat, people often try to limit the quantities of sweets, salted nuts, and other delicacies they keep on hand. Anticipating their temptation to spend cash in their checking accounts, people enroll in payroll-deduction savings plans. Foreseeing the difficulty of putting down a good mystery novel in midstream, many people know better than to start one on the evening before an important meeting. Reformed smokers seek the company of nonsmokers when they first try to kick the habit and are more likely than others to favor laws that limit smoking in public places. The recovering alcoholic avoids cocktail lounges.

Effective as these bootstrap methods of self-control may often be, they are far from perfect. Many people continue to express regret about having overeaten, having drunk and smoked too much, having saved too little, having stayed up too late,

¹⁰According to psychologists, people tend to discount future costs and benefits hyperbolically, not exponentially. Two of the most important earlier papers on this issue are Shin-Ho Chung and Richard Herrnstein, "Choice and Delay of Reinforcement," *Journal of the Experimental Analysis of Behavior* 10 (1967), pp. 67–74; George Ainslie, "Specious Reward: A Behavioral Theory of Impulsiveness and Impulse Control," *Psychological Bulletin* 82 (July 1975), pp. 463–496. See also Jon Elster, "Don't Burn Your Bridge before You Come to It: Seven Types of Ambiguity in Precommitment," Texas Law Review Symposium on Precommitment Theory, Bioethics, and Constitutional Law, September 2002; Thomas Schelling, "The Intimate Contest for Self-Command," *The Public Interest*, Summer 1980, pp. 94–118; R. Thaler and H. Shefrin, "An Economic Theory of Self-Control," *Journal of Political Economy* 89 (1981), pp. 392–405; Gordon Winston, "Addiction and Backsliding: A Theory of Compulsive Consumption," *Journal of Economic Behavior and Organization* 1 (1980), pp. 295–394. For a comprehensive review of work discussing the economic applications of the hyperbolic discounting model, see Shane Frederick, George Loewenstein, and Ted O'Donoghue, "Time Discounting and Time Preference: A Critical Review," *Journal of Economic Literature* 40, no. 2 (June 2002), pp. 351–401.

having watched too much television, and so on. Traditional economic models urge us to dismiss these expressions as sour grapes. But behavioral evidence suggests that these same expressions are often coherent. In each case, the actor chose an inferior option when a better one was available and later feels genuinely sorry about it. As with behaviors that stem from systematic cognitive errors, those that stem from impulse-control problems may be described as “departures from rational choice with regret.”

In view of the obvious difficulties to which excessive discounting gives rise, it might seem puzzling that this particular feature of human motivation is so widespread. As with cognitive errors that stem from reliance on judgmental heuristics, however, here too it is important to stress that a behavioral tendency leading to bad outcomes some of the time does not imply that it is maladaptive in a global sense. Heavy discounting of future events forcefully directs our attention to immediate costs and benefits. If forced to choose between a motivational structure that gave heavy weight to current payoffs and an alternative that gave relatively greater weight to future payoffs, it is easy to imagine why evolutionary forces might have favored the former. After all, in highly competitive and uncertain environments, immediate threats to survival were numerous, and there must always have been compelling advantages in directing most of our energies toward meeting them.

In any case, there is clear evidence here, too, that excessive discounting exists and is widespread. Whether or not this motivational structure is disadvantageous on balance, there is a persuasive case for taking explicit account of it if our aim is to predict how people will behave under real-world conditions.

We also note that societies around the world have adopted a variety of policies that are most easily understood as attempts to mitigate the consequences of impulse-control problems. These include the treatment of so-called crimes of passion; prohibitions against gambling, addictive drugs, and prostitution; entrapment laws; and sanctions against adultery. Programs that aim to stimulate savings can also be viewed as a response to impulse-control problems that make it difficult to execute rational savings plans.

RECAP ↑

IMPULSE-CONTROL PROBLEMS

Although many traditional economic models rule out the possibility that people might regret having chosen behaviors whose consequences were perfectly predictable, such regrets are common.

Compelling evidence from psychology suggests that expressions of regret may often be genuine, that people often have difficulty waiting for a larger, better reward when an inferior reward is available immediately. Traditional models also have difficulty explaining why people might pay large sums to attend a weight-loss camp that will dramatically constrain their daily caloric intake. But if people tend to discount future costs and benefits excessively, they know that if tempting foods are readily available, they may lack the willpower to avoid them. Without taking explicit account of self-control problems, it is difficult to predict how people will behave under real-world conditions.

Some of the most important developments in economics in recent decades have entailed revisions to traditional assumptions about the nature of people's preferences. We focus next on three major changes—the introduction of asymmetries in the way people evaluate alternatives involving gains and losses, the relaxation of the assumption that people are self-interested in the narrow sense, and the introduction of concerns about relative position.

LOSS AVERSION AND STATUS QUO BIAS

Traditional economic models hold that the extra satisfaction from gaining a dollar should be roughly equal to the reduction in satisfaction from losing a dollar. In practice, however, people's behavior suggests that losses often weigh much more heavily than gains.

An illustration of this tendency comes in the form of experiments designed to probe the difference in the value people assign to an object they already own versus the value they assign to one they have an opportunity to buy. In one experiment involving 44 subjects, half of participants were chosen at random to receive a coffee mug that they were free to keep or sell as they chose. The other half of the participants did not receive a mug. Participants were then given an opportunity to examine the mugs carefully—their own if they'd been given one, a neighboring subject's if they hadn't.

Subjects were then told to submit bids reflecting how much they'd be willing to pay for a mug if they didn't already have one or how much they'd be willing to accept for selling the mug they already had. In repeated rounds of this experiment, researchers used these bids to determine the prices at which the number of mugs offered would be equal to the number of mugs demanded. Those equilibrium prices varied between \$4.25 and \$4.75.

The researchers used the term *mug lovers* to describe the half of all subjects who happened to like the mugs most among all participants and *mug haters* to describe those who liked the mugs least. The fact that the mugs had been distributed at random implied that half of them, on average, would have ended up in the hands of mug lovers and the other half in the hands of mug haters. Thus, the apparent prediction of traditional economic models was that approximately half of the mugs—which is to say, 11 mugs on average—should be exchanged at the market-clearing prices. But the actual number of observed trades was dramatically smaller—only slightly more than 2 mugs changed hands, on average.

The apparent explanation was that the mere fact of owning a mug caused people to assign much greater value to it. The median price at which mug owners were willing to sell their mugs was \$5.25, for example, more than twice the amount the median non-owner was willing to pay.¹¹

This asymmetry in valuations is often called **loss aversion**. Similar experiments involving the exchange of modestly priced objects suggest that someone who possesses an object values it about twice as high, on average, as someone who does not possess it.

But there is evidence that when the things at stake are significantly more important, loss aversion becomes even more pronounced. In another experiment by Richard Thaler, subjects were asked to imagine that there was a 1 in 1,000 chance they had been exposed to an invariably fatal disease.¹² How much would they be willing to pay for the only available dose of an antidote that would save their lives? Their median response to this question was \$2,000. A second group of subjects was asked how much they would have to be paid to be willing to volunteer to be exposed to that same disease if there was no available antidote. This time the median response was close to \$500,000—more than 250 times higher than in the first case.

Note that in both cases, subjects are being asked how much they value a one-in-a-thousand reduction in their probability of dying. If people treated gains and losses symmetrically, their answers to the two questions would be roughly the same. But in the first case, they were being asked to buy an increase in their survival prospects (a gain), and in the second case, they were asked to sell a health benefit they already owned (a loss). People do not treat gains and losses symmetrically, and it appears that loss aversion becomes much more pronounced when the objects in question are more important.

Because every change in policy generates winners and losers, loss aversion has important implications for public policy. When the total benefits from a change in

loss aversion the tendency to experience losses as more painful than the pleasures that result from gains of the same magnitude

¹¹Daniel Kahneman, Jack L. Knetsch, and Richard H. Thaler, "Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias," *The Journal of Economic Perspectives* 5, no. 1 (Winter 1991), pp. 193–206.

¹²Michael Lewis, "The Economist Who Realized How Crazy We Are," *Bloomberg View*, May 29, 2015.

status quo bias the general resistance to change, often stemming from *loss aversion*

policy exceed the total cost, it will sometimes be practical for the winners to compensate the losers so that everyone comes out ahead. But when large numbers of people are involved, compensation often becomes impractical. And in such cases, resistance from those who stand to lose from the policy change often overwhelms the support of those who stand to gain, even when the total benefits substantially exceed the total costs. Loss aversion is thus an important source of what has been called **status quo bias** in the domain of public policy. The difficulty of enacting policy changes is encapsulated in the fabled *iron law of politics*: The losers always cry louder than the winners sing.

Is status quo bias a bad thing? It might seem reasonable to wonder whether the best posture for cost-benefit analysts might be to simply accept at face value the inflated weights that people assign to losses. In other words, perhaps the widespread resistance to change that we observe in practice should be interpreted to mean that change often does more harm than good.

Weighing against this interpretation, however, is evidence that people's initial estimates of how painful losses will be often fail to take into account the remarkable human tendency to adapt to altered circumstances. Many people say, for example, that they would prefer to be killed in an accident than to survive as a paraplegic. And most people who have become paralyzed have indeed experienced the transition as devastating. But what psychologists were surprised to discover was that paraplegics adapt more quickly and completely than they ever would have thought possible. According to one study, the post-accident happiness levels of paralyzed accident victims were significantly lower than they had been before their accidents, but not statistically different from those of a control group.¹³ Parting with a possession, similarly, is likely to be far more aversive at first than it is after a period of adaptation to living without it.

There are thus reasonable grounds for believing that steps to counteract status quo bias will often serve community interests. Behavioral economists have begun to explore such steps. One effective way to mitigate status quo bias is to focus more carefully on how different alternatives are presented to people. Suppose, for example, that policymakers want to encourage citizens to save a larger proportion of their incomes. Researchers have discovered that an effective and relatively painless way to accomplish that goal is to make participation in employer payroll-deduction savings the default option. Historically, most employers required their employees to take an active step in setting up a payroll-deduction savings plan. But researchers persuaded a sample of firms to make enrollment in the plan automatic, allowing those who didn't want to participate to take the active step of opting out. An influential study by Brigitte Madrian and Dennis Shea found that when participation became the status quo, savings plan enrollments shot up dramatically—in one case, a jump from less than 50 percent participation to 86 percent of all first-year employees.¹⁴ Similar experiments have demonstrated the power of default options to alter behavior in other domains, ranging from the kinds of insurance people buy to their willingness to be organ donors.



The Economic Naturalist 8.3

Why was Obamacare difficult to enact and is harder still to repeal?

The Obama administration's strategy for reforming the American health care system was heavily shaped by its understanding of status quo bias. Most other developed countries guarantee universal health care coverage either by having the government hire health care providers directly (as in the UK) or by having the government reimburse private health care providers for their services (as in France and Canada). But in the United States, most people received health insurance

¹³Philip Brinkman, Dan Coates, and Ronnie Janoff-Bulman, "Lottery Winners and Accident Victims: Is Happiness Relative?," *Journal of Personality and Social Psychology* 36, no. 8 (1978), pp. 917–927.

¹⁴Brigitte C. Madrian and Dennis F. Shea, "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior," *Quarterly Journal of Economics* 116, no. 4 (2001), pp. 1149–1187.

provided by their employers. Although there are sound policy reasons to favor the approach used elsewhere (where health systems deliver care at about half the cost as in the United States), government officials realized that many citizens had grown accustomed to employer-sponsored insurance and would resist attempts to require them to abandon it for a different approach. So the Affordable Care Act as passed in 2010 left the employer-provided insurance system intact, supplementing it with private market exchanges through which people without employer-provided insurance could purchase health insurance at affordable rates. Despite this compromise, the legislative effort to enact the Affordable Care Act was extremely contentious, and it passed by the slimmest of margins.

Once the act was implemented, it came under relentless attack from Republicans in the House of Representatives, who voted scores of times to repeal it. But to the surprise of no one familiar with loss aversion and status quo bias, these efforts also met stiff resistance. Obamacare had boosted the number of Americans with health insurance by more than 20 million, and estimates by the nonpartisan Congressional Budget Office suggested that repeal would reduce that number by a similar amount.

In short, once the Affordable Care Act was enacted into law, the conditions it brought about became the new status quo. So even as loss aversion helps explain why passage of the Affordable Care Act was so difficult in the first place, it also explains why it is so difficult to repeal.

RECAP ↑

LOSS AVERSION AND STATUS QUO BIAS

Widespread behavioral evidence suggests that people tend to weigh losses much more heavily than gains. The minimum price for which people would be willing to sell objects they already own is substantially higher than the price they would be willing to pay to acquire those same objects if they do not already own them.

This asymmetry is often called *loss aversion*. In experiments involving low-priced items, the difference between willingness to accept and willingness to pay is often on the order of two to one. But the ratio can be dramatically higher for high-valued items, such as conditions related to health and safety.

Because every change in policy generates winners and losers, loss aversion implies a large bias in favor of the status quo in public policy decisions. Evidence suggests that this bias often has harmful consequences since people's initial estimates of how painful losses will be often fail to take into account their ability to adapt to altered circumstances.

The strategic use of default options has proven to be an effective way to mitigate *status quo bias*. For example, making participation in employer payroll-deduction savings plans the default option has been shown to significantly reduce the tendency to save too little.

CONCERNS ABOUT RELATIVE POSITION

Many traditional economic models assume that well-being depends on current and future levels of absolute consumption. Considerable evidence suggests, however, that well-being depends as much or more on current and future levels of relative consumption.¹⁵

¹⁵The points in this section are developed in greater detail in Robert H. Frank, "The Demand for Unobservable and Other Nonpositional Goods," *American Economic Review* 75 (March 1985), pp. 101–116; Robert H. Frank, *Choosing the Right Pond* (New York: Oxford University Press, 1985).

The following thought experiment provides an intuitive test of whether relative consumption matters:

Which world would you choose?

World A: You and your family live in a neighborhood with 4,000-square-foot houses, and all other neighborhoods have 6,000-square-foot houses.

World B: You and your family live in a neighborhood with 3,000-square-foot houses, and all other neighborhoods have 2,000-square-foot houses.

If the satisfaction provided by a house depended only on its absolute features, World A would be the unambiguously better choice. Yet significantly more than half of respondents faced with this question choose World B. Some of the minority who choose World A seem to do so out of concern that they *shouldn't* care about others' houses. Even so, few among them seem puzzled that others might choose differently. The point is not that World B is the uniquely correct choice here. But if our goal is to predict people's actual choices, the traditional assumption that only absolute consumption matters is often misleading.

In some domains, however, the traditional assumption serves reasonably well. Consider a second thought experiment, one with exactly the same structure as the one for house size:

Which world would you choose?

World A: You have a 2 in 100,000 chance of dying on the job this year, and others have a 1 in 100,000 chance.

World B: You have a 4 in 100,000 chance of dying, and others have a 6 in 100,000 chance.

As before, this choice is between absolute advantage (World A) and relative advantage (World B). When people face that choice in the domain of housing, most choose relative advantage. But when their choice involves safety, almost everyone opts for absolute advantage.

positional good a good whose value depends relatively heavily on how it compares with other goods in the same category

nonpositional good a good whose value does not depend heavily on how it compares with other goods in the same category

The late British economist Fred Hirsch coined the term **positional good** to describe things that derive their value primarily from their relative scarcity rather than from their absolute characteristics. "The value of my education," he wrote, "depends on how much the man ahead of me in the job line has."¹⁶ In what follows, we use Hirsch's term to mean a good whose value depends relatively heavily on how it compares with other goods in the same category. For people who chose relative over absolute advantage in the first thought experiment, housing would thus be a positional good.

Similarly, we use the term **nonpositional good** to describe one whose value depends relatively little on how it compares with other goods in the same category. For people who chose absolute advantage in the second thought experiment, workplace safety would be a nonpositional good.

To call housing positional is not to say that only the relative attributes of a house matter. Someone who chose World B in the first thought experiment, for example, might well have preferred to live in a world in which her neighborhood had 3,500-square-foot houses and others had houses of 3,000 square feet.

Similarly, choosing a world with high absolute safety doesn't mean that relative safety levels are of no concern. Most people choosing World A would almost certainly notice that their jobs were much more dangerous than others'. But if their only recourse were to move to World B, where their own absolute risk of death would double, they would decline.

Caring more about relative consumption in some domains than in others distorts people's spending decisions. It leads to what we call *positional arms races*—patterns of escalating expenditure focused on positional goods. The resulting dynamic is exactly analogous to the one that drives military arms races. In both cases, wasteful spending occurs only because some categories of expenditure are more context-sensitive than others.

¹⁶Fred Hirsch, *Social Limits to Growth* (Cambridge, MA: Harvard University Press, 1976).

Military arms races occur only because relative differences in spending on weapons have more important consequences than relative differences in nonmilitary spending. If, to the contrary, spending less than a rival nation spent on toasters and television sets had more serious consequences than spending less on bombs, we would expect expenditure imbalances of precisely the opposite sort: Countries would spend ever less on armaments in a quest to gain relative advantage in the domain of nonmilitary spending. That is of course not what we see. Military arms races are a consequence, pure and simple, of relative spending on armaments being more important—because it contributes so heavily to the outcome of military conflict.

Similar distortions appear to occur in domestic consumption patterns. Because relative spending matters more, by definition, for positional goods than for nonpositional goods, people often end up spending too much on the former and too little on the latter. To see how the dynamic unfolds, consider a worker's choice between two jobs that offer different levels of risk to safety. Both because safety devices are costly and because workers prefer safe jobs to risky ones, employers who want to attract workers to riskier jobs can and must pay higher wages. Choosing additional safety on the job therefore comes at the expense of having less income to spend on other things. Choosing a riskier job would enable a worker to buy a nicer house. If she has children, that option might look doubly attractive since she knows that in almost every jurisdiction, better schools are those located in more expensive neighborhoods.

So let's imagine that she chooses the higher-paying but riskier job. The same logic would also lead similarly situated workers to make the same choice. And if they too chose riskier jobs to be able to afford houses in better school districts, their increased bidding would serve only to bid up the prices of those houses. Half of all children would still attend bottom-half schools, the same as if no one had sacrificed safety for higher wages. All will have sold their safety in pursuit of a goal that none achieves.

Since Adam Smith's day, classical economic theory has held that well-informed workers in competitive markets will sensibly navigate this trade-off between income and safety. They will accept additional risk in return for higher pay only if the satisfaction from having what they can buy with the extra income is greater than the corresponding loss in satisfaction from reduced safety. According to proponents of this view, regulations that mandate higher safety levels harm workers by forcing them to buy safety they value at less than its cost.

But why, then, does virtually every country in the world regulate workplace safety?¹⁷ (Even the poorest countries have at least rudimentary safety requirements.) Classical economic theory doesn't have a good answer. It portrays such regulations as either anomalous, or else needed because workers are uninformed, or because markets aren't competitive enough. Yet we regulate many safety risks that workers clearly understand. Most coal miners, for example, know their work entails risk of black lung disease since their fathers and grandfathers died of it.

Nor is imperfect competition the problem. Most safety regulations, after all, have their greatest impact in precisely those markets that most closely approximate the competitive ideal. They have little effect on engineers in Silicon Valley or on investment bankers on Wall Street, whose conditions of employment are substantially safer than required by the Occupational Safety and Health Administration. It's the employers who compete most fiercely with one another to attract the workers they need, such as fast-food restaurants and manufacturing firms that employ unskilled labor, that seldom go long without a visit from a safety inspector.

Analogous reasoning suggests that individual incentives are misleadingly high for many other things we might sell for money—our leisure time, future job security, and a host of other environmental amenities. The problem is not that we exchange these items for money, but that our incentives lead us to sell them too cheaply. In the case of workplace safety, the solution is not to ban risk, but to make it less attractive to individuals. As concerns the length of the workweek, the best solution is not strict limits

¹⁷Scott Coulter, "Advancing Safety around the World," *Occupational Health and Safety*, February 1, 2009.

on hours worked, but rather a change in incentives (such as overtime laws) that make long hours less attractive.

One rationale for policy measures to stimulate additional savings is that impulse-control problems often prevent people from adhering to rationally chosen savings plans. But savings shortfalls also stem from a second source, one that is difficult to address by unilateral action. The following thought experiment illustrates the basic problem:

If you were society's median earner, which of these two worlds would you prefer?

World A. You save enough to support a comfortable standard of living in retirement, but your children attend a school whose students score in bottom 20 percent on standardized tests in reading and math.

World B. You save too little to support a comfortable standard of living in retirement, but your children attend a school whose students score in the 50th percentile on those tests.

Because the concept of a "good" school is inescapably relative, this thought experiment captures an essential element of the savings decision confronting most middle-income families. If others bid for houses in better school districts, failure to do likewise will often consign one's children to inferior schools. But as noted earlier, no matter how much each family spends, half of all children must attend schools in the bottom half. The choice posed by the thought experiment is one that most parents would prefer to avoid. But when forced to choose, most say they would pick the second option.

Like workplace safety decisions, savings decisions are thus also examples of the collective-action problem inherent in a military arms race. Each nation knows that it would be better if all spent less on arms. Yet if others keep spending, it is simply too dangerous not to follow suit. Curtailing an arms race thus requires an enforceable agreement. Similarly, unless all families can bind themselves to save more, those who do so unilaterally will pay the price of having to send their children to inferior schools. In these observations lie the seeds of a possible answer to the question posed in The Economic Naturalist 8.4.



The Economic Naturalist 8.4

Why have attempts to privatize Social Security proved so politically unpopular in the United States?

Since its inception during the Great Depression, the Social Security system has required universal participation. People's labor earnings are subject to the payroll tax, which is then used to finance pension payments for older Americans in retirement. Early in his second term in office, President George W. Bush promoted a proposal to privatize Social Security. Under this proposal, people would be free to decide for themselves how much to save each month and how much to spend. To the president's surprise, however, this proposal proved deeply unpopular. Why didn't most people welcome the additional freedom that privatization would have offered?

As discussed, if parents had complete access to their retirement savings accounts, some would undoubtedly withdraw money from those accounts in order to bid for houses in better school districts. Rather than see their own children fall behind, many other parents would feel compelled to follow suit. But the increased bidding would serve only to drive up the prices of the houses in better school districts, even as many parents would end up with inadequate savings to support themselves in retirement. Mandatory participation in the Social Security system solves this dilemma. The money paid in payroll taxes, which is unavailable to bid for houses in better school districts, remains available to finance pension payments to retirees.

Positional concerns also appear to alter the distribution between public and private spending. Public goods are fundamentally different from private goods: When some people spend more on houses or wedding celebrations, for example, others may feel pressure to follow suit, but the same dynamic does not occur for public goods, which are available on essentially equal terms to all citizens. Because public goods do not lend themselves to interpersonal comparisons, they are nonpositional almost by definition.

In any event, there is little question that levels of public investment have not kept pace in recent decades. For example, engineers report record maintenance backlogs not only in the nation's roads and bridges, but also in airports, schools, drinking water, and sewage systems.¹⁸ And in the face of these backlogs, proposals to launch significant new infrastructure initiatives, such as high-speed rail service or a smart electric grid, have consistently failed in Congress. More troubling, we have not yet launched the investments required to confront the climate crisis, by far the biggest challenge facing the planet.

In short, substantial evidence suggests that most people would be happier if a larger share of national income were spent on nonpositional goods, including long-neglected public investment. Such misallocations tend to be stubborn, however, because they stem largely from the kinds of collective-action problems we earlier described as “departures from rational choice without regret.” Just as the practice of standing at concerts persists because it is rational for each individual to continue standing when others are doing so, spending imbalances that arise from collective-action problems cannot be remedied by individual action. People must turn to the political system in search of effective ways to act collectively.

In contrast, welfare losses that arise from cognitive errors can be attacked at the individual level. Once people realize that it's irrational to take sunk costs into account, for example, they may become both motivated to—and able to—alter their choices unilaterally. That's why we earlier described choices that are driven by cognitive errors as “departures from rational choice with regret.”

As noted, many of the tax and regulatory policies already in effect, such as workplace safety laws and pollution taxes, can be interpreted as positional arms control agreements. But if, as suggested above, additional implementation of such policies could in fact promote changes in spending patterns that would increase welfare substantially, we're left with the challenging question posed in The Economic Naturalist 8.5.

The Economic Naturalist 8.5



If prosperous voters would be happier if they spent less on positional goods and lived in environments with more generously funded public sectors, why haven't they elected politicians who would deliver what they want?

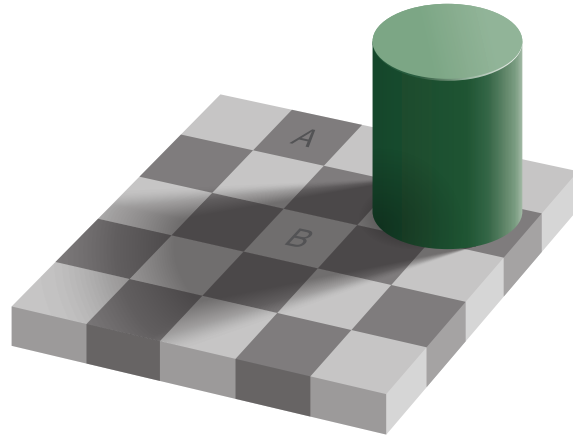
A partial answer to this question might begin with the observation that voters are extremely reluctant to pay the higher taxes necessary to refurbish decaying infrastructure and underwrite massive investment in green energy. But that begs the question: If the gains from the additional investment would indeed be much larger than the corresponding losses from reduced positional consumption, why would voters be opposed?

Here we consider the possibility that opposition is itself rooted in a powerful cognitive illusion. But first, we offer a few general observations about the nature of cognitive illusions: Life is complicated. Because we are bombarded by terabytes of information each day, our nervous systems employ shortcuts, or rules of thumb,

¹⁸2017 Infrastructure Report Card: D+, www.infrastructurereportcard.org.

for evaluating it. As noted earlier, these heuristics work reasonably well most of the time—but not always.

In the diagram below, which square is darker, A or B? If you think A looks darker, your eyes and brain are functioning normally. But in this instance, your judgment is incorrect. In what's called the checker-shadow illusion, Square A is exactly the same shade of gray as Square B. Look at the figure carefully. If your reaction is typical, you might be thinking, "That can't be true!" And yet it is.

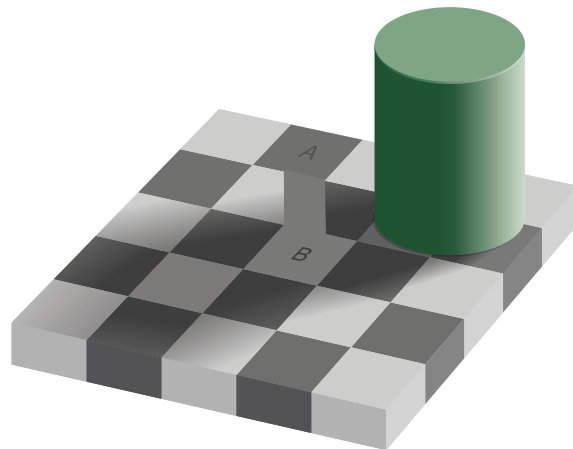


The psychologist Richard Wiseman offers this explanation:

Your eyes and brain see that the two squares are the same shade of gray, but then think, "Hold on—if a square in a shadow reflects the same amount of light as a square outside of the shadow, then in reality [Square B] must be a much lighter shade of gray." As a result, your brain alters your perception of the image so that you see what it thinks is out there in the real world.¹⁹

This is of course an adaptive response, something we'd want our brains to do. Yet Wiseman's explanation, plausible though it sounds, is insufficient to convince most people.

Now study the amended image below, and note the complete lack of contrast between Squares A and B and the added strip that joins them. Only upon seeing this second image are many people even able to consider the possibility that A and B might actually be the same shade of gray.



As the checker-shadow illusion dramatically illustrates, a statement that seems incontrovertibly true ("Square A is darker than Square B") may in fact be

¹⁹Richard Wiseman, "Ten of the Greatest Optical Illusions," *Daily Mail*, September 25, 2010.

false. This example should affirm at least the possibility that self-evidently true beliefs about the effect of higher taxes may be false as well.

When someone asks, “How will an event affect me?” the natural first step is to try to recall the effects of similar events in the past. When parents are trying to decide whether to take their children to Disney World, for example, they might try to summon memories of how well they had enjoyed past visits to similar theme parks. In like fashion, when high-income people try to imagine the impact of higher taxes, Plan A is to summon memories of how they felt in the wake of past tax increases.

But that strategy doesn’t work in the current era, since most high-income people alive today have experienced steadily declining tax rates. In WWII, the top marginal tax rate in the United States was 92 percent. In 1966, it was 70 percent, and in 1982, just 50 percent. Currently it is 37 percent. Apart from brief and isolated increases almost too small to notice, top marginal tax rates have fallen steadily throughout the lifetimes of voters born after WWII.

When Plan A fails, we go to Plan B. Because paying higher taxes means having less money to spend on other things, a plausible alternative cognitive strategy is to estimate the effect of tax hikes by recalling earlier events that resulted in lower disposable income—an occasional business reverse, for example, or a losing lawsuit, or a divorce, or a house fire, maybe even a health crisis. Rare is the life history that is completely devoid of events like these, which share a common attribute: They make people feel miserable.

More important, such events share a second feature, one that is absent from an increase in taxes: They reduce our own incomes while leaving others’ incomes unaffected. Higher taxes, in contrast, reduce all incomes in tandem. This difference holds the key to understanding why people think higher taxes would be so painful.

As most prosperous people would themselves be quick to concede, they have everything anybody might reasonably be said to need. If higher taxes pose any threat, it would be to make it more difficult for them to buy life’s special extras. But as we have seen, “special” is an inescapably relative concept. To be special means to stand out in some way from what is expected. And almost without exception, special things are in limited supply. There are only so many penthouse apartments with sweeping views of Central Park, for instance. To get one, a wealthy person must outbid peers who also want it. The outcomes of such bidding contests depend almost exclusively on relative purchasing power. And since relative purchasing power is completely unaffected when the wealthy all pay higher taxes, the same penthouses end up in the same hands as before.

A plausible objection is that higher tax rates on prosperous Americans would put them at a disadvantage relative to oligarchs from other countries in the bidding wars for trophy properties in the United States. But that disadvantage could be eliminated easily by the imposition of a purchase levy on nonresident buyers.

In short, the answer to the question of why voters haven’t supported the taxes necessary to finance desperately needed public investment may be that many suffer from the illusory belief that higher taxes would make it harder for them to buy what they want. Measured by its consequences for public welfare, it would be no exaggeration to call this false belief *the mother of all cognitive illusions*.

RECAP ↑

CONCERNS ABOUT RELATIVE POSITION

Evidence suggests that well-being depends not only on absolute levels of current and future consumption, but also on the corresponding levels of relative consumption.

Adding a taste for relative consumption alters many of the predictions and prescriptions of traditional economic models. Traditional models hold, for example, that the optimal amount of workplace safety will be provided in competitive labor markets. Critics respond that because labor markets are not effectively competitive, safety regulation is needed to prevent employers from exploiting workers. The puzzle for both camps is why safety regulations are most binding in those labor markets that, on traditional measures, are the most highly competitive.

Models that incorporate concerns about relative position address this puzzle by noting that by accepting riskier jobs, parents can use their higher pay to bid for housing in better school districts. But because school quality is an inherently relative concept, when other parents follow suit, the collective effect of their efforts is merely to bid up the prices of houses in better school districts. As before, half of all children end up attending bottom-half schools.

Workers confronting these incentives might well prefer an alternative state of the world in which all enjoyed greater safety, even at the expense of all having lower wages. Analogous reasoning suggests that individual incentives are misleadingly high for many other things we might sell for money—our leisure time, future job security, and a host of other environmental amenities. The problem is not that we exchange these items for money, but rather that we tend to sell them too cheaply.

SUMMARY

- Because it is costly to gather and analyze information, it often makes sense to employ rules of thumb when making complex assessments and decisions. Although the rules that most of us use appear to work reasonably well under a broad range of circumstance, they sometimes lead to error. (LO1)
- Behavioral economists have identified numerous systematic errors that are associated with the availability and representativeness heuristics. Anchoring and adjustment can also lead to biased assessments, as can failure to take account of regression to the mean or the tendency for extreme events to be followed by more nearly normal ones. It is also common for people to misinterpret contextual cues in their assessments and choices, as, for example, by reckoning the importance of a price reduction not by its absolute magnitude, but by its fraction of the original price. (LO1, LO2)
- Traditional economic models typically rule out the possibility that people might regret having chosen behaviors whose consequences were perfectly predictable, yet such expressions of regret appear genuine. Evidence suggests that people tend to discount future costs and benefits excessively leading them to choose imminent, though inferior, rewards instead of substantially larger rewards that require waiting. (LO3)
- Although traditional models assume that the reduction in happiness from a very small decline in wealth should be approximately the same as the gain in happiness from a similar gain in wealth, evidence suggests that losses weigh much more heavily than gains. Called loss aversion, this asymmetry in valuations is approximately two to one for minor changes, but can be substantially larger when important changes are at stake. Because every change in policy generates winners and losers, loss aversion implies a large bias in favor of the status quo in public policy decisions. Evidence suggests that this bias often has harmful consequences since people's initial estimates of how painful losses will be often fail to take into account their ability to adapt to altered circumstances. (LO4)
- Many traditional models assume that well-being depends only on absolute levels of current and future consumption. But evidence suggests that it also depends heavily on the corresponding levels of relative consumption. Adding concerns about relative position alters many of the predictions and prescriptions of traditional economic models. When well-being depends only on absolute levels of consumption and certain other conditions are met, traditional models hold that individual incentives produce socially optimal patterns

of spending. But when reward depends on relative position, individual incentives often result in mutually offsetting, and hence partially wasteful spending patterns. (LO5)

- The findings of behavioral economists in recent decades have contributed to our understanding of laws and other institutions. Societies around the world, for example, have adopted a variety of policies that are most easily understood as attempts to mitigate the consequences of impulse-control problems. These include special treatment of so-called crimes of passion; prohibitions against
- gambling, addictive drugs, and prostitution; entrapment laws; sanctions against adultery; and programs to stimulate savings. (LO6)
- Many other widely adopted laws, norms, and regulations are parsimoniously interpreted as attempts to mitigate collective-action problems caused by concerns about relative position. These include workplace safety regulations, limitations on work hours, and programs to stimulate savings, all of which reduce individual incentives to engage in mutually offsetting patterns of spending. (LO6)

KEY TERMS

anchoring and adjustment
availability heuristic
fungibility
homo economicus
judgmental and decision
heuristics

loss aversion
nonpositional good
positional good
regression to the mean
representativeness
heuristic

satisficing
status quo bias
Weber-Fechner law

REVIEW QUESTIONS

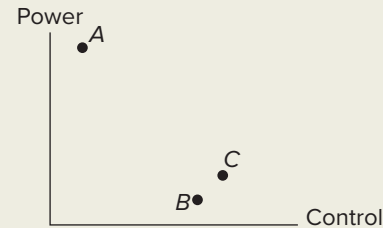
1. How does the representativeness heuristic explain why people might think, mistakenly, that a randomly chosen shy person is more likely to be a librarian than a salesperson? (LO1)
2. Explain why thinking of costs or benefits in proportional terms might lead to suboptimal choices. (LO2)
3. Explain why traditional economic models find it difficult to explain why people would pay to attend weight-loss camps that restrict their daily calorie intake. (LO3)
4. How does loss aversion help explain why attempts to repeal the Affordable Care Act have met such strong resistance? (LO4)
5. How does the concept of a positional good help explain why some people rush to purchase the latest iPhone even if the one they have is perfectly good? (LO5)
6. Cite two examples of laws, norms, and regulations that can be interpreted as attempts to mitigate collective-action problems caused by concerns about relative position. (LO6)

PROBLEMS



1. Only two moving companies, United and North American, provide local delivery service in a small western city. United operates 20 percent of the trucks, North American the remaining 80 percent. On a dark and rainy night, a pedestrian is run over and killed by a moving van. The lone witness to the incident testifies that the van was owned by United. An independent laboratory hired by the court finds that under dark and rainy conditions, the witness is able to identify the owner of a moving van with 90 percent accuracy. What is the probability that the van that struck the witness was owned by United? (LO1)
2. Kaela, a tennis player, has been struggling to develop a more consistent serve. She made the following remark to her partner during the second set of a recent match: "I feel like I'm making progress. I haven't double-faulted once today." She then served two double faults, which caused her to say, "Every time I say I haven't double-faulted, I immediately start to." Kaela's perception may have been influenced by (LO1)
 - a. the sunk cost effect.
 - b. regression to the mean.
 - c. the availability heuristic.

- d. More than one of the above are correct.
e. None of the above.
3. Studies have shown that in the New York City subways crime rates fall in the years following increased police patrols. Does this pattern suggest that the increased patrols are the cause of the crime reductions? (LO1)
4. Dalgliesh the detective fancies himself a shrewd judge of human nature. In careful tests, it has been discovered that he is right 80 percent of the time when he says that a suspect is lying. Dalgliesh says that Jones is lying. The polygraph expert, who is right 100 percent of the time, says that 40 percent of the subjects interviewed by Dalgliesh are telling the truth. What is the probability that Jones is lying? (LO1)
5. Shobhana is a gourmet. She makes it a point never to visit a restaurant a second time unless she has been served a superb meal on her first visit. She is puzzled at how seldom the quality of her second meal is as high as the first. Should she be? (LO1)
6. Maria will drive across town to take advantage of a 40 percent off sale on a \$40 blouse but will not do so to take advantage of a 10 percent off sale on a \$1,000 stereo. Assuming that her alternative is to pay list price for both products at the department store next to her home, is her behavior rational? (LO2)
7. Tom has said he would be willing to drive across town in order to save \$10 on the purchase price of a \$20 clock radio. If Tom is rational, this implies that he (LO2)
- believes that the opportunity cost of the trip is not more than \$10.
 - should also drive across town to save \$10 on a \$500 television set.
 - should not drive across town to save \$10 on a \$500 television set.
 - should drive across town only if the savings on a \$500 television set is at least \$250.
 - More than one of the above are correct.
8. Marco is having difficulty choosing between two tennis racquets, *A* and *B*. As shown in the diagram, *A* has more power than *B*, but less control. According to the rational choice model, how will the availability of a third alternative—racquet *C*—influence Marco's final decision? If Marco behaves like most ordinary decision makers in this situation, how will the addition of *C* to his choice set matter? (LO2)



9. In the fall, Crusoe puts 50 coconuts from his harvest into a cave just before a family of bears goes in to hibernate. As a result, he is unable to get the coconuts out before the bears emerge the following spring. Coconuts spoil at the same rate no matter where he stores them, and yet he continues this practice each year. Why might he do this? (LO3)
10. When students in a large class were surveyed about how much they would be willing to pay for a coffee mug with their university's logo on it, their median willingness to pay was \$5. At random, half of the students in this class were then given such a coffee mug, and each of the remaining students were given \$5 in cash. Students who got mugs were then offered an opportunity to sell them to students who had not gotten one. According to standard economic models, how many mugs would be expected to change hands? How, if at all, would a behavioral economist's prediction differ? (LO4)
11. Explain why rational voters whose happiness depends on relative consumption might favor laws that require them to save a certain portion of each year's earnings for retirement. (LO5, LO6)

ANSWERS TO SELF-TESTS

- 8.1 For every 100 taxis in a dark alley, 15 will be green, 85 blue. The witness will identify $0.8(15) = 12$ of the green taxis as green, the remaining 3 green taxis as blue; he will identify $0.8(85) = 68$ of the blue taxis as blue, the remaining 17 blue taxis as green. The probability that the cab in question was green, given that the witness said it was, is thus equal to $12/(12 + 17) = 0.413$ and as this is less than half, the Green Taxi Company should not be held liable. (LO1)
- 8.2 Regression to the mean suggests that a month with an unusually high number of burglaries is likely to be followed by a month in which the number of burglaries is more nearly normal. So the increased patrols were not necessarily the cause of the observed reduction in burglaries. (LO1)

9

LEARNING OBJECTIVES

After reading this chapter, you should be able to:

- LO1** Define negative and positive externalities and analyze their effect on resource allocation.
- LO2** Explain and discuss the Coase theorem.
- LO3** Explain how the effects of externalities can be remedied and discuss why the optimal amount of an externality is almost never zero.
- LO4** Illustrate the tragedy of the commons and show how private ownership is a way of preventing it.
- LO5** Define positional externalities and their effects and show how they can be remedied.

Externalities and Property Rights

A droll television ad for a British brand of pipe tobacco opens with a distinguished-looking gentleman sitting quietly on a park bench, smoking his pipe and reading a book of poetry. Before him lies a pond, unrippled except for a mother duck swimming peacefully with her ducklings. Suddenly a raucous group of teenage boys bursts onto the scene with a remote-controlled toy warship. Yelling and laughing, they launch their boat and maneuver it in aggressive pursuit of the terrified ducks.

Interrupted from his reverie, the gentleman looks up from his book and draws calmly on his pipe as he surveys the scene before him. He then reaches into his bag, pulls out a remote control of his own, and begins manipulating the joystick. The scene shifts underwater, where a miniature submarine rises from the depths of the pond. Once the boys' boat is in the sub's sights, the gentleman pushes a button on his remote control. Seconds later, the boat is blown to smithereens by a torpedo. The scene fades to a close-up of the tobacco company's label.

EXTERNAL COSTS AND BENEFITS

External costs (or **negative externality**) and **external benefits** (or **positive externality**)—**externalities**, for short—are activities that generate costs or benefits that accrue to people not directly involved in those activities. These effects are generally unintended. From the pipe smoker's point of view, the noise generated by the marauding boys was an external cost. Had others been disturbed by the boys' rowdiness, they may well have regarded the pipe smoker's retaliatory gesture as an external benefit.

This chapter focuses on how externalities affect the allocation of resources. Adam Smith's theory of the invisible hand applies to an ideal marketplace in which externalities do not exist. In such situations, Smith argued, the self-interested actions of individuals would lead to socially efficient outcomes. We will see that when the

external cost (or **negative externality**) a cost of an activity that falls on people other than those who pursue the activity

external benefit (or **positive externality**) a benefit of an activity received by people other than those who pursue the activity

externality an external cost or benefit of an activity

parties affected by externalities can easily negotiate with one another, the invisible hand will still produce an efficient outcome.

But in many cases, such as the scene depicted in the tobacco ad, negotiation is impractical. In those cases, the self-serving actions of individuals won't lead to efficient outcomes. The need to deal with externalities and other collective-action problems is one of the most important rationales for the existence of government.

How Externalities Affect Resource Allocation

The following examples illustrate the ways in which externalities distort the allocation of resources.

EXAMPLE 9.1

Positive Externalities

Does the honeybee keeper face the right incentives?, Part I

Phoebe earns her living as a keeper of honeybees. Her neighbors on all sides grow apples. Because bees pollinate apple trees as they forage for nectar, the more hives Phoebe keeps, the larger the harvests will be in the surrounding orchards. If Phoebe takes only her own costs and benefits into account in deciding how many hives to keep, will she keep the socially optimal number of hives?

Phoebe's hives constitute an external benefit, or a positive externality, for the orchard owners. If she takes only her own personal costs and benefits into account, she will add hives only until the added revenue she gets from the last hive just equals the cost of adding it. But because the orchard owners also benefit from additional hives, the total benefit of adding another hive at that point will be greater than its cost. Phoebe, then, will keep too few hives.

As we will discuss later in the chapter, problems like the one in Example 9.1 have several possible solutions. One is for orchard owners to pay beekeepers for keeping additional hives. But such solutions often require complex negotiations between the affected parties. For the moment, we assume that such negotiations are not practical.

EXAMPLE 9.2

Negative Externalities

Does the honeybee keeper face the right incentives?, Part 2

As in Example 9.1, Phoebe earns her living as a keeper of honeybees. Now her neighbors are not apple growers but an elementary school and a nursing home. The more hives Phoebe keeps, the more students and nursing home residents will be stung by bees. If Phoebe takes only her own costs and benefits into account in deciding how many hives to keep, will she keep the socially optimal number of hives?

For the students and nursing home residents, Phoebe's hives constitute an external cost, or a negative externality. If she considers only her own costs and benefits in deciding how many hives to keep, she will continue to add hives until the added revenue from the last hive is just enough to cover its cost. But since Phoebe's neighbors also incur costs when she adds a hive, the benefit of the last hive at that point will be smaller than its cost. Phoebe, in other words, will keep too many hives.

Every activity involves costs and benefits. When all the relevant costs and benefits of an activity accrue directly to the person who carries it out—that is, when the activity generates no externalities—the level of the activity that is best for the individual will be best for society as a whole. But when an activity generates externalities, be they positive or negative, individual self-interest does not produce the best allocation of resources. Individuals who consider only their own costs and benefits will tend to engage too much in activities that generate negative externalities and too little in activities that generate positive externalities. When an activity generates both positive and negative externalities, private and social interests will coincide only in the unlikely event that the opposing effects offset one another exactly.

THE COASE THEOREM

To say that a situation is inefficient means that it can be rearranged in a way that would make at least some people better off without harming others. Such situations, we have seen, are a source of creative tension. The existence of inefficiency, after all, means that there is cash on the table, which usually triggers a race to see who can capture it. For example, we saw that because monopoly pricing results in an inefficiently low output level, the potential for gain gave monopolists an incentive to make discounts available to price-sensitive buyers. As the next examples illustrate, the inefficiencies that result from externalities create similar incentives for remedial action.

EXAMPLE 9.3 Inefficiencies That Result from Externalities

Will Abercrombie dump toxins in the river?, Part I

Abercrombie's factory produces a toxic waste by-product. If Abercrombie dumps it in the river, he causes damage to Fitch, a fisherman located downstream. The toxins are short-lived and cause no damage to anyone other than Fitch. At a cost, Abercrombie can filter out the toxins, in which case Fitch will suffer no damage at all. The relevant gains and losses for the two individuals are listed in Table 9.1.

TABLE 9.1
Costs and Benefits of Eliminating Toxic Waste, Part 2

	With filter	Without filter
Gains to Abercrombie	\$100/day	\$130/day
Gains to Fitch	\$100/day	\$ 50/day

If the law does not penalize Abercrombie for dumping toxins in the river, and if Abercrombie and Fitch cannot communicate with one another, will Abercrombie operate with or without a filter? Is that choice socially efficient?

Abercrombie has an incentive to operate without a filter since he earns \$30 per day more than if he operates with a filter. But the outcome when he does so is socially inefficient. Thus, when Abercrombie operates without a filter, the total daily gain to both parties is only $\$130 + \$50 = \$180$, compared to $\$100 + \$100 = \$200$ if Abercrombie had operated with a filter. The daily cost of the filter to Abercrombie is only $\$130 - \$100 = \$30$, which is smaller than its daily benefit to Fitch of $\$100 - \$50 = \$50$. The fact that Abercrombie does not install the filter implies a squandered daily surplus of \$20.

EXAMPLE 9.4 **Negotiating Efficient Solutions to Externalities****Will Abercrombie dump toxins in the river?, Part 2**

Suppose the costs and benefits of using the filter are as in the previous example except that Abercrombie and Fitch can now communicate with one another at no cost. Even though the law does not require him to do so, will Abercrombie use a filter?

This time, Abercrombie will use a filter. Recall from Chapter 5, *Efficiency, Exchange, and the Invisible Hand*, the observation that when the economic pie grows larger, everyone can have a larger slice. Because use of a filter would result in the largest possible economic surplus, it would enable both Abercrombie and Fitch to have a larger net gain than before. Fitch thus has an incentive to *pay* Abercrombie to use a filter. Suppose, for instance, that Fitch offers Abercrombie \$40 per day to compensate him for operating with a filter. Both Abercrombie and Fitch will then be exactly \$10 per day better off than before, for a total daily net gain of \$20.

SELF-TEST 9.1

In Example 9.4, what is the largest whole-dollar amount by which Fitch could compensate Abercrombie for operating with a filter and still be better off than before?

Coase theorem if, at no cost, people can negotiate the purchase and sale of the right to perform activities that cause externalities, they can always arrive at efficient solutions to the problems caused by externalities

Ronald Coase, a professor at the University of Chicago Law School, was the first to see clearly that if people can negotiate with one another at no cost over the right to perform activities that cause externalities, they will always arrive at an efficient solution. This insight, which is often called the **Coase theorem**, is a profoundly important idea, for which Coase (rhymes with “rose”) was awarded the 1991 Nobel Prize in Economics.

Why, you might ask, should Fitch pay Abercrombie to filter out toxins that would not be there in the first place if not for Abercrombie’s factory? The rhetorical force of this question is undeniable. Yet Coase points out that externalities are reciprocal in nature. The toxins do harm Fitch, to be sure, but preventing Abercrombie from emitting them would penalize Abercrombie, by exactly \$30 per day. Why should Fitch necessarily have the right to harm Abercrombie? Indeed, as the next example illustrates, even if Fitch had that right, he would exercise it only if filtering the toxins proved the most efficient outcome.

EXAMPLE 9.5 **Social Efficiency****Will Abercrombie dump toxins in the river?, Part 3**

Suppose the law says that Abercrombie may *not* dump toxins in the river unless he has Fitch’s permission. If the relevant costs and benefits of filtering the toxins are as shown in Table 9.2, and if Abercrombie and Fitch can negotiate with one another at no cost, will Abercrombie filter out the toxins?

TABLE 9.2**Costs and Benefits of Eliminating Toxic Waste, Part 3**

	With filter	Without filter
Gains to Abercrombie	\$100/day	\$150/day
Gains to Fitch	\$100/day	\$ 70/day

Note that this time, the most efficient outcome is for Abercrombie to operate without a filter because the total daily surplus in that case will be \$220 as compared to only \$200 with a filter. Under the law, however, Fitch has the right to insist that Abercrombie use a filter. We might expect him to exercise that right since his own gain would rise from \$70 to \$100 per day if he did so. But because this outcome would be socially inefficient, we know that each party can do better.

Suppose, for example, that Abercrombie gives Fitch \$40 per day in return for Fitch's permission to operate without a filter. Each would then have a net daily gain of \$110, which is \$10 better for each of them than if Fitch had insisted that Abercrombie use a filter. Abercrombie's pollution harms Fitch, sure enough. But failure to allow the pollution would have caused even greater harm to Abercrombie.

The Coase theorem tells us that regardless of whether the law holds polluters liable for damages, the affected parties will achieve efficient solutions to externalities if they can negotiate costlessly with one another. Note carefully that this does not imply that affected parties will be indifferent about whether the law holds polluters responsible for damages. If polluters are liable, they will end up with lower incomes and those who are injured by pollutants will end up with higher incomes than if the law does not hold polluters liable—even though the same efficient production methods would be adopted in each case. When polluters are held liable, they must remove the pollution at their own expense. When they are not held liable, those who are injured by pollution must pay polluters to cut back.

Externalities are hardly rare and isolated occurrences. On the contrary, finding examples of actions that are altogether free of them is difficult. And because externalities can distort the allocation of resources, it is important to recognize them and deal intelligently with them. Consider the following example of an externality that arises because of shared living arrangements.

EXAMPLE 9.6**Cost-Benefit Principle—Shared Living Expenses****Will Adriana and Sofia share an apartment?**

Adriana and Sofia can live together in a two-bedroom apartment for \$1,000 per month, or separately in 2 one-bedroom apartments, each for \$600 per month. If the rent paid were the same for both alternatives, the two women would be indifferent between living together or separately, except for one problem: Adriana often has friends over late at night when Sofia is trying to sleep. Adriana would pay up to \$250 per month for this privilege. Sofia, for her part, would pay up to \$150 per month to consistently get a good night's sleep. Should they live together or separately?

Adriana and Sofia should live together only if the benefit of doing so exceeds the cost. The benefit of living together is the reduction in their rent. Since 2 one-bedroom apartments would cost a total of \$1,200 per month, compared to \$1,000 for a two-bedroom unit, their benefit from living together is \$200 per month. Their cost of living together is the least costly accommodation they can make to Adriana's objectionable habits. Adriana would be willing to pay up to \$250 per month to avoid changing her behavior, so the \$200 rent saving is too small to persuade her to change. But Sofia is willing to put up with Adriana's behavior for a compensation payment of only \$150 per month. Since that amount is smaller than the total saving in rent, the least costly solution to the problem is for Sofia to live with Adriana and simply put up with her behavior.

Table 9.3 summarizes the relevant costs and benefits of this shared living arrangement. The Cost-Benefit Principle tells us that Adriana and Sofia should live together if and only if the benefit of living together exceeds the cost. The cost of the shared living arrangement is not the sum of all possible costs but the least costly accommodation to the problem (or problems) of shared living. Since the \$200 per month saving in rent exceeds the least costly accommodation to the phone problem, Adriana and Sofia can reap a total gain in economic surplus of \$50 per month by sharing their living quarters.

TABLE 9.3
The Gain in Surplus from Shared Living Arrangements

Benefits of Shared Living			
Total cost of separate apartments	Total cost of shared apartment		Rent savings from sharing
2 × \$600/month \$1,200/month	\$1,000/month		\$200/month
Costs of Shared Living			
Problem	Adriana's cost of solving problem	Sofia's cost of solving problem	Least costly solution to the problem
Adriana inviting friends over late at night	Limiting when her friends come over: \$250/month	Tolerating the noise at night: \$150/month	Sofia tolerates the noise made by Adriana's friends: \$150/month
Gain in Surplus from Shared Living			
Rent savings (\$200/month)	—	Least costly accommodation to shared living problems (\$150/month)	= Gain in surplus: (\$50/month)

Some people might conclude that Adriana and Sofia should not live together because if the two share the rent equally, Sofia would end up paying \$500 per month—which when added to the \$150 cost of putting up with Adriana's behavior comes to \$50 more than the cost of living alone. As persuasive as that argument may sound, however, it is mistaken. The source of the error, as the

following example illustrates, is the assumption that the two must share the rent equally.

EXAMPLE 9.7 Cost-Benefit Principle—Paying Unequal Rent Amounts

What is the highest rent Sofia would be willing to pay for the two-bedroom apartment?

In Example 9.6, Sofia's alternative is to live alone, which would mean paying \$600 per month, her reservation price for a living arrangement in which she is not woken up by Adriana's friends. Since the most she would be willing to pay to avoid this problem is \$150 per month, the highest monthly rent she would be willing to pay for the shared apartment is $\$600 - \$150 = \$450$. If she pays that amount, Adriana will have to pay the difference, namely, \$550 per month, which is clearly a better alternative for Adriana than paying \$600 to live alone.

EXAMPLE 9.8 Cost-Benefit Principle—Splitting Economic Surplus

How much should Adriana and Sofia pay if they agree to split their economic surplus equally?

As we saw in Table 9.3, the total rent saving from the shared apartment is \$200, and since the least costly solution to the problem is \$150, the monthly gain in economic surplus is \$50. We know from Example 9.7 that Adriana's reservation price for living together is \$600 per month and Sofia's is \$450. So if the two women want to split the \$50 monthly surplus equally, each should pay \$25 less than her reservation price. Adriana's monthly rent will thus be \$575 and Sofia's, \$425. The result is that each is \$25 per month better off than if she had lived alone.

SELF-TEST 9.2

As in Examples 9.6 and 9.7, Adriana and Sofia can live together in a two-bedroom apartment for \$1,000 per month or separately in 2 one-bedroom apartments, each for \$600 per month. Adriana would pay up to \$250 per month rather than moderate her social habits, and Sofia would pay up to \$150 per month to consistently get a good night's sleep. Now, suppose Sofia would also be willing to pay up to \$60 per month to avoid the loss of privacy that comes with shared living space. Should the two women live together?

REMEDIES FOR EXTERNALITIES

Laws and Regulations

We have seen that efficient solutions to externalities can be found whenever the affected parties can negotiate with one another at no cost. But negotiation is not always practical. A motorist with a noisy muffler imposes costs on others, yet they cannot

flag him down and offer him a compensation payment to fix his muffler. In recognition of this difficulty, most governments simply require that cars have working mufflers. Indeed, the explicit or implicit purpose of a large share—perhaps the lion's share—of laws is to solve problems caused by externalities. The goal of such laws is to help people achieve the solutions they might have reached had they been able to negotiate with one another.

When negotiation is costless, the task of adjustment generally falls on the party who can accomplish it at the lowest cost. For instance, in our examples, Sofia put up with Adriana's annoying habits because doing so was less costly than asking Adriana to change her habits. Many municipal noise ordinances also place the burden of adjustment on those who can accomplish it at the lowest cost. Consider, for example, the restrictions on loud party music, which often take effect at a later hour on weekends than on weekdays. This pattern reflects both the fact that the gains from loud music tend to be larger on weekends and the fact that such music is more likely to disturb people on weekdays. By setting the noise curfew at different hours on different days of the week, the law places the burden on partygoers during the week and on sleepers during the weekend. Similar logic explains why noise ordinances allow motorists to honk their horns in most neighborhoods, but not in the immediate vicinity of a hospital.

The list of laws and regulations that may be fruitfully viewed as solutions to externalities is a long one. When a motorist drives his car at high speed, he endangers not just his own life and property, but also the lives and property of others. Speed limits, no-passing zones, right-of-way rules, and a host of other traffic laws may be seen as reasoned attempts to limit the harm one party inflicts on another. Many jurisdictions even have laws requiring that motorists install snow tires on their cars by the first of November. These laws promote not just safety, but also the smooth flow of traffic: If one motorist can't get up a snow-covered hill, he delays not only himself, but also the motorists behind him.

Similar reasoning helps us understand the logic of zoning laws that restrict the kinds of activities that take place in various parts of cities. Because many residents place a high value on living in an uncongested neighborhood, some cities have enacted zoning laws specifying minimum lot sizes. In places like Manhattan, where a shortage of land encourages developers to build very large and tall buildings, zoning laws limit both a building's height and the proportion of a lot it may occupy. Such restrictions recognize that the taller a building is, and the greater the proportion of its lot that it occupies, the more it blocks sunlight from reaching surrounding properties. The desire to control external costs also helps explain why many cities establish separate zones for business and residential activity. Even within business districts, many cities limit certain kinds of commercial activity. For example, in an effort to revitalize the Times Square neighborhood, New York City enacted a zoning law banning adult bookstores and pornographic movie theaters from the area.

Limitations on the discharge of pollutants into the environment are perhaps the clearest examples of laws aimed at solving problems caused by externalities. The details of these laws reflect the principle of placing the burden of adjustment on those who can accomplish it at least cost. The discharge of toxic wastes into rivers, for example, tends to be most strictly regulated on those waterways whose commercial fishing or recreational uses are most highly valued. On other waterways, the burden of adjustment is likely to fall more heavily on fishermen, recreational boaters, and swimmers. Similarly, air-quality regulations tend to be strictest in the most heavily populated regions of the country, where the marginal benefit of pollution reduction is the greatest.

The following example suggests an additional way in which Coase's insights about how societies deal with externalities provide rich fodder for the economic naturalist.

The Economic Naturalist 9.1



What is the purpose of free speech laws?

The First Amendment's protection of free speech and the pattern of exceptions to that protection are another illustration of how legal remedies are used to solve the problems caused by externalities. The First Amendment acknowledges the decisive value of open communication, as well as the practical difficulty of identifying and regulating acts of speech that cause more harm than good. Yet there are some important exceptions. The Supreme Court has ruled, for instance, that the First Amendment does not allow someone to yell "fire" in a crowded theater if there is no fire, nor does it allow someone to advocate the violent overthrow of the government. In those instances, the external benefits of free speech are far too small to justify the external costs.



Why does the U.S. Constitution protect the right of free speech?

The Optimal Amount of Negative Externalities Is Not Zero

Curbing pollution and other negative externalities entails both costs and benefits. The best policy is to curtail pollution until the marginal cost of further abatement just equals the marginal benefit. In general, the marginal cost of abatement rises with the amount of pollution eliminated, since polluters use the cheapest cleanup methods first and then turn to more expensive ones. And the law of diminishing marginal utility suggests that beyond some point, the marginal benefit of pollution reduction tends to fall as more pollution is removed. As a result, the marginal cost and marginal benefit curves almost always intersect at less than the maximum amount of pollution reduction.

The intersection of the two curves marks the socially optimal level of pollution reduction. If pollution is curtailed by any less than that amount, society will gain more than it will lose by pushing the cleanup effort a little further. But if regulators push beyond the point at which the marginal cost and benefit curves intersect, society will incur costs that exceed the benefits. The existence of a socially optimal level of pollution reduction implies the existence of a socially optimal level of pollution, and that level will almost always be greater than zero.

Because people have been conditioned to think of pollution as bad, many cringe when they hear the phrase "socially optimal level of pollution." How can any positive level of pollution be socially optimal? *But to speak of a socially optimal level of pollution is not the same as saying that pollution is good.* It is merely to recognize that society has an interest in cleaning up the environment, but only up to a certain point. The underlying idea is no different from the idea of

an optimal level of dirt in an apartment. After all, even if you spent the whole day, every day, vacuuming your apartment, there would be *some* dirt left in it. And because you have better things to do than vacuum all day, you probably tolerate substantially more than the minimal amount of dirt. A dirty apartment is not good, nor is pollution in the air you breathe. But in both cases, the cleanup effort should be expanded only until the marginal benefit equals the marginal cost.

Compensatory Taxes and Subsidies

As noted, when transaction costs prohibit negotiation among affected parties, negative externalities lead to excessive output levels because activities that produce negative externalities are misleadingly attractive to those who engage in them. One solution to this problem, proposed by the British economist A. C. Pigou, is to make such activities less attractive by taxing them.

Although many critics insist that taxes always reduce economic efficiency, a tax can actually make the economy *more* efficient. A tax can have that effect because it forces producers to take explicit account of the fact that each additional unit of output they produce imposes an external cost on the rest of society.

Similar reasoning suggests that a subsidy to producers can serve to counteract misallocations that result from positive externalities.

A subsidy makes the economy more efficient because it induces producers to take account of a relevant benefit that they otherwise would have ignored.



The Economic Naturalist 9.2



Why does the government subsidize scientific research?

Why does the government subsidize private property owners to plant trees on their hillsides?

Societies use laws not only to discourage activities that generate negative externalities, but also to encourage activities that generate positive externalities. The planting of trees on hillsides, for example, benefits not just the landowner, but also his neighbors by limiting the danger of flooding. In recognition of this fact, many jurisdictions subsidize the planting of trees. Similarly, Congress budgets millions of dollars each year in support of basic research—an implicit acknowledgment of the positive externalities associated with the generation of new knowledge.

RECAP ↑

EXTERNAL COSTS AND BENEFITS

Externalities occur when the costs or benefits of an activity accrue to people other than those directly involved in the activity. The Coase theorem says that when affected parties can negotiate with one another without cost, activities will be pursued at efficient levels, even in the presence of positive or negative externalities. But when negotiation is prohibitively costly, inefficient behavior generally results. Activities that generate negative externalities are pursued to excess, while those that generate positive externalities are pursued too little. Laws and regulations, including taxes and subsidies, are often adopted in an effort to alter inefficient behavior that results from externalities.

PROPERTY RIGHTS AND THE TRAGEDY OF THE COMMONS

People who grow up in industrialized nations tend to take the institution of private property for granted. Our intuitive sense is that people have the right to own any property they acquire by lawful means and to do with that property as they see fit. In reality, however, property laws are considerably more complex in terms of the rights they confer and the obligations they impose.

The Problem of Unpriced Resources

To understand the laws that govern the use of property, let's begin by asking why societies created the institution of private property in the first place. The following examples, which show what happens to property that nobody owns, suggest an answer.

EXAMPLE 9.9 Individual Income

How many steers will villagers send onto the commons?

A village has five residents, each of whom has accumulated savings of \$100. Each villager can use the money to buy a government bond that pays 13 percent interest per year or to buy a year-old steer, send it onto the commons to graze, and sell it after 1 year. The price the villager will get for the 2-year-old steer depends on the amount of weight it gains while grazing on the commons, which in turn depends on the number of steers sent onto the commons, as shown in Table 9.4.

TABLE 9.4
The Relationship between Herd Size and Steer Price

Number of steers on the commons	Price per 2-year-old steer (\$)	Income per steer (\$/year)
1	126	26
2	119	19
3	116	16
4	113	13
5	111	11

The price of a 2-year-old steer declines with the number of steers grazing on the commons because the more steers, the less grass available to each. The villagers make their investment decisions one at a time, and the results are public. If each villager decides how to invest individually, how many steers will be sent onto the commons, and what will be the village's total income?

If a villager buys a \$100 government bond, he'll earn \$13 of interest income at the end of 1 year. Thus, he should send a steer onto the commons if and only if that steer will command a price of at least \$113 as a 2-year-old. When each villager chooses in this self-interested way, we can expect four villagers to send a steer onto the commons. (Actually, the fourth villager would be indifferent between investing in a steer or buying a bond since he would earn \$13 either way. For the sake of discussion, we'll assume that in the case of a tie, people choose to be cattle owners.) The fifth villager, seeing that he would earn only \$11 by sending a fifth steer onto the commons, will choose instead to buy a government bond. As a result of these decisions, the total village income will be \$65 per year—\$13 for the one bondholder and $4(\$13) = \52 for the four cattle owners.

Has Adam Smith’s invisible hand produced the most efficient allocation of these villagers’ resources? We can tell at a glance that it has not, since their total village income is only \$65—precisely the same as it would have been had the possibility of cattle raising not existed. The source of the difficulty will become evident in the following example.

EXAMPLE 9.10 Maximizing Total Group Income

What is the socially optimal number of steers to send onto the commons?

Suppose the five villagers in the previous example confront the same investment opportunities as before, except that this time they are free to make their decisions as a group rather than individually. How many steers will they send onto the commons, and what will be their total village income?

This time the villagers’ goal is to maximize the income received by the group as a whole. When decisions are made from this perspective, the criterion is to send a steer onto the commons only if its marginal contribution to village income is at least \$13, the amount that could be earned from a government bond. As the entries in column 5 indicate, the first steer clearly meets this criterion since it contributes \$26 to total village income. But the second steer does not. Sending that steer onto the commons raises the village’s income from cattle raising from \$26 to \$38, a gain of just \$12. The \$100 required to buy the second steer would thus have been better invested in a government bond. Worse, the collective return from sending a third steer is only \$10; from a fourth, only \$4; and from a fifth, only \$3.

TABLE 9.5
Marginal Income and the Socially Optimal Herd Size

(1) Number of steers on the commons	(2) Price per 2-year-old steer (\$)	(3) Income per steer (\$/year)	(4) Total village income (\$/year)	(5) Marginal income (\$/year)
				26
1	126	26	26	12
2	119	19	38	10
3	116	16	48	4
4	113	13	52	3
5	111	11	55	

In sum, when investment decisions are made with the goal of maximizing total village income, the best choice is to buy four government bonds and send only a single steer onto the commons. The resulting village income will be \$78: \$26 from sending the single steer and \$52 from the four government bonds. That amount is \$13 more than the total income that resulted when villagers made their investment decisions individually. Once again, the reward from moving from an inefficient allocation to an efficient one is that the economic pie grows larger. And when the pie grows larger, everyone can get a larger slice. For instance, if the villagers agree to pool their income and share it equally, each will get \$15.60, or \$2.60 more than before.

SELF-TEST 9.3

How would your answers to Examples 9.9 and 9.10 change if the interest rate were 11 percent per year rather than 13 percent?

Why do the villagers in Examples 9.9 and 9.10 do better when they make their investment decisions collectively? The answer is that when individuals decide alone, they ignore the fact that sending another steer onto the commons will cause existing steers to gain less weight. Their failure to consider this effect makes the return from sending another steer seem misleadingly high to them.

The grazing land on the commons is a valuable economic resource. When no one owns it, no one has any incentive to take the opportunity cost of using it into account. And when that happens, people will tend to use it until its marginal benefit is zero. This problem, and others similar to it, is known as the **tragedy of the commons**. The essential cause of the tragedy of the commons is the fact that one person's use of commonly held property imposes an external cost on others by making the property less valuable. Each individual villager behaves rationally by sending an additional steer onto the commons, yet the overall outcome falls far short of the attainable ideal.

tragedy of the commons

the tendency for a resource that has no price to be used until its marginal benefit falls to zero

The Effect of Private Ownership

As the following example illustrates, one solution to the tragedy of the commons is to place the village grazing land under private ownership.

EXAMPLE 9.11 Private Ownership**How much will the right to control the village commons sell for?**

Suppose the five villagers face the same investment opportunities as before, except that this time they decide to auction off the right to use the commons to the highest bidder. Assuming that villagers can borrow as well as lend at an annual interest rate of 13 percent, what price will the right to use the commons fetch? How will the owner of that property right use it, and what will be the resulting village income?

To answer these questions, simply ask yourself what you would do if you had complete control over how the grazing land were used. As we saw earlier, the most profitable way to use this land is to send only a single steer to graze on it. If you do so, you will earn a total of \$26 per year. Since the opportunity cost of the \$100 you spent on the single yearling steer is the \$13 in interest you could have earned from a bond, your economic profit from sending a single steer onto the commons will be \$13 per year, provided you can use the land for free. But you cannot; to finance your purchase of the property right, you must borrow money (since you used your \$100 savings to buy a year-old steer).

What is the most you would be willing to pay for the right to use the commons? Its use generates an income of \$26 per year—or \$13 more than the opportunity cost of your investment in the steer—so the most you would pay is \$100 (since that amount used to purchase a bond that pays 13 percent interest would also generate income of \$13 per year). If the land were sold at auction, \$100 is precisely the amount you would have to pay. Your annual earnings from the land would be exactly enough to pay the \$13 interest on your loan and cover the opportunity cost of not having put your savings into a bond.

Note that when the right to use the land is auctioned to the highest bidder, the village achieves a more efficient allocation of its resources because the owner has a strong incentive to take the opportunity cost of more intensive grazing fully into account. Total village income in this case will again be \$78. If the annual interest on the \$100 proceeds from selling the land rights is shared equally among the five villagers, each will again have an annual investment income of \$15.60.

The logic of economic surplus maximization helps explain why the most economically successful nations have all been ones with well-developed private property laws. Property that belongs to everyone belongs, in effect, to no one. Not only is its potential economic value never fully realized; it usually ends up being of no value at all.

Bear in mind, however, that in most countries the owners of private property are not free to do *precisely* as they wish with it. For example, local zoning laws may give the owner of a residential building lot the right to build a three-story house but not a taller one. Here, too, the logic of economic surplus maximization applies, for a fully informed and rational legislature would define property rights so as to create the largest possible total economic surplus. In practice, of course, such ideal legislatures never really exist. Yet the essence of politics is the cutting of deals that make people better off. If a legislator could propose a change in the property laws that would enlarge the total economic surplus, she could also propose a scheme that would give each of her constituents a larger slice, thus enhancing her chances for reelection.

As an economic naturalist, challenge yourself to use this framework when thinking about the various restrictions you encounter in private property laws: zoning laws that constrain what you can build and what types of activities you can conduct on your land; traffic laws that constrain what you can do with your car; employment and environmental laws that constrain how you can operate your business. Your understanding of these and countless other laws will be enhanced by the insight that everyone can gain when the private property laws are defined so as to create the largest total economic surplus.

When Private Ownership Is Impractical

Don't be misled into thinking that the law provides an *ideal* resolution of all problems associated with externalities and the tragedy of the commons. Defining and enforcing efficient property rights entail costs, after all, and sometimes, as in the following examples, the costs outweigh the gains.



The Economic Naturalist 9.3

Why do blackberries in public parks get picked too soon?

Wild blackberries grow profusely at the edge of a wooded area in a crowded city park. The blackberries will taste best if left to ripen fully, but they still taste reasonably good if picked and eaten a few days early. Will the blackberries be left to ripen fully?

Obviously, the costs of defining and enforcing the property rights to blackberries growing in a public park are larger than the potential gains, so the blackberries will remain common property. That means that whoever picks them first gets them. Even though everyone would benefit if people waited until the berries were fully ripe, everyone knows that those who wait are likely to end up with no berries at all. And that means that the berries will be eaten too soon.



► Visit your instructor's Connect course and access your eBook to view this video.



Why do blackberries in public parks get picked before they're completely ripe?

The Economic Naturalist 9.4

Why are shared milkshakes consumed too quickly?

Sara and Susan are identical twins who have been given a chocolate milkshake to share. If each has a straw and each knows that the other is self-interested, will the twins consume the milkshake at an optimal rate?

Because drinking a milkshake too quickly chills the taste buds, the twins will enjoy their shake more if they drink it slowly. Yet each knows that the other will drink any part of the milkshake she doesn't finish herself. The result is that each will consume the shake at a faster rate than she would if she had half a shake all to herself.



Why are shared milkshakes drunk too quickly?

Here are some further examples in which the tragedy of the commons is not easily solved by defining private ownership rights.

Harvesting Timber on Remote Public Land

On remote public lands, enforcing restrictions against cutting down trees may be impractical. Each tree cutter knows that a tree that is not harvested this year will be bigger, and hence more valuable, next year. But he also knows that if he doesn't cut the tree down this year, someone else might do so. In contrast, private companies that grow trees on their own land have no incentive to harvest timber prematurely and a strong incentive to prevent outsiders from doing so.

Harvesting Whales in International Waters

Each individual whaler knows that harvesting an extra whale reduces the breeding population, and hence the size of the future whale population. But the whaler also knows that any whale that is not harvested today may be taken by some other whaler. The solution would be to define and enforce property rights to whales. But the oceans are vast, and the behavior of whalers is hard to monitor. And even if their behavior could be monitored, the concept of national sovereignty would make the international enforcement of property rights problematic.

More generally, the animal species that are most severely threatened with extinction tend to be those that are economically valuable to humans but that are not privately owned by anyone. This is the situation confronting whales as well as elephants. Contrast this with the situation confronting chickens, which are also economically valuable to humans but which, unlike whales, are governed by traditional laws of private property. This difference explains why no one worries that Colonel Sanders might threaten the extinction of chickens.

Controlling Multinational Environmental Pollution

Each individual polluter may know that if he and all others pollute, the damage to the environment will be greater than the cost of not polluting. But if the environment is common property into which all are free to dump, each has a powerful incentive to pollute. Enforcing laws and regulations that limit the discharge of pollution may be practical if all polluters live under the jurisdiction of a single government. But if polluters come from many different countries, solutions are much more difficult to implement. Thus, the Mediterranean Sea has long suffered serious pollution since none of the many nations that border it has an economic incentive to consider the effects of its discharges on other countries.

As the world's population continues to grow, the absence of an effective system of international property rights will become an economic problem of increasing significance.

RECAP ↑

PROPERTY RIGHTS AND THE TRAGEDY OF THE COMMONS

When a valuable resource has a price of zero, people will continue to exploit it as long as its marginal benefit remains positive. The tragedy of the commons describes situations in which valuable resources are squandered because users are not charged for them. In many cases, an efficient remedy is to define and enforce rights to the use of valuable property. But this solution is difficult to implement for resources such as the oceans and the atmosphere because no single government has the authority to enforce property rights for these resources.

POSITIONAL EXTERNALITIES

Former tennis champion Steffi Graf received more than \$1.6 million in tournament winnings in 1992; her endorsement and exhibition earnings totaled several times that amount. By any reasonable measure, the quality of her play was outstanding, yet she consistently lost to archrival Monica Seles. But in April 1993, Seles was stabbed in the back by a deranged fan and forced to withdraw from the tour. In the ensuing months, Graf's tournament winnings accumulated at almost double her 1992 pace, despite little change in the quality of her play.

Payoffs That Depend on Relative Performance

In professional tennis and a host of other competitive situations, the rewards people receive typically depend not only on how they perform in absolute terms but also on how they perform relative to their closest rivals. In these situations, competitors have an incentive to take actions that will increase their odds of winning. For example, tennis players can increase their chances of winning by hiring personal fitness trainers and sports psychologists to travel with them on the tour. Yet the simple mathematics of competition tells us that the sum of all individual payoffs from such investments will be larger than the collective payoff. In any tennis match, for example, each contestant will get a sizable payoff from money spent on fitness trainers and sports psychologists, yet each match will have exactly one winner and one loser, no matter how much players spend. The overall gain to tennis spectators is likely to be small, and the overall gain to players as a group must be zero. To the extent that each contestant's payoff depends on his or her relative performance, then, the incentive to undertake such investments will be excessive, from a collective point of view.

Consider the following example.



The Economic Naturalist 9.5

Why do football players take anabolic steroids?

The offensive linemen of many National Football League teams currently average more than 330 pounds. In the 1970s, by contrast, offensive linemen in the league averaged barely 280 pounds, and the all-decade linemen of the 1940s averaged only 229 pounds. One reason that today's players are so much heavier is that

players' salaries have escalated sharply over the last several decades, which has intensified competition for the positions. Size and strength are the two cardinal virtues of an offensive lineman, and other things being equal, the job will go to the larger and stronger of two rivals.

Size and strength, in turn, can be enhanced by the consumption of anabolic steroids. But if all players consume these substances, the rank ordering of players by size and strength—and hence the question of who lands the jobs—will be largely unaffected. And because the consumption of anabolic steroids entails potentially serious long-term health consequences, football players as a group are clearly worse off if they consume these drugs. So why do football players take steroids?

The problem here is that contestants for starting berths on the offensive line confront a prisoner's dilemma, like the ones analyzed in Chapter 7, *Games and Strategic Behavior*. Consider two closely matched rivals—Smith and Jones—who are competing for a single position. If neither takes steroids, each has a 50 percent chance of winning the job and a starting salary of \$1 million per year. If both take steroids, each again has a 50 percent chance of winning the job. But if one takes steroids and the other doesn't, the first is sure to win the job. The loser ends up selling insurance for \$60,000 per year. Neither likes the fact that the drugs may have adverse health consequences, but each would be willing to take that risk in return for a shot at the big salary. Given these choices, the two competitors face a payoff matrix like the one shown in Table 9.6.



Why do so many football players take steroids?

TABLE 9.6
Payoff Matrix for Steroid Consumption

		Jones	
		Don't take steroids	Take steroids
Smith	Don't take steroids	Second best for each	Best for Jones Worst for Smith
	Take steroids	Best for Smith Worst for Jones	Third best for each

Clearly, the dominant strategy for both Smith and Jones is to take steroids. Yet when they do, each gets only the third-best outcome, whereas they could have gotten the second-best outcome by not taking the drugs—hence the attraction of rules that forbid the consumption of anabolic steroids.

positional externality this occurs when an increase in one person's performance reduces the expected reward of another's in situations in which reward depends on relative performance

positional arms race a series of mutually offsetting investments in performance enhancement that is stimulated by a positional externality

positional arms control agreement an agreement in which contestants attempt to limit mutually offsetting investments in performance enhancement

Positional Arms Races and Positional Arms Control Agreements

The steroid problem is an example of a **positional externality**. Whenever the payoffs to one contestant depend at least in part on how he or she performs relative to a rival, any step that improves one side's relative position must necessarily worsen the other's. The shouting-at-parties example discussed in Chapter 7, *Games and Strategic Behavior*, is another instance of a positional externality. Just as the invisible hand of the market is weakened by the presence of standard externalities, it is also weakened by positional externalities.

We have seen that positional externalities often lead contestants to engage in an escalating series of mutually offsetting investments in performance enhancement. We call such spending patterns **positional arms races**.

Because positional arms races produce inefficient outcomes, people have an incentive to curtail them. Steps taken to reduce positional arms races, such as blue laws and rules against anabolic steroids, may therefore be thought of as **positional arms control agreements**.

Once you become aware of positional arms races, you will begin to see them almost everywhere. You can hone your skills as an economic naturalist by asking these questions about every competitive situation you observe: What form do the investments in performance enhancement take? What steps have contestants taken to limit these investments? Sometimes positional arms control agreements are achieved by the imposition of formal rules or by the signing of legal contracts. Some examples of this type of agreement follow.

Campaign Spending Limits

In the United States, presidential candidates routinely spend hundreds of millions of dollars on advertising. Yet if both candidates double their spending on ads, each one's odds of winning will remain essentially the same. Recognition of this pattern led Congress to adopt strict spending limits for presidential candidates. (That those regulations have proved difficult to enforce does not call into question the logic behind the legislation.)

Roster Limits

Major League Baseball permits franchises to have only 25 players on the roster during the regular season. The National Football League sets its roster limit at 53; the National Basketball Association at 12. Why these limits? In their absence, any team could increase its chance of winning by simply adding players. Inevitably, other teams would follow suit. On the plausible assumption that, beyond some point, larger rosters do not add much to the entertainment value for fans, roster limits are a sensible way to deliver sports entertainment at a more reasonable cost.

Arbitration Agreements

In the business world, contracting parties often sign a binding agreement that commits them to arbitration in the event of a dispute. By doing so, they sacrifice the option of pursuing their interests as fully as they might wish to later, but they also insulate themselves from costly legal battles. Other parties in the legal system may sometimes take steps to limit spending on litigation. For example, a federal judge in South Dakota announced—presumably to the approval of litigants—that he would read only the first 15 pages of any brief submitted to his court.

Mandatory Starting Dates for Kindergarten

A child who is a year or so older than most of her kindergarten classmates is likely to perform better, in relative terms, than if she had entered school with children her own age. And since most parents are aware that admission to prestigious universities

and eligibility for top jobs upon graduation depend largely on *relative* academic performance, many are tempted to keep their children out of kindergarten a year longer than necessary. Yet there is no social advantage in holding *all* children back an extra year since their relative performance would essentially be unaffected. In most jurisdictions, therefore, the law requires children who reach their fifth birthday before December 1 of a given year to start kindergarten the same year.

Social Norms as Positional Arms Control Agreements

In some cases, social norms may take the place of formal agreements to curtail positional arms races. Some familiar examples follow.

Nerd Norms

Some students care more—in the short run, at least—about the grades they get than how much they actually learn. When such students are graded on the curve—that is, on the basis of their performance relative to other students—a positional arms race ensues because if all students were to double the amount of time they studied, the distribution of grades would remain essentially the same. Students who find themselves in this situation are often quick to embrace “nerd norms,” which brand as social misfits those who “study too hard.”

Fashion Norms

Social norms regarding dress and fashion often change quickly because of positional competition. Consider, for instance, the person who wishes to be on the cutting edge of fashion. In some American social circles during the 1950s, that goal could be accomplished by having pierced ears. But as more and more people adopted the practice, it ceased to communicate avant-garde status. At the same time, those who wanted to make a conservative fashion statement gradually became freer to have their ears pierced.

For a period during the 1960s and 1970s, one could be on fashion’s cutting edge by wearing two earrings in one earlobe. But by the 1990s multiple ear piercings had lost much of their social significance, the threshold of cutting-edge status having been raised to upward of a dozen piercings of each ear or a smaller number of piercings of the nose, eyebrows, or other body parts. A similar escalation has taken place in the number, size, and placement of tattoos.

The increase in the required number of tattoos or body piercings has not changed the value of avant-garde fashion status to those who desire it. Being on the outer limits of fashion has much the same meaning now as it once did. To the extent that there are costs associated with body piercings, tattoos, and other steps required to achieve avant-garde status, the current fashions are wasteful compared to earlier ones. In this sense, the erosion of social norms against tattoos and body piercings has produced a social loss. Of course, the costs associated with this loss are small in most cases. Yet since each body piercing entails a small risk of infection, the costs will continue to rise with the number of piercings. And once those costs reach a certain threshold, support may mobilize on behalf of social norms that discourage these activities.

Norms of Taste

Similar cycles occur with respect to behaviors considered to be in bad taste. In the 1950s, for example, prevailing norms prevented major national magazines from accepting ads that featured nude photographs. Naturally, advertisers had a powerful incentive to chip away at such norms in an effort to capture the reader’s limited attention. And indeed, taboos against nude photographs have eroded in the same way as with those against body piercings and tattoos.



druvo/Vetta/Getty Images

Is being on fashion’s cutting edge more valuable now than in the 1950s?

Consider, for instance, the evolution of perfume ads. First came the nude silhouette; then, increasingly well-lit and detailed nude photographs; and more recently, photographs of what appear to be group sex acts. Each innovation achieved just the desired effect: capturing the reader's instant and rapt attention. Inevitably, however, other advertisers followed suit, causing a shift in our sense of what is considered attention-grabbing. Photographs that once would have shocked readers now often draw little more than a bored glance.

Opinions differ, of course, about whether this change is an improvement. Many believe that the earlier, stricter norms were ill-advised, the legacy of a more prudish and repressive era. Yet even people who take that view are likely to believe that *some* kinds of photographic material ought not to be used in magazine advertisements. Obviously, what is acceptable will differ from person to person, and each person's threshold of discomfort will depend in part on current standards. But as advertisers continue to break new ground in their struggle to capture attention, the point may come when people begin to mobilize in favor of stricter standards of "public decency." Such a campaign would provide yet another case of a positional arms control agreement.

Norms against Vanity

Cosmetic and reconstructive surgery has produced dramatic benefits for many people, enabling badly disfigured accident victims to recover a normal appearance. It also has eliminated the extreme self-consciousness felt by people born with strikingly unusual features. Such surgery, however, is by no means confined to the conspicuously disfigured. Increasingly, "normal" people are seeking surgical improvements to their appearance. For example, in the United States, over 18 million cosmetic procedures were performed in 2019, and demand has continued to grow. Once a carefully guarded secret, these procedures are now offered as prizes in Southern California charity raffles.

In individual cases, cosmetic surgery may be just as beneficial as reconstructive surgery is for accident victims. Buoyed by the confidence of having a straight nose or a wrinkle-free complexion, patients sometimes go on to achieve much more than they ever thought possible. But the growing use of cosmetic surgery also has had an unintended side effect: It has altered the standards of normal appearance. A nose that once would have seemed only slightly larger than average may now seem jarringly big. The same person who once would have looked like an average 55-year-old may now look nearly 70. And someone who once would have tolerated slightly thinning hair or an average amount of cellulite may now feel compelled to undergo hair transplantation or liposuction. Because such procedures shift people's frame of reference, their payoffs to individuals are misleadingly large. From a social perspective, therefore, reliance on them is likely to be excessive.

Legal sanctions against cosmetic surgery are difficult to imagine. But some communities have embraced powerful social norms against cosmetic surgery, heaping scorn and ridicule on the consumers of face-lifts and tummy tucks. In individual cases, such norms may seem cruel. Yet without them, many more people might feel compelled to bear the risk and expense of cosmetic surgery.

RECAP ↑

POSITIONAL EXTERNALITIES

Positional externalities occur when an increase in one person's performance reduces the expected reward of another person in situations in which reward depends on relative performance. Positional arms races are a series of mutually offsetting investments in performance enhancement that are stimulated by a positional externality. Positional arms control agreements are sometimes enacted in an attempt to limit positional arms races. In some cases, social norms can act as positional arms control agreements.

SUMMARY

- Externalities are the costs and benefits of activities that accrue to people who are not directly involved in those activities. When all parties affected by externalities can negotiate with one another at no cost, the invisible hand of the market will produce an efficient allocation of resources. (LO1)
- According to the Coase theorem, the allocation of resources is efficient in such cases because the parties affected by externalities can compensate others for taking remedial action. (LO2)
- Negotiation over externalities is often impractical, however. In these cases, the self-serving actions of individuals typically will not lead to an efficient outcome. The attempt to forge solutions to the problems caused by externalities is one of the most important rationales for collective action. Sometimes collective action takes the form of laws and government regulations that alter the incentives facing those who generate, or are affected by, externalities. Such remedies work best when they place the burden of accommodation on the parties who can accomplish it at the lowest cost. Traffic laws, zoning laws, environmental protection laws, and free speech laws are examples. (LO3)
- Curbing pollution and other negative externalities entails costs as well as benefits. The optimal amount of pollution reduction is the amount for which the marginal benefit of further reduction just equals the marginal cost. In general, this formula implies that the socially optimal level of pollution, or of any other negative externality, is greater than zero. (LO3)
- When grazing land and other valuable resources are owned in common, no one has an incentive to take into account the opportunity cost of using those resources. This problem is known as the tragedy of the commons. Defining and enforcing private rights governing the use of valuable resources is often an effective solution to the tragedy of the commons. Not surprisingly, most economically successful nations have well-developed institutions of private property. Property that belongs to everyone belongs, in effect, to no one. Not only is its potential economic value never fully realized; it usually ends up having no value at all. (LO4)
- The difficulty of enforcing property rights in certain situations explains a variety of inefficient outcomes such as the excessive harvest of whales in international waters and the premature harvest of timber on remote public lands. The excessive pollution of seas that are bordered by many countries also results from a lack of enforceable property rights. (LO4)
- Situations in which people's rewards depend on how well they perform in relation to their rivals give rise to positional externalities. In these situations, any step that improves one side's relative position necessarily worsens the other's. Positional externalities tend to spawn positional arms races—escalating patterns of mutually offsetting investments in performance enhancement. Collective measures to curb positional arms races are known as positional arms control agreements. These collective actions may take the form of formal regulations or rules such as rules against anabolic steroids in sports, campaign spending limits, and binding arbitration agreements. Informal social norms can also curtail positional arms races. (LO5)

KEY TERMS

Coase theorem
external benefit
(or positive externality)

external cost (or negative externality)
externality
positional arms control agreement

positional arms race
positional externality
tragedy of the commons

REVIEW QUESTIONS

1. If Congress could declare illegal any activity that imposes external costs on others, would such legislation be advisable? (LO2)
2. What incentive problem explains why the freeways in cities like Los Angeles suffer from excessive congestion? (LO3)
3. How would you explain to a friend why the optimal amount of freeway congestion is not zero? (LO3)
4. Why might it be easier to reduce pollution levels in the Great Salt Lake, which is located solely in the state of Utah, than in Lake Erie, which is bordered by several states and Canada? (LO4)
5. Explain why the wearing of high-heeled shoes might be viewed as the result of a positional externality. (LO5)

PROBLEMS



- For each of the goods listed below, discuss whether the good is likely to entail an external cost or an external benefit. In addition, discuss whether the market is likely to provide more or less than the socially optimal quantity of the good. (LO1)
 - Vaccinations.
 - Cigarettes.
 - Antibiotics.
- Suppose the law says that Jones may *not* emit smoke from his factory unless he gets permission from Smith, who lives downwind. The relevant costs and benefits of filtering the smoke from Jones's production process are as shown in the following table. If Jones and Smith can negotiate with one another at no cost, will Jones emit smoke? (LO2)

	Jones emits smoke	Jones does not emit smoke
Surplus for Jones	\$200	\$160
Surplus for Smith	\$400	\$420

- John and Karl can live together in a two-bedroom apartment for \$900 per month, or each can rent a single-bedroom apartment for \$750 per month. Aside from the rent, the two would be indifferent between living together and living separately, except for one problem: John leaves dirty dishes in the sink every night. Karl would be willing to pay up to \$175 per month to avoid John's dirty dishes. John, for his part, would be willing to pay up to \$225 to be able to continue his sloppiness. (LO2)
 - Should John and Karl live together? If they do, will there be dirty dishes in the sink? Explain.
 - What if John would be willing to pay up to \$30 per month to avoid giving up his privacy by sharing quarters with Karl? Should John and Karl live together?
- Khatia and Gwendolyn are neighbors in an apartment complex. Khatia is a concert pianist, and Gwendolyn is a poet. Khatia rehearses her concert pieces on the baby grand piano in her front room, which is directly above Gwendolyn's study. The following matrix shows the monthly payoffs to Khatia and Gwendolyn when Khatia's front room is and is not soundproofed. The soundproofing will be effective only if it is installed in Khatia's apartment. (LO2)

	Soundproofed	Not soundproofed
Gains to Khatia	\$100/month	\$150/month
Gains to Gwendolyn	\$120/month	\$ 80/month

- If Khatia has the legal right to make any amount of noise she wants and she and Gwendolyn can negotiate with one another at no cost, will Khatia install and maintain soundproofing? Explain. Is her choice socially efficient?
 - If Gwendolyn has the legal right to peace and quiet and can negotiate with Khatia at no cost, will Khatia install and maintain soundproofing? Explain. Is her choice socially efficient?
 - Does the attainment of an efficient outcome depend on whether Khatia has the legal right to make noise, or Gwendolyn the legal right to peace and quiet?
- Refer to Problem 4. Khatia decides to buy a full-size grand piano. The new payoff matrix is as follows. (LO2)

	Soundproofed	Not soundproofed
Gains to Khatia	\$100/month	\$150/month
Gains to Gwendolyn	\$120/month	\$ 60/month

- If Gwendolyn has the legal right to peace and quiet and Khatia and Gwendolyn can negotiate at no cost, will Khatia install and maintain soundproofing? Explain. Is this outcome socially efficient?
 - Suppose that Khatia has the legal right to make as much noise as she likes and that negotiating an agreement with Khatia costs \$15 per month. Will Khatia install and maintain soundproofing? Explain. Is this outcome socially efficient?
 - Suppose Gwendolyn has the legal right to peace and quiet, and it costs \$15 per month for Gwendolyn and Khatia to negotiate any agreement. (Compensation for noise damage can be paid without incurring negotiation cost.) Will Khatia install and maintain soundproofing? Is this outcome socially efficient?
 - Why does the attainment of a socially efficient outcome now depend on whether Khatia has the legal right to make noise?
- Which of the following statements is correct? (LO3)
 - Taxes and subsidies are the only tools that governments use to regulate activities that generate externalities.
 - The government typically does not tax activities that generate negative externalities since taxes lower economic efficiency.
 - Taxes on activities that generate negative externalities increase economic efficiency.
 - Taxes on activities that generate positive externalities have no impact on economic efficiency.

- 7* A village has six residents, each of whom has accumulated savings of \$100. Each villager can use this money either to buy a government bond that pays 15 percent interest per year or to buy a year-old llama, send it onto the commons to graze, and sell it after 1 year. The price the villager gets for the 2-year-old llama depends on the quality of the fleece it grows while grazing on the commons. That in turn depends on the animal's access to grazing, which depends on the number of llamas sent to the commons, as shown in the following table.

Number of llamas on the commons	Price per 2-year-old llama (\$)
1	122
2	118
3	116
4	114
5	112
6	109

The villagers make their investment decisions one after another, and their decisions are public. (LO4)

- If each villager decides individually how to invest, how many llamas will be sent onto the commons, and what will be the resulting village income?
- What is the socially optimal number of llamas for this village? Why is that different from the actual

number? What would village income be if the socially optimal number of llamas were sent onto the commons?

- The village committee votes to auction the right to graze llamas on the commons to the highest bidder. Assuming villagers can both borrow and lend at 15 percent annual interest, how much will the right sell for at auction? How will the new owner use the right, and what will be the resulting village income?
8. Suppose that Lance and Jan are the two top cyclists in the world. Both are scheduled to compete in an upcoming cycling competition in which the winner will receive \$100,000 in prize money while the rest of the competitors receive nothing. Both cyclists are very talented, but they can increase their chances of winning by doping (i.e., taking performance-enhancing drugs). The cost to each of doping is \$25,000 (this includes not just the cost of the drugs, but also both the health costs and the expected damage to their reputations if they are caught). Each has a 50 percent chance of winning the race either if both of them dope or if neither of them dopes. On the other hand, if only one of them dopes, then the one who dopes will win the race for sure. (LO5)
- Discuss whether this situation involves a positional externality.
 - Write down the expected payoff matrix for this game, assuming that both Lance and Jan make their decisions simultaneously.
 - Will the outcome of this game be socially optimal? Explain.

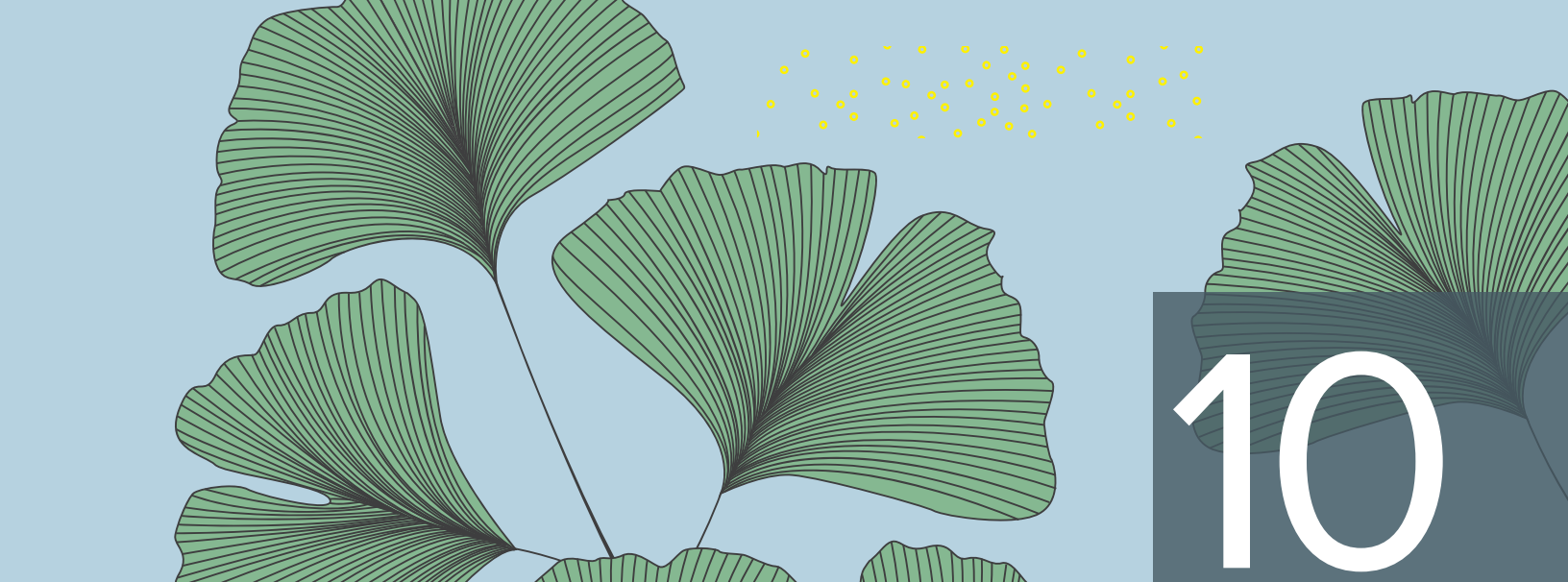
*Denotes more difficult problem.

ANSWERS TO SELF-TESTS

- 9.1 Since Fitch gains \$50 per day when Abercrombie operates with a filter, he could pay Abercrombie as much as \$49 per day and still come out ahead. (LO2)
- 9.2 If the two were to live together, the most efficient way to resolve the problem would be as before, for Sofia to tolerate sometimes being woken up by Adriana's friends. But on top of that cost, which is \$150, Sofia would also bear a \$60 cost from the loss of her privacy. The total cost of their living together would thus be \$210 per month. Since that amount is greater than the \$200 saving in rent, the two should live separately. (LO2)
- 9.3 The income figures from the different levels of investment in cattle would remain as before, as shown in the table. What is different is the opportunity cost of investing in each steer, which is now \$11 per year instead of \$13. Column 5 of the table shows that the socially optimal number of steers is

now 2 instead of 1. And if individuals still favor holding cattle, all other things being equal, they will now send 5 steers onto the commons instead of 4, as shown in column 3. (LO4)

(1) Number of steers on the commons	(2) Price per 2-year-old steer (\$)	(3) Income per steer (\$/year)	(4) Total village income (\$/year)	(5) Marginal income (\$/year)
1	126	26	26	26
2	119	19	38	12
3	116	16	48	10
4	113	13	52	4
5	111	11	55	3



10

LEARNING OBJECTIVES

After reading this chapter, you should be able to:

- LO1** Describe how the Cost-Benefit Principle applies even to choices involving health.
- LO2** Explain why health care costs have been rising so rapidly.
- LO3** Compare and contrast the ways in which taxes and tradable permits can be used to reduce pollution.
- LO4** Assess the economic pros and cons of the various methods used to reduce poverty in the United States.

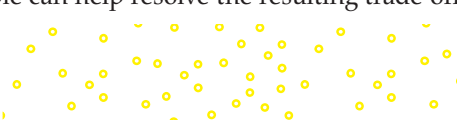
Using Economics to Make Better Policy Choices

In 1979, in the wake of the second major oil supply interruption in a decade, officials in the Carter administration met to discuss policies for reducing the risks to domestic security inherent in U.S. dependence on foreign oil. The proposal they ultimately put forward was a gasoline tax of 50 cents per gallon. Anticipating objections that the tax would impose an unacceptable hardship on the poor, policymakers proposed to return the revenues from the tax to the citizenry by reducing the payroll tax—the tax on wages that supports the Social Security system.

Proponents of the gasoline tax argued that in addition to reducing the nation's dependence on foreign oil, the tax would reduce air pollution and ease highway congestion. But critics ridiculed the proposal, charging that if the revenues from the tax were returned to the people, the quantity of gasoline demanded would remain essentially the same. Their argument tipped the debate, and officials never managed to implement the proposal.

Whatever the ultimate merits of the administration's proposal, there was no merit at all in the argument the critics used to attack it. True, the proposed tax rebate meant that people *could* have bought just as much gasoline as before the tax. Yet the tax would have given them a powerful incentive not to do so. Consumers can change their behavior to escape the effects of a steep rise in the after-tax price of gasoline—by switching to cars with smaller, more fuel-efficient engines; forming carpools; and so on. Such changes free up money to spend on other goods and services, which become relatively more attractive because they are not taxed.

No society can hope to formulate and implement intelligent economic policies unless its citizens and leaders share an understanding of basic economic principles. Our aim in this chapter is to explore how careful application of these principles can help us design policies that both expand the economic pie and make everyone's slice larger. We will explore the economics of health care delivery, environmental regulation, and programs for providing assistance to low-income persons. The unifying thread running through these issues is the problem of scarcity. In each case, we will explore how the Cost-Benefit Principle can help resolve the resulting trade-offs.



THE ECONOMICS OF HEALTH CARE

Political leaders are often reluctant to discuss expenditures on public health programs in cost-benefit terms. But because we live in a world of scarcity, we cannot escape the fact that spending more on health means spending less on other things of value.

Illnesses, like accidents, are costly to prevent. The socially optimal expenditure on a health measure that reduces a specific illness is that amount for which the marginal benefit to society of the measure exactly equals its marginal cost. For example, in deciding how much to spend on vaccinating against measles, a rational public health policy would expand the proportion of the population vaccinated until the marginal cost of an additional vaccination was exactly equal to the marginal value of the illnesses thus prevented.

As we explain in the following section, however, the decision of whether to become vaccinated looks very different from each individual's perspective.

The Case for Mandatory Immunization Laws

Being vaccinated against a childhood illness entails a small but potentially serious risk. The vaccine against pertussis (whooping cough), for example, is believed to cause some form of permanent brain damage in 1 out of every 110,000 children vaccinated. Contracting the disease itself also poses serious health risks, and in an environment in which infections were sufficiently likely to occur, individuals would have a compelling reason to bear the risk of being vaccinated in order to reduce the even larger risk from infection. The problem is that in an environment in which most children were vaccinated, infection rates would be low, making the risk of vaccination loom relatively large in the eyes of individual families.

The ideal situation from the perspective of any individual family would be to remain unvaccinated in an environment in which all other families were vaccinated. But as more and more families decided to forgo vaccination, infection rates would mount. Eventually the vaccination rate would stabilize at the point at which the additional risk to the individual family of becoming vaccinated would be exactly equal to the risk from remaining unvaccinated. But this calculation ignores the fact that a decision to remain unvaccinated poses risk not just to the individual decision maker, but also to others who have decided to become vaccinated (since no vaccine affords 100 percent protection against infection).

Relegating the vaccination decision to individuals thus results in a suboptimally low vaccination rate because individual decision makers fail to take adequate account of the cost that their becoming infected will impose on others. It is for this reason that most states require vaccinations against specific childhood illnesses. Proof of immunization against diphtheria, measles, poliomyelitis, and rubella, for example, is now universally required for entry into American public schools. Most states also require immunization against tetanus (49 states), pertussis (44 states), mumps (43 states), and hepatitis B (26 states).

Even these laws, however, allow parents to apply for exemptions on religious or philosophical grounds. Communities vary in the extent to which parents avail themselves of these exemptions. In Colorado, for instance, Boulder County heads the list of parents who opt to exempt their children from taking the pertussis vaccine (with an exemption rate of 8.4 percent, more than four times the rate statewide). Not surprisingly, the incidence of whooping cough is much higher in Boulder (34.7 cases per year per 100,000 people) than in the state as a whole (22.4 cases per year per 100,000 people).¹

¹From the Colorado School and Childcare Immunization Data 2016–2017, www.cohealthdata.dphe.state.co.us/Data/Details/1; and the Colorado Department of Public Health and Environment, www.colorado.gov/pacific/cdphe/colorado-health-indicators.

EXPLAINING RISING HEALTH CARE COSTS

In the United States, real health care expenditures per capita have grown more rapidly than real income per capita for as long as the relevant data have been available. As a share of national income, health care costs have risen from only 4 percent in 1940 to roughly 17 percent today. Part of this increase is the result of costly new health care technologies and procedures. Diagnostic tests have grown more expensive and sophisticated, and procedures like coronary bypass surgery and organ transplantation have grown far more common. Yet a great deal of medical expenditure inflation has nothing to do with these high-tech developments. Rather, it is the result of fundamental changes in the way we pay for medical services.

The most important change has been the emergence of the so-called third-party payment system. Earlier in the twentieth century, many people insured themselves against catastrophic illness but purchased routine medical care out of their own pockets, just as they did food, clothing, and other consumer goods. Starting after World War II, and increasingly since the mid-1960s, people have come to depend on insurance for even routine medical services. Some insurance is provided by employers, some by the government. Medicaid covers the medical expenses of the poor and Medicare, those of people who are older or disabled.

The spread of medical insurance, especially government-financed medical insurance, owes much to the belief that an inability to pay should not prevent people from receiving medical care they need. Indeed, medical insurance has surely done much to shelter people from financial hardship. The difficulty is that in its most common form, it also has spawned literally hundreds of billions of dollars of waste each year.

To understand the nature of this waste, we must recognize that although medical services differ from other services in many ways, they are in one fundamental respect the same: The cost-benefit test is the only sensible criterion for deciding which services ought to be performed. The fact that a medical procedure has *some* benefit does not, by itself, imply that the procedure should be performed. Rather, it should be performed only if its marginal benefit, broadly construed, exceeds its marginal cost.

The costs of medical procedures are relatively easy to measure, using the same methods applied to other goods and services. But the usual measure of the benefit of a good or service, a person's willingness to pay, may not be acceptable in the case of many medical services. For example, most of us would not conclude that a lifesaving appendectomy that costs \$2,000 is unjustified merely because one person who needs it can afford to pay only \$1,000. When someone lacks the resources to pay for what most of us would consider an essential medical service, society has at least some responsibility to help. Hence the proliferation of government-sponsored medical insurance.

Many other medical expenditures are not as pressing as an emergency appendectomy, however. Following such surgery, for example, the patient requires a period of recuperation in the hospital. How long should that period last—2 days? 5? 10? The Cost-Benefit Principle is critically important to thinking intelligently about such questions. But as Example 10.1 illustrates, the third-party payment system has virtually eliminated cost-benefit thinking from the medical domain.

EXAMPLE 10.1 The Impact of a Third-Party Payment System on Cost-Benefit Thinking

How long should David stay in the hospital?

To eliminate recurrent sore throats, David plans to have his tonsils removed. His surgeon tells him that the average hospital stay after this procedure is two days (some people stay only one day, while others stay three, four, or even five days). Hospital rooms cost \$300 per day. If David's demand curve for days in the hospital

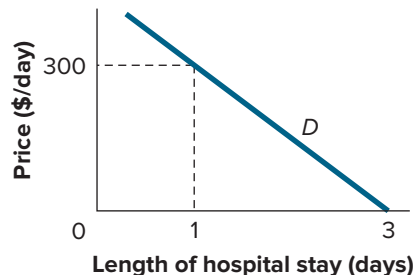
is as shown in Figure 10.1, how many days will he stay if he must pay for his hospital room himself? How many days will he stay if his medical insurance fully covers the cost of his hospital room?

If David must pay for his hospital room himself, his best option will be to stay for just one day. But if the cost of his hospital room is completely covered by insurance, the marginal cost to *him* will be zero. In that case, he will stay for three days.

FIGURE 10.1

The Demand for Hospital Care.

The demand curve for postoperative hospital care is downward-sloping, just like any other demand curve. At higher prices, people choose shorter hospital stays, not because there is no benefit to a longer stay, but because they prefer to spend their money in other ways.



SELF-TEST 10.1

In Example 10.1, how long would David choose to stay in the hospital if his health insurance covered 50 percent of the cost of his hospital room?

Should we be concerned that people choose longer hospital stays when their expenses are fully insured? The Cost-Benefit Principle tells us that a hospital stay should be extended another day only if the benefit of doing so would be at least as great as the cost of the resources required to extend the stay. But when hospital costs are fully covered by insurance, the decision maker sees a marginal cost of zero, when in fact the marginal cost is several hundred dollars. According to the Cost-Benefit Principle, then, full insurance coverage leads to wastefully long hospital stays. That is not to say that the additional days in the hospital do no good at all. Rather, their benefit, as measured by the decision maker's willingness to pay for them, is less than their cost (\$300 per day).

Designing a Solution

The amount of waste caused by full insurance coverage depends on the price elasticity of demand for medical services—the more elastic the demand, the greater the waste. Proponents of full coverage believe that the demand for medical services is almost completely inelastic with respect to price and that the resulting waste is therefore negligible. Critics of full coverage argue that the demand for medical services is actually quite sensitive to price and that the resulting waste is significant.

Who is right? One way to determine this is to examine whether people who lack full insurance coverage spend significantly less than those who have it. The economist W. G. Manning and several co-authors did so by performing an experiment in which they assigned subjects randomly to one of two different kinds of medical insurance policy.² The first group of subjects received **first-dollar coverage**, meaning that 100 percent of their

first-dollar insurance coverage insurance that pays all expenses generated by the insured activity

²W. G. Manning, J. P. Newhouse, E. B. Keeler, A. Liebowitz, and M. S. Marquis, "Health Insurance and the Demand for Medical Care," *American Economic Review* 77 (June 1987), pp. 251–277.

medical expenses was covered by insurance. The second group got “\$1,000-deductible” coverage, meaning that only expenses beyond the first \$1,000 a year were covered. (For example, someone with \$1,200 of medical bills would receive \$1,200 from his insurance company if he belonged to the first group, but only \$200 if he belonged to the second.) In effect, since most people incur less than \$1,000 a year in medical expenses, most subjects in the second group effectively paid full price for their medical services, while subjects in the first group paid nothing. Manning and his colleagues found that *people with \$1,000-deductible policies spent between 40 and 50 percent less on health care than subjects with first-dollar coverage. More important, there were no measurable differences in health outcomes between the two groups.*

Taken at face value, the results of the Manning study suggest that a large share of the inflation in medical expenditures since World War II has been caused by growth in first-dollar medical insurance. The problem with first-dollar coverage is that it completely ignores people’s incentives. Why not simply abandon first-dollar coverage in favor of high deductibles? People would still be protected against financial catastrophe but would have a strong incentive to avoid medical services whose benefit does not exceed their cost.

Some would say that Medicaid and Medicare should not carry high deductibles because the resulting out-of-pocket payments would impose too great a burden on poor families. But as in other instances in which concern for the poor is offered in defense of an inefficient policy, an alternative can be designed that is better for rich and poor alike. For example, all health insurance could be written to include high deductibles, and the poor could be given an annual stipend to defray the initial medical expenses not covered by insurance. At year’s end, any unspent stipend would be theirs to keep. Here again, concern for the well-being of the poor is no reason for not adopting the most efficient policy. Again, when the economic pie grows larger, it is possible for everyone to have a larger slice.

The HMO Revolution

During the 1990s, the high cost of conventional health insurance led many people to switch to **health maintenance organizations (HMOs)**. An HMO is a group of physicians that provides its patients with medical services in return for a fixed annual fee. As The Economic Naturalist 10.1 illustrates, the incentive to provide any given medical service is weaker under the standard HMO contract than under conventional health insurance.

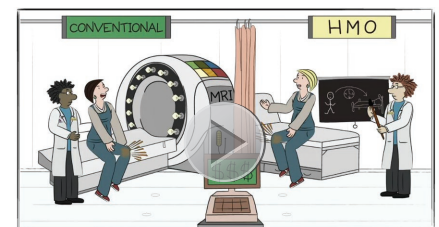
health maintenance organization (HMO) a group of physicians that provides health services to individuals and families for a fixed annual fee

The Economic Naturalist 10.1



Why is a patient with a sore knee more likely to receive an MRI exam if he has conventional health insurance than if he belongs to a health maintenance organization?

When a patient visits his physician complaining of a sore knee, the physician has several options. After hearing the patient describe his symptoms and examining the knee manually, the physician may prescribe anti-inflammatory drugs and advise the patient to abstain from vigorous physical activity for a period; or she may advise the patient to undergo a magnetic resonance imaging (MRI) exam, a costly diagnostic procedure that generates images of the inner workings of the injured joint. The physician in an HMO receives no additional revenue if she orders the MRI because all services are covered by the patient’s fixed annual fee. Under conventional health insurance, in contrast, the physician will be reimbursed at a fixed rate, usually well above her marginal cost, for each additional service performed.



Why might a patient be more likely to receive an expensive magnetic resonance imaging (MRI) exam for a sore knee if covered under a conventional health insurance plan rather than a Health Maintenance Organization (HMO) plan?

In many instances, the most prudent course of treatment is unambiguous, and in such cases physicians will make the same recommendation despite this striking difference in incentives. But in many other cases, it may not be obvious which decision is best. And in these cases, HMO physicians are less likely to order expensive tests.

People who switch to HMOs pay less for their health plans than those who stick with conventional health insurance since the HMO contract provides a strong incentive for doctors not to prescribe nonessential services. Many people fear, however, that the very same incentive may sometimes result in their not receiving valuable care. These concerns have led to proposed legislation granting patients rights of appeal when they are denied care by an HMO.

The Problem with Health Care Provision through Private Insurance

It is troubling, but perhaps not surprising, that access to medical care is extremely limited in many of the world's poorest nations. After all, citizens of those nations lack enough income to buy adequate food, shelter, and many other basic goods and services. What is surprising, however, is that despite the movement to less-expensive HMO plans, almost 50 million Americans still had no health coverage of any kind when President Barack Obama took office in 2009. Politicians in both parties agreed that something had to be done. But why were so many without health coverage in the first place?

The answer to that question is rooted in the fact that the United States was almost alone among the world's nations in its reliance on unregulated private insurance markets to orchestrate the delivery of health care to its citizens. This approach was essentially a historical accident, a consequence of the fact that many labor unions managed to negotiate employer-provided health insurance as part of their compensation packages during the rapidly growing economy of the immediate post-World War II years. Under government policy, employer expenditures for health insurance were nontaxable. Employer-provided insurance was thus much cheaper for employees than private insurance purchased individually with income on which they had already been taxed. That incentive induced nonunion employers to join their union counterparts in offering employer-provided health insurance. And as long as health care spending was a fairly small share of total income, coverage was broad and the system functioned reasonably well.

An important policy detail was that eligibility for the tax exemption was conditional on insurance being made available to all employees irrespective of preexisting medical conditions. Given the high cost of treating individuals with chronic medical problems, private insurance companies are generally reluctant to issue policies to people with serious health problems. But by covering large groups of employees, only a small percentage of whom would be likely to have serious health problems during any year, insurance companies could issue these policies without taking unacceptable risks. Indeed, the large new employer-provided insurance market was sufficiently lucrative that most insurance companies were eager to participate in it.

Yet unregulated private insurance markets are a deeply flawed mechanism for providing access to health care, and no other industrial nation relies on unregulated markets for this purpose. The tax exemption, coupled with the requirement of group coverage, has enabled participants on both sides of the insurance market to sidestep the problem that undermines private insurance markets.

Economists call it the **adverse-selection** problem. It arises in individual insurance markets because individuals generally know much more about their own health status than the companies that sell insurance. To remain in business, a private

adverse selection the pattern in which insurance tends to be purchased disproportionately by those who are most costly for companies to insure

insurance company must collect enough in premiums to cover the cost of the medical treatments it covers. If its rates are based on the expected medical expenses of a person with average health status, its policy will seem like a bargain to potential customers who know themselves to be in bad health. At the same time, its policies will seem overpriced to those who know themselves to be in excellent health.

The upshot is that a disproportionate share of the customers it attracts will have below-average health status, which means its initial premiums will be too low to cover its costs. To stay in business, it will have to raise its rates. But then potential customers in good health will find its policies even less attractive. A downward spiral often ensues, with the end result that insurance becomes unaffordable for most people.

Although the employer-provided group insurance approach helped keep the adverse-selection problem at bay for many years, this approach began to unravel as medical costs continued to rise relative to all other goods and services. With health insurance premiums taking a bigger and bigger bite out of workers' paychecks and heightened competition forcing companies to look for new ways to cut costs, some began offering higher wages in lieu of employer-provided health coverage. Younger, healthier workers—for whom medical expenses are normally small—found these offers increasingly tempting.

Parents who didn't buy health insurance for their families were once viewed as irresponsible, but this stigma lost some of its sting as the number of uninsured grew. As more and more people took jobs without health coverage, going without insurance became more socially acceptable. Making matters worse was the changing composition of the pool of the insured. As more healthy families took jobs without coverage, those left tended to be sicker and more costly to treat, forcing premiums to rise still more rapidly. In short, our health insurance system was caught in a long-term death spiral.

The Affordable Care Act of 2010

Passed by Congress and signed into law by President Obama in March 2010, the Affordable Care Act was the government's first serious attempt to halt that death spiral. It contained three main provisions, each one of which was essential for reform to succeed. First, it required insurance companies to offer coverage to everyone on roughly equal terms, irrespective of preexisting medical conditions. Without this provision, the economic imperative of every private insurance company would have been to take every step possible to deny coverage to anyone expected to incur significant medical expenses. Any insurance system that couldn't cover those who most need care would clearly be unacceptable.

Because insurance companies cannot cover their costs if they insure only the least-healthy people, it was also necessary for the Affordable Care Act to include a mandate requiring everyone to buy health insurance. Without such a mandate, healthy individuals would face strong incentives to go without health insurance until they got sick, since they would then be able to buy affordable insurance from companies that were forbidden to charge high rates based on preexisting conditions.

The third major feature of the Affordable Care Act was to provide for subsidies to low-income families. You can't require people to buy insurance if they can't afford it. With health care costs already high and rapidly rising, it was essential to include some provision to ease the burden on those who are unable to pay.

The act contained numerous other provisions, many of which were designed to reduce the growth rate of health care costs by requiring more streamlined medical record keeping and supporting research on the questions of which treatments were most effective. But the essence of the act lies in its three main provisions—nondiscrimination on the basis of preexisting conditions, the mandate, and subsidies for low-income families. Without all three of these provisions, the health insurance industry's death spiral would have surely continued.

In tax legislation enacted in late 2017, Congress included a measure repealing the Affordable Care Act's mandate. This step promises to reduce participation rates of healthy young people, with adverse effects on the costs of serving those who remain insured. Without further legislative action to maintain the diversity of the insured pool, the health care industry's long-term death spiral may be poised to resume.

RECAP ↑

THE ECONOMICS OF HEALTH CARE DELIVERY

The rapid escalation in medical expenditures since World War II is attributable in large part to the spread of first-dollar insurance coverage, which encourages people to behave as if medical services were free of charge. Total economic surplus would be larger if we switched to insurance coverage with high deductibles because such policies provide an incentive to use only those services whose benefit exceeds their cost.

The switch to HMOs addresses this problem because the standard HMO contract provides a strong incentive for physicians not to prescribe non-essential services. Some voice concern, however, that HMO contracts may lead physicians to withhold services that satisfy the cost-benefit test.

Mounting insurance premiums have caused many people in good health to do without health coverage, resulting in higher premiums for those who remain insured. The Affordable Care Act of 2010 was enacted in an attempt to remedy market failures that exist in attempts to provide health care access through unregulated private insurance contracts. In the wake of congressional repeal of the ACA's mandate provision in late 2017, the historical decline in the proportion of Americans with health insurance appears poised to resume.

USING PRICE INCENTIVES IN ENVIRONMENTAL REGULATION

As discussed earlier, goods whose production generates negative externalities, such as atmospheric pollution, tend to be overproduced whenever negotiation among private parties is costly. Suppose we decide, as a society, that the best attainable outcome would be to have half as much pollution as would occur under completely unregulated conditions. In that case, how should the cleanup effort be distributed among those firms that currently discharge pollution into the environment?

The most efficient—and hence best—distribution of effort is the one for which each polluter's marginal cost of abatement is exactly the same. To see why, imagine that under current arrangements, the cost to one firm of removing a ton of pollution from the air is larger than the cost to another firm. Society could then achieve the same total reduction in pollution at lower cost by having the first firm discharge one ton more into the air and the second firm one ton less.

Unfortunately, government regulators seldom have detailed information on how the cost of reducing pollution varies from one firm to another. Many pollution laws therefore require all polluters simply to cut back their emissions by the same proportion or to meet the same absolute emissions standards. If different polluters have different marginal costs of pollution abatement, however, these approaches will not be efficient.

Taxing Pollution

Fortunately, alternative policies can distribute the cleanup more efficiently, even if the government lacks detailed information about how much it costs different firms to curtail pollution. One method is to tax pollution and allow firms to decide for themselves how much pollution to emit. The following example illustrates the logic of this approach.

EXAMPLE 10.2 **Taxing Pollution****What is the least costly way to cut pollution by half?**

Two firms, Sludge Oil and Northwest Lumber, have access to five production processes, each of which has a different cost and produces a different amount of pollution. The daily costs of the processes and the number of tons of smoke emitted are as shown in Table 10.1. Pollution is currently unregulated, and negotiation between the firms and those who are harmed by pollution is impossible, which means that each firm uses process *A*, the least costly of the five. Each firm emits 4 tons of pollution per day, for a total of 8 tons of pollution per day.

TABLE 10.1
Costs and Emissions for Different Production Processes

Process (smoke)	<i>A</i> (4 tons/day)	<i>B</i> (3 tons/day)	<i>C</i> (2 tons/day)	<i>D</i> (1 ton/day)	<i>E</i> (0 tons/day)
Cost to Sludge Oil (\$/day)	100	200	600	1,300	2,300
Cost to Northwest Lumber (\$/day)	300	320	380	480	700

The government is considering two options for reducing total emissions by half. One is to require each firm to curtail its emissions by half. The other is to set a tax of \$*T* per ton of smoke emitted each day. How large must *T* be to curtail emissions by half? What would be the total cost to society under each alternative?

If each firm is required to cut pollution by half, each must switch from process *A* to process *C*. The result will be 2 tons per day of pollution for each firm. The cost of the switch for Sludge Oil will be \$600 per day – \$100 per day = \$500 per day. The cost to Northwest Lumber will be \$380 per day – \$300 per day = \$80 per day, for a total cost of \$580 per day.

Consider now how each firm would react to a tax of \$*T* per ton of pollution. If a firm can cut pollution by 1 ton per day, it will save \$*T* per day in tax payments. Whenever the cost of cutting a ton of pollution is less than \$*T*, then, each firm has an incentive to switch to a cleaner process. For example, if the tax were set at \$40 per ton, Sludge Oil would stick with process *A* because switching to process *B* would cost \$100 per day extra but would save only \$40 per day in taxes. Northwest Lumber, however, would switch to process *B* because the \$40 saving in taxes would be more than enough to cover the \$20 cost of switching.

The problem is that a \$40-per-day tax on each ton of pollution results in a reduction of only 1 ton per day, 3 short of the 4-ton target. Suppose instead that the government imposed a tax of \$101 per ton. Sludge Oil would then adopt process *B* because the \$100 extra daily cost of doing so would be less than the \$101 saved in taxes. Northwest Lumber would adopt process *D* because, for every process up to and including *C*, the cost of switching to the next process would be less than the resulting tax saving.

Overall, then, a tax of \$101 per ton would result in the desired pollution reduction of 4 tons per day. The total cost of the reduction would be only \$280 per day (\$100 per day for Sludge Oil and \$180 per day for Northwest Lumber), or \$300 per day less than when each firm was required to cut its pollution by half. (The taxes paid by the firms do not constitute a cost of pollution reduction because the money can be used to reduce whatever taxes would otherwise need to be levied on citizens.)

SELF-TEST 10.2

In Example 10.2, if the tax were \$61 per ton of pollution emitted each day, which production processes would the two firms adopt?

The advantage of the tax approach is that it concentrates pollution reduction in the hands of the firms that can accomplish it at least cost. Requiring each firm to cut emissions by the same proportion ignores the fact that some firms can reduce pollution much more cheaply than others. Note that under the tax approach, the cost of the last ton of smoke removed is the same for each firm, so the efficiency condition is satisfied.

One problem with the tax approach is that unless the government has detailed knowledge about each firm’s cost of reducing pollution, it cannot know how high to set the pollution tax. A tax that is too low will result in too much pollution, while a tax that is too high will result in too little. Of course, the government could start by setting a low tax rate and gradually increase the rate until pollution is reduced to the target level. But because firms often incur substantial sunk costs when they switch from one process to another, that approach might be even more wasteful than requiring all firms to cut their emissions by the same proportion.

Auctioning Pollution Permits

Another alternative is to establish a target level for pollution and then auction off permits to emit that level. The virtues of this approach are illustrated in the following example.

EXAMPLE 10.3 Pollution Permits

How much will pollution permits sell for?

Two firms, Sludge Oil and Northwest Lumber, again have access to the production processes described earlier (which are reproduced in Table 10.2). The government’s goal is to cut the current level of pollution, 8 tons per day, by half. To do so, the government auctions off four permits, each of which entitles the bearer to emit 1 ton of smoke per day. No smoke may be emitted without a permit. What price will the pollution permits fetch at auction, how many permits will each firm buy, and what will be the total cost of the resulting pollution reduction?

TABLE 10.2
Costs and Emissions for Different Production Processes

Process (smoke)	A (4 tons/day)	B (3 tons/day)	C (2 tons/day)	D (1 ton/day)	E (0 tons/day)
Cost to Sludge Oil (\$/day)	100	200	600	1,300	2,300
Cost to Northwest Lumber (\$/day)	300	320	380	480	700

If Sludge Oil has no permits, it must use process E, which costs \$2,300 per day to operate. If it had one permit, it could use process D, which would save it \$1,000 per day. Thus, the most Sludge Oil would be willing to pay for a single 1-ton

pollution permit is \$1,000 per day. With a second permit, Sludge Oil could switch to process C and save another \$700 per day; with a third permit, it could switch to process B and save another \$400; and with a fourth permit, it could switch to process A and save another \$100. Using similar reasoning, we can see that Northwest Lumber would pay up to \$220 for one permit, up to \$100 for a second, up to \$60 for a third, and up to \$20 for a fourth.

Suppose the government starts the auction at a price of \$90. Sludge Oil will then demand four permits and Northwest Lumber will demand two, for a total demand of six permits. Since the government wishes to sell only four permits, it will keep raising the price until the two firms together demand a total of only four permits. Once the price reaches \$101, Sludge Oil will demand three permits and Northwest Lumber will demand only one, for a total quantity demanded of four permits. Compared to the unregulated alternative, in which each firm used process A, the daily cost of the auction solution is \$280: Sludge Oil spends \$100 switching from process A to process B, and Northwest Lumber spends \$180 switching from A to D. This total is \$300 less than the cost of requiring each firm to reduce its emissions by half. (Again, the permit fees paid by the firms do not constitute a cost of cleanup because the money can be used to reduce taxes that would otherwise have to be collected.)

The auction method has the same virtue as the tax method: It concentrates pollution reduction in the hands of those firms that can accomplish it at the lowest cost. But the auction method has other attractive features that the tax approach does not. First, it does not induce firms to commit themselves to costly investments that they will have to abandon if the cleanup falls short of the target level. And second, it allows private citizens a direct voice in determining where the emission level will be set. For example, any group that believes the pollution target is too lenient could raise money to buy permits at auction. By keeping those permits locked away in a safe, the group could ensure that they will not be used to emit pollution.

Several decades ago, when economists first proposed the auctioning of pollution permits, reactions of outrage were widely reported in the press. Most of those reactions amounted to the charge that the proposal would “permit rich firms to pollute to their hearts’ content.” Such an assertion betrays a total misunderstanding of the forces that generate pollution. Firms pollute not because they *want* to pollute but because dirty production processes are cheaper than clean ones. Society’s only real interest is in keeping the total amount of pollution from becoming excessive, not in *who* actually does the polluting. And in any event, the firms that do most of the polluting under an auction system will not be rich firms, but those for whom pollution reduction is most costly.

Economists have argued patiently against these misinformed objections to the auction system, and their efforts have finally borne fruit. The sale of pollution permits is now common in several parts of the United States, and there is growing interest in the approach in other countries.

The problem of acid rain is an instructive case in point. Electric power plants in the Midwest had long burned high-sulfur coal, in the process discharging large quantities of sulfur dioxide (SO₂) from their smokestacks. Prevailing winds carried these emissions eastward, where they precipitated out over New York State and New England as acid rain, killing large quantities of trees and fish.

Beginning in the 1960s, economists advocated addressing this problem by requiring firms to acquire marketable permits for each ton of SO₂ they emitted. But as discussed earlier, economists’ proposals, despite their compelling logic, languished in Congress for more than three decades. Only with the passage of amendments to the Clean Air Act in 1995 did Congress require the establishment of a market for permits for discharging SO₂.

As economists had predicted, the permit system worked spectacularly well. Acid rain caused by SO₂ emissions quickly plummeted, and at about one-sixth the cost that more complex regulatory approaches would have entailed.³ Once people have to pay for their emissions, they quickly find ingenious ways of reducing them.

Climate Change and Carbon Taxes

Growing concentrations of carbon dioxide (CO₂) in the atmosphere are widely recognized as a principal contributor to global warming. Concerns about the consequences of climate change have led to proposals to tax CO₂ emissions or require marketable permits for them. Critics of these proposals emphasize that forecasts involving climate change are highly uncertain, a fact they view as arguing against taking action. But uncertainty is a two-edged sword. Climate researchers themselves readily concede that estimates based on their models are extremely uncertain. But that means that although the actual outcome might be much better than their median forecast, it might also be significantly worse.

The Paris Agreement, which was adopted by representatives from nearly every nation on December 12, 2015, seeks to limit global warming to 3.6°F by the end of the twenty-first century. But even an increase that small would cause deadly harm, and the most respected climate change models estimate that there is essentially no chance that average temperature will rise by less than that amount if we take no action. In 2019, the Trump administration submitted a formal notice of intention to withdraw the U.S. from the agreement.

According to estimates published in 2013 by the Intergovernmental Panel on Climate Change (IPCC), the median forecast is for an average global temperature climb of 6.7°F by century's end, in the absence of effective countermeasures.⁴ The IPCC also estimated that we face a 5 percent chance of temperatures rising by more than 8.6°F by 2100. Temperature increases of that magnitude would be accompanied by sea level rises that would make much of the world's most densely populated coastal regions uninhabitable. Climate scientists also believe that severe storms and droughts caused by climate change are already causing enormous damage.⁵

Again, long-term forecasts from climate models are highly uncertain. Things might not be as bad as predicted, but they could also be much worse. In other domains, uncertainty doesn't counsel against taking action. Few would recommend, for example, that we disband the military simply because adversaries might not invade. In the climate arena, the only remaining uncertainty is how much worse things will get. Should we take action? To respond to that question, we must ask, how much would it cost? The answer, as it turns out, is astonishingly little.

Early estimates by the IPCC suggest that a tax of between \$20 and \$80 per ton on carbon emissions would be needed by 2030 to achieve climate stability by 2100. The organization also noted, however, that technical advances might reduce the required levy to as little as \$5 per ton.⁶ Under a carbon tax, the prices of goods would rise in proportion to their carbon footprints. A tax of \$80 per ton, for example, would raise the price of gasoline by about 70 cents a gallon, while a tax of \$5 per ton would raise prices by less than 5 cents a gallon. Even a dramatically higher tax on carbon—say, \$300 a ton—would raise the price of gasoline by less than \$3 a gallon.

As American motorists saw in 2008, a sudden price increase of that magnitude could indeed be painful. But if phased in gradually, it would cause much less harm. Facing steadily increasing fuel prices, for example, manufacturers would scramble to develop more fuel-efficient vehicles. Many Europeans now pay \$4 a gallon more for gas than Americans do. But precisely because of that fact, European automakers

³See the United States Environmental Protection Agency, www.epa.gov/airmarkets.

⁴See Intergovernmental Panel on Climate Change, *Fifth Assessment Report—Climate Change 2013: The Physical Science Basis*, www.ipcc.ch/report/ar5/wg1.

⁵The Environmental Defense Fund, "How Climate Change Plunders the Planet," www.edf.org/climate/how-climate-change-plunders-planet.

⁶See Intergovernmental Panel on Climate Change, *Climate Change 2007—Synthesis Report*, pp. 59–61, www.ipcc.ch/pdf/assessment-report/ar4/syr/ar4_syr.pdf.

have pioneered development of many of the world's most fuel-efficient cars. Europeans actually spend less on gas than Americans do, yet seem no less happy with their rides.

If a family traded in its aging Ford Explorer (15 mpg) for a Ford Focus wagon (32 mpg), it would spend less on gas than before, even if it drove just as much. The tax could be phased in slowly, to give people time to adjust. People would also move closer to work, form carpools, choose less distant vacation destinations, and so on. Some of the revenue from the tax could be used to send checks to low-income families to ease the burden of higher gas prices. Portions of it could help pay down debt and rebuild crumbling infrastructure, or reduce other taxes.

In 2009 the U.S. House of Representatives actually passed an energy bill that included a comprehensive carbon cap and trade system, the functional equivalent of a carbon tax. But the bill couldn't win Senate approval, and seasoned congressional observers now see little chance that meaningful climate legislation could win passage in Congress anytime soon.

Some argue that taxing carbon emissions would be a waste of time unless other countries did likewise. It's a fair point. But World Trade Organization officials have indicated that countries could tax imported goods in proportion to their carbon dioxide emissions if exporting countries failed to enact carbon taxes at home. Access to the American market is a potent bargaining chip. Countries that sell in the U.S. would be quick to enact carbon taxes of their own rather than allow the U.S. to reap the revenue from a carbon tax levied on goods imported into the U.S. Moreover, four out of five of America's top trading partners (Canada, Germany, Japan, and Mexico) already have carbon taxes.

In short, the economist's cost-benefit framework suggests that our failure to enact carbon taxes constitutes a mystery of the highest order and is a stark example of how status quo bias and our tendency to discount future events excessively (see Chapter 8, *An Introduction to Behavioral Economics*) can lead to suboptimal policy choices.

RECAP ↑

USING PRICE INCENTIVES IN ENVIRONMENTAL REGULATION

An efficient program for reducing pollution is one for which the marginal cost of abatement is the same for all polluters. Taxing pollution has this desirable property, as does the auction of pollution permits. The auction method has the advantage that regulators can achieve a desired abatement target without having detailed knowledge of the abatement technologies available to polluters.

Climate scientists warn that increasing atmospheric concentrations of greenhouse gases threatens to cause catastrophic global warming. That risk could be averted by imposition of a carbon tax or equivalent carbon permit system.

METHODS OF INCOME REDISTRIBUTION

Although we as a society have an interest in limiting income inequality, programs for reducing it are often fraught with practical difficulties. The challenge is to find ways to raise the living standard of low-income people, without at the same time undermining their incentive to work, and without using scarce resources to subsidize those who are not poor. Of course, some people simply cannot work, or cannot find work that pays enough to live on. In a world of perfect information, the government could make generous cash payments to those people, and withhold support from those who can fend for themselves. In practice, however, the two groups are often hard to

distinguish from each other. And so we must choose among imperfect alternative measures.

Welfare Payments and In-Kind Transfers

in-kind transfer a payment made not in the form of cash, but in the form of a good or service

Cash transfers and in-kind transfers are at the forefront of antipoverty efforts around the globe. **In-kind transfers** are direct transfers of goods or services to low-income individuals or families, such as the Supplemental Nutrition Assistance Program (also known as SNAP or food stamps), public housing, subsidized school lunches, and Medicaid.

From the mid-1960s until 1996, the most important federal program of cash transfers was Aid to Families with Dependent Children (AFDC), which provided cash payments to poor single-parent households. A sharp rise in AFDC caseloads in the early 1990s raised concerns among many lawmakers that AFDC recipients had become dependent on welfare and were trapped in a “cycle of poverty.” Critics of the program also argued that AFDC undermined family stability because in many states a poor mother became ineligible for benefits if her husband or other able-bodied adult male lived with her and her children. This provision confronted many long-term unemployed fathers with an agonizing choice: They could leave their families, making them eligible for public assistance, or they could remain, making them ineligible. Even fathers who deeply loved their families understandably chose to leave.

Personal Responsibility and Work Opportunity Reconciliation Act

the 1996 federal law that transferred responsibility for welfare programs from the federal level to the state level and placed a five-year lifetime limit on payment of TANF benefits to any given recipient

Concern about work incentives led Congress to pass the **Personal Responsibility and Work Opportunity Reconciliation Act** in 1996, which abolished the AFDC program and replaced it with Temporary Assistance for Needy Families (TANF). Under the new law, the federal government makes lump-sum cash grants to the states, which then have broad discretion in determining TANF benefit levels and eligibility requirements. In addition, the new law set a five-year lifetime limit on the receipt of TANF benefits.

Supporters of the Personal Responsibility and Work Reconciliation Act argue that it has already reduced the nation’s welfare rolls considerably and that it will encourage greater self-reliance over the long run. Skeptics fear that the five-year lifetime limit on benefits may eventually impose severe hardships on the poor, especially poor children, if overall economic conditions deteriorate. Indeed, debate continues about the extent to which the observed increases in poverty rates during the economic downturns of 2001 and 2008–2009 were attributable to the Personal Responsibility and Work Opportunity Reconciliation Act. What is clear, however, is that the new law has not eliminated the need to discover efficient ways of providing assistance to low-income families.

Means-Tested Benefit Programs

means-tested a benefit program whose benefit level declines as the recipient earns additional income

Many welfare programs, including TANF, are **means-tested**, which means that the more income a family has, the smaller are the benefits it receives under these programs. The purpose of means testing is to avoid paying benefits to those who don’t really need them. But because of the way welfare programs are administered, means testing often has a pernicious effect on work incentives.

Consider, for example, an unemployed participant in four welfare programs: food stamps, rent stamps, energy stamps, and day care stamps. Each program gives him \$100 worth of stamps per month, which he is then free to spend on food, rent, energy, and day care. If he gets a job, his benefits in each program are reduced by 50 cents for each dollar he earns. Thus, if he accepts a job that pays \$50 weekly, he’ll lose \$25 in weekly benefits from each of the four welfare programs, for a total benefit reduction of \$100 per week. Taking the job thus leaves him \$50 per week worse off than before. Low-income persons need no formal training in economics to realize that seeking gainful employment does not pay under these circumstances.

What is more, means-tested programs of cash and in-kind transfers are extremely costly to administer. If the government were to eliminate all existing welfare and social service agencies that are involved in these programs, the resulting savings would be enough to lift every poor person out of poverty. One proposal to do precisely this is the negative income tax.

The Negative Income Tax

Under the **negative income tax (NIT)**, every man, woman, and child—rich or poor—would receive a substantial income tax credit, say \$5,000 per year. A person who earns no income would receive this credit in cash. People who earn income would receive the same initial credit, and their income would continue to be taxed at some rate less than 100 percent.

The negative income tax would do much less than current programs to weaken work incentives because, unlike current programs, it would ensure that someone who earned an extra dollar would keep at least a portion of it. And because the program would be administered by the existing Internal Revenue Service, administrative costs would be far lower than under the current welfare system.

Despite these advantages, however, the negative income tax is by no means a perfect solution to the income-transfer problem. Although the incentive problem under the program would be less severe than under current welfare programs, it would remain a serious difficulty. To see why, note that if the negative income tax were the *sole* means of insulating people against poverty, the payment to people with no earned income would need to be at least as large as the government's official **poverty threshold**.

The poverty threshold is the annual income level below which a family is officially classified as "poor" by the government. The threshold is based on government estimates of the cost of the so-called economy food plan, the least costly of four nutritionally adequate food plans designed by the Department of Agriculture. The department's 1955 Household Food Consumption Survey found that families of three or more people spent approximately one-third of their after-tax income on food, so the government pegs the poverty threshold at three times the cost of the economy food plan. In 2020, that threshold was approximately \$26,200 for a family of four.

For a family of four living in a city, \$26,200 a year is scarcely enough to make ends meet. But suppose a group of, say, eight families were to pool their negative tax payments and move to the mountains of northern New Mexico. With a total of \$209,600 per year to spend, plus the fruits of their efforts at gardening and animal husbandry, such a group could live very nicely indeed.

Once a small number of experimental groups demonstrated the feasibility of quitting their jobs and living well on the negative income tax, others would surely follow suit. But two practical difficulties would ensue. First, as more and more people left their jobs to live at government expense, the program would eventually become prohibitively costly. And second, the political cost of the program would almost surely force supporters to abandon it long before that point. Reports of people living lives of leisure at taxpayers' expense would be sure to appear on the nightly news. People who work hard at their jobs all day long would wonder why their tax dollars were being used to support those who are capable of holding paying jobs, yet choose not to work. If the resulting political backlash did not completely eliminate the negative income tax program, it would force policymakers to cut back the payment so that members of rural communes could no longer afford to live comfortably. And that would mean the payment would no longer support an urban family. This difficulty has led policymakers to focus on other ways to increase the incomes of the working poor.

Minimum Wages

The United States and many other industrialized countries have sought to ease the burden of low-wage workers by enacting minimum wage legislation—laws that

negative income tax (NIT)

a system under which the government would grant every citizen a cash payment each year, financed by an additional tax on earned income

poverty threshold the level of income below which the federal government classifies a family as poor

prohibit employers from paying workers less than a specified hourly wage. The federal minimum wage in the United States is currently set at \$7.25 per hour, and several states have set minimum wage levels significantly higher. For example, the minimum wage in the state of Washington was \$13.50 per hour for 2020.

At one point, economists were almost unanimous in their opposition to minimum wage laws, arguing that those laws reduce total economic surplus, as do other regulations that prevent markets from reaching equilibrium. Some economists, however, have softened their opposition to minimum wage laws, citing studies that have failed to show significant reductions in employment following increases in minimum wage levels. But as we saw in Chapter 5, *Efficiency, Exchange, and the Invisible Hand in Action*, any policy that prevents a market from reaching equilibrium causes a reduction in total economic surplus—which means society ought to be able to find a more effective policy for helping low-wage workers.

The Earned-Income Tax Credit

earned-income tax credit (EITC) a policy under which low-income workers receive credits on their federal income tax

One such policy is the **earned-income tax credit (EITC)**, which gives low-wage workers a credit on their federal income tax each year. The EITC was enacted into law in 1975, and in the years since has drawn praise from both liberals and conservatives. The program is essentially a wage subsidy in the form of a credit against the amount a family owes in federal income taxes. For example, in 2019, a married couple with two children under the age of 18 and a total annual earnings of \$24,000 could have received an annual tax credit of approximately \$5,800 under this program. That is, the program would have reduced the annual federal income tax payment of this family by roughly that amount. Families who earned more would have received smaller tax credit, with no credit at all for families of four earning more than \$56,000. Families whose tax credit exceeds the amount of tax owed actually receive a check from the government for the difference. The EITC is thus essentially the same as a negative income tax, except that eligibility for the program is confined to people who work.

Like both the negative income tax and the minimum wage, the EITC puts extra income into the hands of workers who are employed at low wage levels. But unlike the minimum wage, the earned-income tax credit creates no incentive for employers to lay off low-wage workers.

Public Employment for the Poor

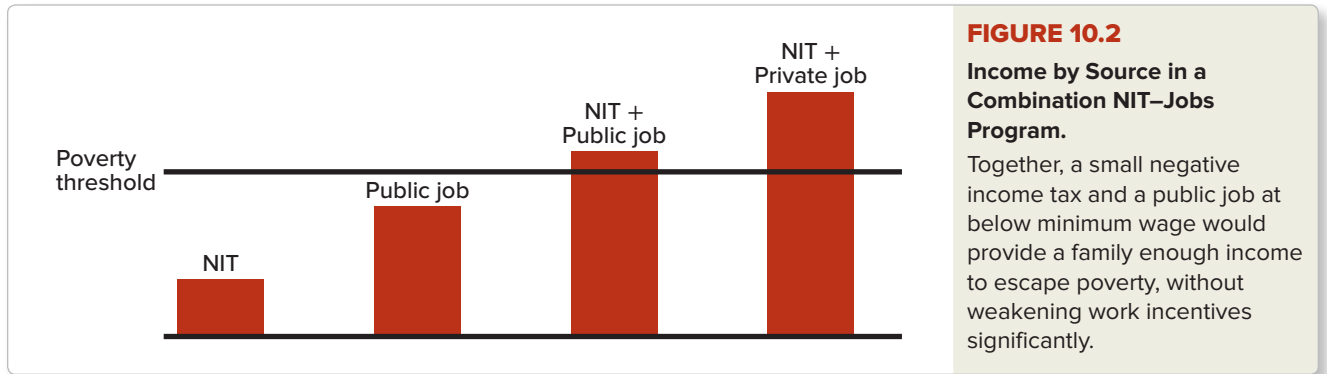
The main shortcoming of the EITC is that it does nothing for the unemployed poor. The negative income tax lacks that shortcoming but may substantially weaken work incentives. There is yet another method of transferring income to the poor that avoids both shortcomings. Government-sponsored jobs could pay wages to the unemployed poor for useful work. With public service employment, the specter of people living lives of leisure at public expense simply does not arise.

But public service employment has difficulties of its own. Evidence shows that if government jobs pay the same wages as private jobs, many people will leave their private jobs in favor of government jobs, apparently because they view government jobs as being more secure. Such a migration would make public service employment extremely expensive. Other worrisome possibilities are that such jobs might involve make-work tasks, and that they would prompt an expansion in government bureaucracy.

Acting alone, government-sponsored jobs for the poor, the EITC, or the negative income tax cannot solve the income-transfer problem. But a combination of these programs might do so.

A Combination of Methods

Consider a negative income tax whose cash grant is far too small for anyone to live on, but that is supplemented if necessary by a public service job at below



minimum wage. Keeping the wage in public service jobs well below the minimum wage would eliminate the risk of a large-scale exodus from private jobs. And while living well on either the negative income tax or the public service wage would be impossible, the two programs together could lift people out of poverty (see Figure 10.2).

To prevent an expansion of the bureaucracy, the government could solicit bids from private management companies to oversee the public service employment program. The fear that this program would inevitably become a make-work project is allayed by evidence that unskilled workers can, with proper supervision, perform many valuable tasks that would not otherwise be performed in the private sector. They can, for example, do landscaping and maintenance in public parks; provide transportation for older adults and people with disabilities; fill potholes in city streets and replace burned-out street lamps; transplant seedlings in erosion control projects; remove graffiti from public places and paint government buildings; recycle newspapers and containers; staff day care centers; and so on.

This combination of a small negative income tax payment and public service employment at a subminimum wage would not be cheap. But the direct costs of existing welfare programs are also large, and the indirect costs, in the form of perverse work incentives and misguided attempts to control prices, are even larger. In economic terms, dealing intelligently with the income-transfer problem may in fact prove relatively inexpensive, once society recognizes the enormous opportunity cost of failing to deal with it intelligently.



Ariel Skelley/Blend Images/Getty Images

Can unskilled workers perform useful public service jobs?

METHODS OF INCOME REDISTRIBUTION

Minimum wage laws reduce total economic surplus by contracting employment. The earned-income tax credit boosts the incomes of the working poor without that drawback, but neither policy provides benefits for those who are not employed.

Other instruments in the battle against poverty include in-kind transfers such as food stamps, subsidized school lunches, Medicaid, and public housing as well as cash transfers such as Temporary Assistance for Needy Families. Because benefits under most of these programs are means-tested, beneficiaries often experience a net decline in income when they accept paid employment.

RECAP ↑

SUMMARY

- Our aim in this chapter has been to apply basic microeconomic principles to a variety of government policy questions. These principles help show how different methods of paying for health care affect the efficiency with which medical services are delivered. In the case of health care, the gains from marginal cost pricing can often be achieved through insurance policies with large deductibles. (LO1, LO2)
- An understanding of the forces that give rise to environmental pollution can help identify those policy measures that will achieve a desired reduction in pollution at the lowest possible cost. Both the taxing of pollution and the sale of transferable pollution rights promote this goal. Each distributes the cost of the environmental cleanup effort so that the marginal cost of pollution abatement is the same for all polluters. (LO3)
- Policies and programs for reducing poverty include minimum wage laws, the earned-income tax credit, food stamps, subsidized school lunches, Medicaid, public housing, and Temporary Assistance for Needy Families. Of these, all but the earned-income tax credit fail to maximize total economic surplus, either by interfering with work incentives or by preventing markets from reaching equilibrium. (LO4)
- The negative income tax works much like the earned-income tax credit, except that it includes those who are not employed. A combination of a small negative income tax and access to public service jobs at sub-minimum wages could ensure adequate living standards for the poor without significantly undermining work incentives. (LO4)

KEY TERMS

adverse selection
 earned-income tax credit (EITC)
 first-dollar insurance coverage
 health maintenance organization (HMO)

in-kind transfer
 means-tested
 negative income tax (NIT)

Personal Responsibility and Work Opportunity Reconciliation Act
 poverty threshold

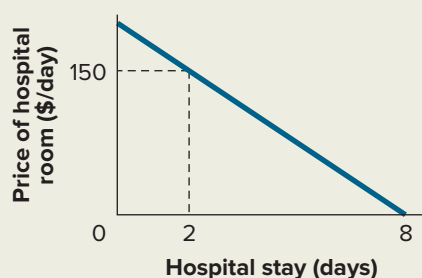
REVIEW QUESTIONS

1. Why is vaccination against many childhood illnesses a legal requirement for entry into public schools? (LO1)
2. Why is first-dollar health care coverage inefficient? (LO2)
3. Why do economists believe that pollution taxes and effluent permits are a more efficient way to curb pollution than laws mandating across-the-board cutbacks? (LO3)
4. Why is a negative income tax, by itself, unlikely to be successful policy for maintaining the living standards of the poor? (LO4)

PROBLEMS

McGraw Hill **connect**

1. In the event he requires an appendectomy, David's demand for hospital accommodations is as shown in the diagram below. David's current insurance policy fully covers the cost of hospital stays. The marginal cost of providing a hospital room is \$150 per day. (LO1, LO2)



- a. If David's only illness this year results in an appendectomy, how many days will he choose to stay in the hospital?
 - b. How many days would David have chosen to stay in the hospital if his insurance covered only the cost of hospital stays that exceed \$1,000 per illness? Explain why that choice would have failed to satisfy the cost-benefit test.
2. David's willingness to pay for hospital recovery time following his upcoming appendectomy is \$300 for the first day, \$200 for the second, \$100 for the third, \$50 for the fourth, and nothing for days beyond four. (LO2)

- a. If David is uninsured and the cost of providing a hospital room is \$250 per day, how many days will he stay in the hospital following his operation?
 - b. How many days would he stay if his insurance covered half the cost of each hospital day?
 - c. Describe a change in the terms of his insurance policy that would benefit both David and his insurer.
3. Two firms, Sludge Oil and Northwest Lumber, have access to five production processes, each one of which has a different cost and gives off a different amount of pollution. The daily costs of the processes and the corresponding number of tons of smoke emitted are as shown in the following table. (LO3)

Process (smoke)	A (4 tons/day)	B (3 tons/day)	C (2 tons/day)	D (1 ton/day)	E (0 tons/day)
Cost to Sludge Oil (\$/day)	50	70	120	200	500
Cost to Northwest Lumber (\$/day)	100	180	500	1,000	2,000

- a. If pollution is unregulated, which process will each firm use, and what will be the total daily smoke emission?
 - b. The city council wants to curb smoke emissions by 50 percent. To accomplish this, it requires each firm to curb its emissions by 50 percent. What will be the increase in total cost to society due to this policy?
 - c. The city council again wants to curb emissions by half. This time, it sets a tax of \$ T per day on each ton of smoke emitted. How large will T have to be to effect the desired reduction? What is the total cost to society of this policy?
4. Refer to Problem 3. Instead of taxing pollution, the city council decides to auction off four permits, each of which entitles the bearer to emit 1 ton of smoke

per day. No smoke may be emitted without a permit. Suppose the government conducts the auction by starting at \$1 and asking how many permits each firm wants to buy at that price. If the total is more than four, it then raises the price by \$1 and asks again, and so on, until the total quantity of demanded permits falls to four. How much will each permit sell for in this auction? How many permits will each firm buy? What will be the total cost to society of this reduction in pollution? (LO3)

5. A senator has argued that we should not spend hundreds of billions of dollars to reduce greenhouse gas emissions because we are not even sure that such

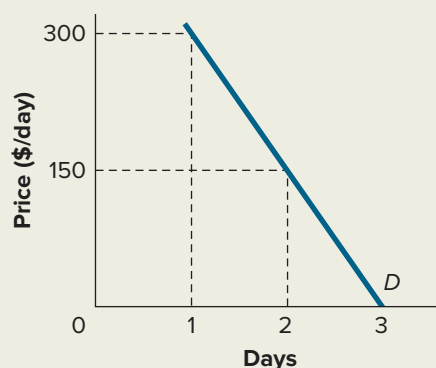
emissions will lead to costly changes in the Earth's climate. Comment critically on this argument. (LO3)

- 6* Suppose employers and workers are risk-neutral, and Congress is about to enact the \$12-per-hour minimum wage. Congressional staff economists have urged legislators to consider adopting an earned-income tax credit instead. Suppose neither workers nor employers would support that proposal unless the expected value of each party's economic surplus would be at least as great as under the minimum wage. Describe an earned-income tax credit (and a tax that would raise enough money to pay for it) that would receive unanimous support from both workers and employers. (LO4)

*Denotes more difficult problem.

ANSWERS TO SELF-TESTS

- 10.1 With 50 percent coverage, David would have to pay \$150 for each additional day in the hospital, so he would choose to stay for two days. (LO2)



10.2 With a tax of \$61 per ton each day, Sludge Oil would adopt process A and Northwest Lumber would adopt process C. (LO3)

Process (smoke)	A (4 tons/day)	B (3 tons/day)	C (2 tons/day)	D (1 ton/day)	E (0 tons/day)
Cost to Sludge Oil (\$/day)	100	200	600	1,300	2,300
Cost to Northwest Lumber (\$/day)	300	320	380	480	700

11

LEARNING OBJECTIVES

After reading this chapter, you should be able to:

- LO1** Explain and apply the concept of comparative advantage and explain how it differs from absolute advantage.
- LO2** Explain how the price of a tradable good is set in a closed versus an open economy, how the quantities of imports or exports are determined, and discuss who are the winners and losers from trade.
- LO3** Illustrate why trade is often politically controversial, even though it promises to increase total income.

International Trade and Trade Policy

On April 13, 1861, Southern troops fired on Fort Sumter in Charleston harbor, initiating the American Civil War. Less than a week later, on April 19, President Lincoln proclaimed a naval blockade of the South. Code-named the Anaconda Plan (after the snake that squeezes its prey to death), the blockade required the Union navy to patrol the Southern coastline, stopping and boarding ships that were attempting to land or depart. The object of the blockade was to prevent the Confederacy from shipping cotton to Europe, where it could be traded for military equipment, clothing, foodstuffs, and other supplies.

Historians are divided on the effectiveness of the Union blockade in choking off Confederate trade. In the early years of the war, the North had too few ships to cover the 3,600-mile Southern coastline, so “running” the blockade was not difficult. But in the latter part of the war the number of Union ships enforcing the blockade increased from about 90 to more than 600, and sailing ships were replaced with faster, more lethal ironclad vessels. Still, private blockade runners attempted to elude the Union navy in small, fast ships. Because the price of raw cotton in Great Britain was between 10 and 20 times what it was in the Confederacy (a differential that indicated disruption in the normal flow of trade), blockade runners enjoyed huge profits when they were successful. But despite their efforts, by 1864 the Southern war effort was seriously hampered by a lack of military equipment and supplies, at least in part as a result of the blockade.

The use of a naval blockade as a weapon of war highlights a paradox in contemporary attitudes toward trade between nations. Presumably, an attempt by a foreign power to blockade U.S. ports today would be considered a hostile act that would elicit a strong response from the U.S. government. Yet one often hears politicians and others arguing that trade with other nations is harmful to the United States and should be restricted—in effect, that the United States should blockade its own ports! In the U.S. presidential election of 2016, for example, both candidates opposed the Trans-Pacific Partnership (TPP)—an agreement intended to increase trade between Australia, Canada, Japan, Mexico, the United States, and seven other, smaller

economies—on the grounds that it would cost American jobs. As another example, politicians from both the Republican and Democratic Parties occasionally complain about China's exports to the United States. Some of them propose taking action against the Chinese government because, they say, it engages in unfair policies that cause Chinese products to sell in the U.S. for prices that are too *cheap*. So is trade a good thing or not? And if it is, why does it sometimes face determined opposition?

This chapter addresses international trade and its effects on the broader economy. We will begin by reviewing the idea of *comparative advantage*. We will show that everyone can enjoy more goods and services if nations specialize in those products in which they have a comparative advantage, and then trade freely among themselves. Furthermore, if trade is unrestricted, market forces will ensure that countries produce those goods in which they have a comparative advantage.

Having shown the potential benefits of trade, we will turn next to the reasons for opposition to trade. Although opening the economy to trade *can*, in our model, increase economic welfare overall, some groups—such as workers in industries that face competition from foreign producers—may in reality be made worse off. The fact that open trade may hurt some groups creates political pressure to enact measures restricting trade, such as taxes on imported goods (called tariffs) and limits on imports (called quotas). We will analyze the effects of these trade restrictions, along with other ways of responding to concerns about affected industries and workers through direct assistance programs.

Economists understand the frustrations with trade expressed by some groups. Indeed, economists recognize that for a displaced worker who does not expect effective direct government assistance to materialize—for example, because the worker does not expect the government to pass, implement, and sustain the relevant measures—opposing trade may appear a reasonable course of action. However, from an economic point of view that takes into account both the overall size of the economic pie and how it is distributed, providing direct assistance to those who are hurt by increased trade is preferable to blocking or restricting trade.

COMPARATIVE ADVANTAGE AS A BASIS FOR TRADE

One of the most important insights of modern economics is that when two people (or two nations) have different opportunity costs of producing different goods and services, they can always increase the total value of available goods and services by trading with one another. Recall from Chapter 1, *Thinking Like an Economist*, that the opportunity cost of spending more time (or other resources) on any one activity is having less time (or other resources) available to spend on others. To illustrate this insight, we will start with an example.

In Example 11.1, economists would say that Wearstler has an **absolute advantage** at designing her web page but a **comparative advantage** at interior design. She has an absolute advantage at designing her web page because she can perform that task in less time than a web designer could. Even so, the web designer has a comparative advantage at designing web pages because his or her opportunity cost of performing that task is lower than Wearstler's.

Example 11.1 makes the implicit assumption that Wearstler would be equally happy to spend an hour designing her web page or working on an interior design project. But suppose she is tired of interior design work and feels it might be enjoyable to try her hand at web design. Designing her own web page might then make perfect sense! But unless she expects to gain extra satisfaction from performing that task, she almost certainly will do better to hire a web designer. The web designer will also benefit, or else he or she wouldn't have offered to design web pages for the stated price.

absolute advantage one person has an absolute advantage over another if he or she takes fewer hours to perform a task than the other person

comparative advantage one person has a comparative advantage over another if his or her opportunity cost of performing a task is lower than the other person's opportunity cost

EXAMPLE 11.1 **Scarcity Principle****Should Kelly Wearstler design her own web page?**

Kelly Wearstler is among the most famous and influential interior designers in the United States today. She has received numerous accolades for her commercial and residential design work, has completed projects for top celebrities such as Cameron Diaz, Gwen Stefani, and Ben Stiller, and boasts more than 700,000 followers on Instagram.

Although Wearstler devotes most of her time and talent to interior design, she is well equipped to do a broad range of other design work. Suppose she could design her own web page in 300 hours, half the time it would take any other web designer. Does that mean that Wearstler should design her own web page?

Suppose that on the strength of her talents as an interior designer, Wearstler earns more than \$1 million a year, implying that the opportunity cost of any time she spent designing her web page would be over \$500 per hour. She would have little difficulty engaging a highly qualified web designer whose hourly wage is considerably less than \$500 per hour. So even though Wearstler's substantial skills might enable her to design her web page more quickly than most web designers, it would not be in her interest to do so.



Amy Graves/Getty Images

Should Kelly Wearstler design her own web page?

In summary, this example demonstrates that by specializing in the activities at which they have comparative advantage (i.e., in the activities with lowest opportunity cost), Wearstler and another web designer could both gain. Indeed, the gains made possible from specialization based on comparative advantage constitute the rationale for market exchange. By concentrating on those tasks at which we are relatively most productive, together we can produce vastly more than if we all tried to be self-sufficient. (In addition, our relative productivity in those activities we concentrate on can itself increase, for example, through specialized training and experience—in turn further increasing our comparative advantage in those activities.) The next example illustrates more concretely how these productivity gains come about.

EXAMPLE 11.2 **Comparative Advantage****Should Ana update her own web page?**

Consider a small community in which Ana is the only professional bicycle mechanic and Xin is the only professional web designer. If the amount of time for each of them to update a web page and repair a bicycle is as shown in Table 11.1, and if each regards the two tasks as equally pleasant (or unpleasant), does the fact that Ana can update a web page faster than Xin imply that Ana should update her own web page?

TABLE 11.1**Productivity Information for Ana and Xin**

	Time to update a web page	Time to complete a bicycle repair
Ana	20 minutes	10 minutes
Xin	30 minutes	30 minutes

The entries in the table show that Ana has an absolute advantage over Xin in both activities. While Ana, the mechanic, needs only 20 minutes to update a web page, Xin, the web designer, needs 30 minutes. Ana's advantage over Xin is even greater when the task is fixing bikes: She can complete a repair in only 10 minutes, compared to Xin's 30 minutes.

But the fact that Ana is a better web designer than Xin does *not* imply that Ana should update her own web page. Xin has a comparative advantage over Ana at web design: She is *relatively* more productive at updating a web page than Ana. Similarly, Ana has a comparative advantage in bicycle repair. (Remember that a person has a comparative advantage at a given task if his or her opportunity cost of performing that task is lower than another person's.)

What is Xin's opportunity cost of updating a web page? Because she takes 30 minutes to update each page—the same amount of time she takes to fix a bicycle—her opportunity cost of updating a web page is one bicycle repair. In other words, by taking the time to update a web page, Xin is effectively giving up the opportunity to do one bicycle repair. Ana, in contrast, can complete two bicycle repairs in the time she takes to update a single web page. For her, the opportunity cost of updating a web page is two bicycle repairs. Ana's opportunity cost to update a web page, measured in terms of bicycle repairs forgone, is twice as high as Xin's. Thus, Xin has a comparative advantage at web design.

The interesting and important implication of the opportunity cost comparison summarized in Table 11.2 is that the total number of bicycle repairs and web updates accomplished if Xin and Ana both spend part of their time at each activity will always be smaller than the number accomplished if each specializes in the activity in which she has a comparative advantage. Suppose, for example, that people in their community demand a total of 16 web page updates per day. If Ana spent half her time updating web pages and the other half repairing bicycles, an 8-hour workday would yield 12 web page updates and 24 bicycle repairs. To complete the remaining 4 updates, Xin would have to spend 2 hours web designing, which would leave her 6 hours to repair bicycles. And because she takes 30 minutes to do each repair, she would have time to complete 12 of them. So when the two women each try to be a jill of all trades, they end up completing a total of 16 web page updates and 36 bicycle repairs.

TABLE 11.2
Opportunity Costs for Ana and Xin

	Opportunity cost of updating a web page	Opportunity cost of a bicycle repair
Ana	2 bicycle repairs	0.5 web page update
Xin	1 bicycle repair	1 web page update

Consider what would have happened had each woman specialized in her activity of comparative advantage. Xin could have updated 16 web pages on her own and Ana could have performed 48 bicycle repairs. Specialization would have created an additional 12 bicycle repairs out of thin air.

When computing the opportunity cost of one good in terms of another, we must pay close attention to the form in which the productivity information is presented. In Example 11.2, we were told how many minutes each person needed to perform each task. Alternatively, we might be told how many units of each task each person can perform in an hour. Work through the following self-test to see how to proceed when information is presented in this alternative format.

SELF-TEST 11.1

Should Miguel update his own web page?

Consider a small community in which Miguel is the only professional bicycle mechanic and Monique is the only professional web designer. If their productivity rates at the two tasks are as shown in the table, and if each regards the two tasks as equally pleasant (or unpleasant), does the fact that Miguel can update a web page faster than Monique imply that Miguel should update his own web page?

	Productivity in web design	Productivity in bicycle repair
Monique	2 web page updates per hour	1 repair per hour
Miguel	3 web page updates per hour	3 repairs per hour

This insight, that specialization and trade among individuals can yield impressive gains in productivity, applies equally well to nations. Factors such as climate, natural resources, technology, workers' skills and education, and culture provide countries with comparative advantages in the production of different goods and services. For example, the large number of leading research universities in the United States gives that nation a comparative advantage in the design of technologically sophisticated computer and mobile hardware and software. Likewise, the wide international use of the English language endows the United States with a comparative advantage in producing popular films and TV shows. Similarly, France's climate and topography, together with the accumulated knowledge of generations of vintners, provides that country a comparative advantage in producing fine wines, while Australia's huge expanses of arable land give that country a comparative advantage in producing grain.

We can all enjoy more goods and services when each country produces according to its comparative advantage, and then trades with other countries. So, while in reality software, films, TV shows, wines, and grain are all produced by each of the three countries, respectively (the U.S., France, and Australia), by concentrating on producing the specific products each country has a comparative at and then trading with other countries, consumers in each of these countries can enjoy a greater variety and quantity of these products.

However, there is an important difference between applying the principle of comparative advantage to individuals and applying it to nations: That a nation as a whole—in the aggregate—enjoys more goods and services does not imply that *everyone* in the nation enjoys more goods and services. Rather, it is possible that within the nation, some individuals will gain from opening up to trade, while others will lose. So when we say that in the transition from a **closed economy**—one that does not trade with the rest of the world—to a more **open economy**—one that does trade with other economies—everybody *can* gain, we are not saying that everybody *will* gain. For everybody to gain, or be better off, the winners from trade will have to share a sufficient amount of their gains with the losers. In itself, opening up trade does not guarantee that the gains will be shared. We return to these issues in the next section.



Climate and long experience give France a comparative advantage in producing fine wines.

closed economy an economy that does not trade with the rest of the world

open economy an economy that trades with other countries

RECAP ↑

COMPARATIVE ADVANTAGE AS A BASIS OF TRADE

Gains from exchange are possible if trading partners have comparative advantage in producing different goods and services. You have a comparative advantage in producing, say, web pages if your opportunity cost of producing a web page—measured in terms of other production opportunities forgone—is smaller than the corresponding opportunity costs of your trading partners. Maximum production is achieved if each person specializes in producing the good or service in which he or she has the lowest opportunity cost. This comparative advantage makes specialization worthwhile even if one trading partner is more productive than others, in absolute terms, in every activity.

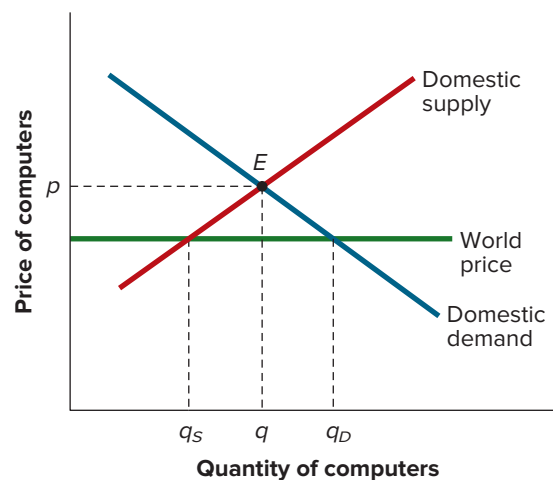
A SUPPLY AND DEMAND PERSPECTIVE ON TRADE

In this section we will look more carefully at how international trade affects supply and demand in the markets for specific goods. We will see that when it is costly for workers and firms to change industries, opening up trade with other countries may create groups of winners and losers among producers, workers, and consumers.

Let's see how trade affects the markets for computers and coffee in a hypothetical economy that produces only these two goods. We will call this economy Costa Rica. Figure 11.1 shows the supply and demand for computers in that economy. As usual, the price is shown on the vertical axis and the quantity on the horizontal axis. For now, think of the price of computers as being measured in terms of coffee rather than in terms of dollars (in other words, we measure the price of computers *relative* to the price of the other good in the economy). As usual, the upward-sloping curve in Figure 11.1 is the supply curve of computers, in this case for computers produced in Costa Rica; and the downward-sloping curve is the demand curve for computers by Costa Rican residents. The supply curve for computers in Costa Rica reflects the opportunity cost of supplying computers (see Chapter 2, *Supply and Demand*). Specifically, at any level of computer production, the relative price at which Costa

FIGURE 11.1**The Market for Computers in Costa Rica.**

If Costa Rica is closed to international trade, the equilibrium price and quantity of computers are determined by the intersection of the domestic supply and demand curves at point *E*. But if Costa Rica is open to trade, the domestic price of computers must equal the world price. At that price, Costa Ricans will demand q_D computers, but domestic producers will supply only q_S computers. Thus $q_D - q_S$ computers must be imported from abroad.



Rican firms are willing to supply an additional computer equals their opportunity cost of doing so. The demand curve, which tells us the number of computers Costa Ricans will purchase at each relative price, reflects the preferences and buying power of Costa Rican consumers.

If the Costa Rican economy is closed to international trade, then market equilibrium occurs where the domestic supply and demand curves intersect, at point E in Figure 11.1. The equilibrium price will be p and the equilibrium quantity, q .

If Costa Rica opens its market to trade, however, the relevant price for computers becomes the **world price** of computers, the price at which computers are traded internationally. The world price for computers is determined by the worldwide supply and demand for computers. If we assume that Costa Rica's computer market is too small to affect the world price for computers very much, the world price can be treated as fixed, and represented by a horizontal line in the figure. Figure 11.1 shows the world price for computers as being lower than Costa Rica's closed-economy price.

If Costa Ricans are free to buy and sell computers on the international market, then the price of computers in Costa Rica must be the same as the world price. (No one in Costa Rica will buy a computer at a price above the world price, and no one will sell one at a price below the world price.) Figure 11.1 shows that at the world price, Costa Rican consumers and firms demand q_D computers, but Costa Rican computer producers will supply only q_S computers. The difference between the two quantities, $q_D - q_S$, is the number of computers that Costa Rica must import from abroad. Figure 11.1 illustrates a general conclusion: *If the price of a good or service in a closed economy is greater than the world price, and that economy opens itself to trade, the economy will tend to become a net importer of that good or service.*

A different outcome occurs in Costa Rica's coffee market, shown in Figure 11.2. The price of coffee (measured relative to the price of computers) is shown on the vertical axis, and the quantity of coffee on the horizontal axis. The downward-sloping demand curve in the figure shows how much coffee Costa Rican consumers want to buy at each relative price, and the upward-sloping supply curve how much coffee Costa Rican producers are willing to supply at each relative price. If Costa Rica's economy is closed to trade with the rest of the world, then equilibrium in the market for coffee will occur at point E , where the domestic demand and supply curves intersect. The quantity produced will be q and the price p .

world price the price at which a good or service is traded on international markets

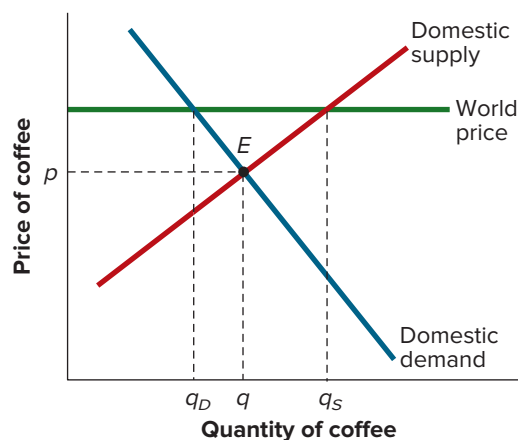


FIGURE 11.2

The Market for Coffee in Costa Rica.

With no international trade, the equilibrium price and quantity of coffee in Costa Rica are determined by the intersection of the domestic supply and demand curves (point E). But if the country opens to trade, the domestic price of coffee must equal the world price. At the higher world price, Costa Ricans will demand the quantity of coffee q_D , less than the amount supplied by Costa Rican producers, q_S . The excess coffee supplied by Costa Rican producers, $q_S - q_D$, is exported.

Now imagine that Costa Rica opens its coffee market to international trade. As in the case of computers, if free trade in coffee is permitted, then the prevailing price for coffee in Costa Rica must be the same as the world price. Unlike the case of computers, however, the world price of coffee as shown in Figure 11.2 is *higher* than the domestic equilibrium price. How do we know that the world price of coffee will be higher than the domestic price? Recall that the price of coffee is measured relative to the price of computers, and vice versa. If the price of computers relative to the price of coffee is higher in Costa Rica than in the world market, then the price of coffee relative to the price of computers must be lower, as each price is the reciprocal of the other. More generally, when two people or two countries trade with each other, neither can have a comparative advantage in *every* good and service. Thus, in an example with only two goods, if non-Costa Rican producers have a comparative advantage in computers, reflected in the lower cost of computers relative to coffee in the world market, then Costa Rican producers must have a comparative advantage in coffee. By definition, this comparative advantage implies that the opportunity cost of coffee in terms of computers must be lower in Costa Rica than in the rest of the world.

Figure 11.2 shows that at the world price for coffee, Costa Rican producers are willing to supply q_S coffee, while Costa Rican consumers want to purchase a smaller amount, q_D . The difference between domestic production and domestic consumption, $q_S - q_D$, is exported to the world market. The general conclusion of Figure 11.2 is this: *If the price of a good or service in a closed economy is lower than the world price, and that economy opens itself to trade, the economy will tend to become a net exporter of that good or service.*

These examples illustrate how the market translates comparative advantage into mutually beneficial gains from trade. If trade is unrestricted, then countries with a comparative advantage in a particular good will profit by supplying that good to the world market and using the revenue earned to import goods in which they do not have a comparative advantage. Thus the workings of the free market automatically ensure that goods will be produced where the opportunity cost is lowest, leading to the highest possible consumption possibilities for the world as a whole.

Winners and Losers from Trade

If trade is so wonderful, why do politicians so often resist free trade and “globalization”? The reason is that although free trade benefits the economy as a whole, specific groups may not benefit. Groups who are hurt by trade, and the politicians they elect, may support policies that restrict the free flow of goods and services across borders.

The supply and demand analyses shown in Figures 11.2 and 11.3 are useful in clarifying who gains and who loses when an economy opens up to trade. Look first at Figure 11.1, which shows the market for computers in Costa Rica. When Costa Rica opens its computer market to international competition, Costa Rican consumers enjoy a larger quantity of computers at a lower price. Clearly, Costa Rican computer users benefit from the free trade in computers. In general, *domestic consumers of imported goods benefit from free trade.* However, Costa Rican computer producers and their workers will not be so happy about opening their market to international competition. The fall in computer prices to the international level implies that less efficient domestic producers will go out of business, and that those who remain will earn lower profits. Unemployment in the Costa Rican computer industry will rise and may persist over time, particularly if displaced computer workers cannot easily move to a new industry. The wages paid to Costa Rican computer workers will also fall, reflecting the lower relative price of computers. We see that, in general, *domestic producers of imported goods, and workers in imported-good sectors, are hurt by free trade.*

Consumers are helped, while producers and workers are hurt, when imports increase. The opposite conclusions apply for an increase in exports (see Figure 11.3). In the example of Costa Rica, an opening of the coffee market raises the domestic price of coffee to the world price and creates the opportunity for Costa Rica to export coffee. Domestic producers of coffee benefit from the increased market (they can now sell coffee abroad as well as at home) and from the higher price of their product. Domestic workers in the coffee sector benefit too, as they are more in demand. In short, *domestic producers of exported goods, and workers in exported-good sectors, benefit from free trade*. Costa Rican coffee drinkers will be less enthusiastic, however, since they must now have to pay the higher world price of coffee, and can therefore consume less. *Thus domestic consumers of exported goods are hurt by free trade*.

Free trade is *efficient* in the sense that it increases the size of the pie available to the economy. Indeed, the efficiency of free trade is an application of the *equilibrium principle*: Markets in equilibrium leave no unexploited opportunities for individuals. Despite the efficiency of free trade, however, some groups may lose from trade, which generates political pressures to block or restrict trade. In the next section we will discuss the major types of policy used to restrict trade.

The Economic Naturalist 11.1



What is the China trade shock?

The China trade shock, a term most commonly associated with economists David Autor, David Dorn, and Gordon Hanson, is used to describe the dramatic change in international trade patterns that resulted from China's rise as a major player in the global economy over the past few decades.

In a series of influential studies, these economists and their collaborators investigated the costs of the shock to U.S. workers. They found that employment has fallen in U.S. industries and regions most exposed to import competition from China—something that our theory in this chapter helps explain. However, they did not find strong evidence of simultaneous offsetting employment increases in other sectors in the same regions, suggesting that the transition of workers into sectors in which the U.S. has comparative advantage has been neither quick nor easy—something that the theory does not emphasize.

Overall, these economists conclude that workers' adjustment to trade shocks is often a slow and difficult process, and that local labor force participation rates and unemployment rates in affected regions may take a decade or more to recover. Moreover, the slow adjustment means that trade shocks could lead to prolonged economic and social problems in affected communities.

While the research underlying these conclusions is still new and is still being examined, it serves as a reminder that for many workers, the short-term costs of trade may outweigh the short-term benefits. While opposition to trade among such workers is understandable, we should not conclude that the overall costs of trade outweigh the overall benefits. Rather, as these economists conclude in one of their studies:

Better understanding when and where trade is costly, and how and why it may be beneficial, is a key item on the research agenda for trade and labor economists. Developing effective tools for managing and mitigating the costs of trade adjustment should be high on the agenda for policymakers and applied economists.¹

¹David Autor, David Dorn, and Gordon Hanson, "The China Shock: Learning from Labor Market Adjustment to Large Changes in Trade," *Annual Review of Economics* 8 (October 2016), pp. 205–240.

RECAP ↑**A SUPPLY AND DEMAND PERSPECTIVE ON TRADE**

- For a closed economy, the domestic supply and demand for a good or service determine the equilibrium price and quantity of that good or service.
- In an open economy, the price of a good or service traded on international markets equals the world price. If the domestic quantity supplied at the world price exceeds the domestic quantity demanded, the difference will be exported to the world market. If the domestic quantity demanded at the world price exceeds the domestic quantity supplied, the difference will be imported.
- Generally, if the price of a good or service in a closed economy is lower than the world price and the economy opens to trade, the country will become a net exporter of that good or service. If the closed-economy price is higher than the world price and the economy opens to trade, the country will tend to become a net importer of the good or service.
- Consumers of imported goods, producers of exported goods, and workers in exported-good sectors benefit from trade, whereas consumers of imported goods, producers of imported goods, and workers in imported-good sectors are hurt by trade. Those groups that are hurt, and the politicians they elect, may support enacting barriers to trade.

TRADE WINNERS AND LOSERS**Winners**

- Consumers of imported goods
- Producers of exported goods
- Workers in exported-good sectors

Losers

- Consumers of imported goods
- Producers of exported goods
- Workers in imported-good sectors

protectionism the view that free trade is injurious and should be restricted

tariff a tax imposed on an imported good

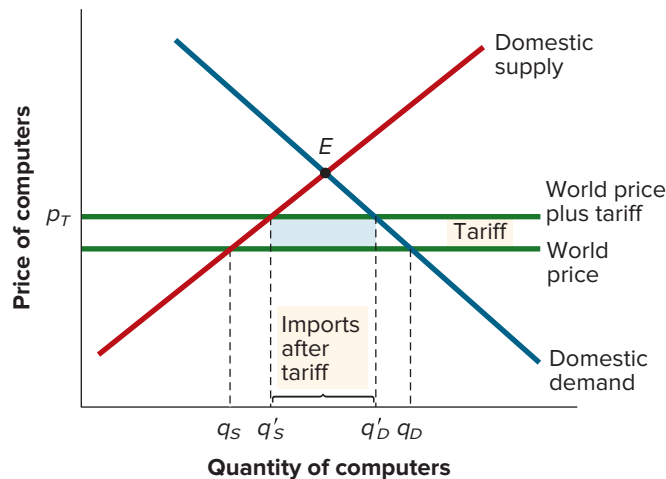
quota a legal limit on the quantity of a good that may be imported

PROTECTIONIST POLICIES: TARIFFS AND QUOTAS

The view that free trade is injurious and should be restricted is known as **protectionism**. Supporters of this view believe the government should attempt to “protect” domestic markets by raising legal barriers to imports. (Interestingly, protectionists rarely attempt to restrict exports, even though they hurt consumers of the exported good.) Two of the most common types of such barriers are tariffs and quotas. A **tariff** is a tax imposed on an imported good. A **quota** is a legal limit on the quantity of a good that may be imported.

Tariffs

The effects of tariffs and quotas can be explained using supply and demand diagrams. Suppose that Costa Rican computer makers, dismayed by the penetration of “their” market by imported computers, persuade their government to impose a tariff—that is, a tax—on every computer imported into the country. Computers

**FIGURE 11.3****The Market for Computers after the Imposition of an Import Tariff.**

The imposition of a tariff on imported computers raises the price of computers in Costa Rica to the world price plus tariff, p_T , represented by the upper horizontal line. Domestic production of computers rises from q_S to q'_S , domestic purchases of computers fall from q_D to q'_D , and computer imports fall from $q_D - q_S$ to $q'_D - q'_S$. Costa Rican consumers are worse off and Costa Rican computer producers are better off. The Costa Rican government collects revenue from the tariff equal to the area of the pale blue rectangle.

produced in Costa Rica will be exempt from the tax. Figure 11.3 shows the likely effects of this tariff on the domestic Costa Rican computer market. The lower of the two horizontal lines in the figure indicates the world price of computers, not including the tariff. The higher of the two lines indicates the price Costa Rican consumers will actually pay for imported computers, including the tariff. We refer to the price of computers including the tariff as p_T . The vertical distance between the two lines equals the amount of the tariff that is imposed on each imported computer.

From the point of view of domestic Costa Rican producers and consumers, the imposition of the tariff has the same effects as an equivalent increase in the world price of computers. Because the price (including the tariff) of imported computers has risen, Costa Rican computer producers will be able to raise the price they charge for their computers to the world price plus tariff, p_T . Thus the price Costa Rican consumers must pay—whether their computers are imported or not—equals p_T , represented by the upper horizontal line in Figure 11.3.

The rise in the price of computers created by the tariff affects the quantities of computers supplied and the quantities demanded by Costa Ricans. Domestic computer producers, facing a higher price for computers, increase their production from q_S to q'_S (see Figure 11.3). Costa Rican consumers, also reacting to the higher price, reduce their computer purchases from q_D to q'_D . As a result, the number of imported computers—the difference between domestic purchases and domestic production—falls from $q_D - q_S$ to $q'_D - q'_S$.

Who are the winners and the losers from the tariff, then? Relative to an environment with free trade and no tariff, the winners are the domestic computer producers, who sell more computers and receive a higher price for them. The clearest losers are Costa Rican consumers, who must now pay more for their computers. Another winner is the government, which collects revenue from the tariff. The blue area in Figure 11.3 shows the amount of revenue the government collects, equal to the quantity of computer imports after the imposition of the tariff, $q'_D - q'_S$, times the amount of the tariff.

McGraw Hill **connect**

► Visit your instructor's Connect course and access your eBook to view this video.



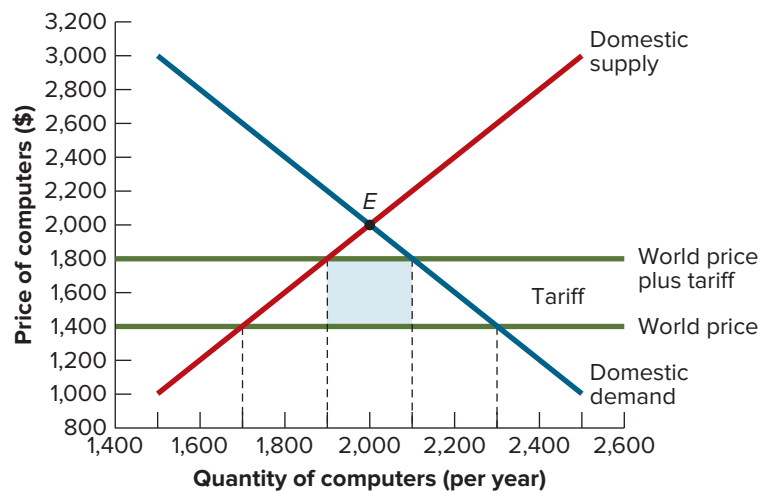
Why do consumers in the United States often pay more than double the world price for sugar?

EXAMPLE 11.3 A Tariff on Imported Computers**What are the effects of a tariff on trade?**

Suppose that the Costa Rican market for computers is represented by Figure 11.4. Thus, if the Costa Rican economy is closed to trade, the equilibrium price for computers would be \$2,000, and 2,000 computers would be bought and sold in the Costa Rican computer market every year (point *E*). Assuming that the world price of computers is \$1,400, how would this market be affected by opening to trade, and how would it be affected by the imposition of a tariff of \$400 per computer?

FIGURE 11.4**The Market for Computers in Costa Rica after the Imposition of an Import Tariff.**

A tariff of \$400 per computer raises the price of computers by \$400 and reduces imports by 400 computers per year.



If the economy opens to trade, the domestic price of computers must equal the world price of \$1,400. At this price, the domestic quantity demanded is 2,300 computers per year and the domestic quantity supplied is 1,700 computers per year. Imports equal the difference between the domestic quantities demanded and supplied, or $2,300 - 1,700 = 600$ computers per year.

The imposition of a tariff of \$400 per computer raises the price from \$1,400 to \$1,800. This price rise causes Costa Rican computer producers to increase their production from 1,700 to 1,900 computers per year, and it causes Costa Rican consumers to reduce their computer purchases from 2,300 to 2,100. As a result, the number of imported computers—the difference between domestic purchases and domestic production—falls from 600 to 200 ($2,100 - 1,900$).

Thus the tariff has raised the price of computers by \$400 and reduced imports by 400 computers per year. The tariff revenue collected by the government is \$400 per imported computer times 200 computers per year = \$80,000 per year.

SELF-TEST 11.2

Repeat Example 11.3 under the assumption that the world price of computers is \$1,200 (the tariff is still \$400). What happens if the world price is \$1,800?

Quotas

An alternative to a tariff is a quota, or legal limit, on the number or value of foreign goods that can be imported. One means of enforcing a quota is to require importers to obtain a license or permit for each good they bring into the country. The

government then distributes exactly the same number of permits as the number of goods that may be imported under the quota.

How does the imposition of a quota on, say, computers affect the domestic market for computers? Figure 11.5, which is similar to Figure 11.3, illustrates the effect of a quota on imported computers. As before, assume that at first there are no restrictions on trade. Consumers pay the world price for computers, and $q_D - q_S$ computers are imported. Now suppose once more that domestic computer producers complain to the government about competition from foreign computer makers, and the government agrees to act. However, this time, instead of a tariff, the government imposes a quota on the number of computers that can be imported. For comparability with the tariff analyzed in Figure 11.3, let's assume that the quota permits the same level of imports as entered the country under the tariff: specifically, $q'_D - q'_S$ computers. What effect does this ruling have on the domestic market for computers?

After the imposition of the quota, the quantity of computers supplied to the Costa Rican market is the production of domestic firms plus the $q'_D - q'_S$ imported computers allowed under the quota. Figure 11.5 shows the quantity of computers supplied inclusive of the quota. The total supply curve, labeled "Domestic supply plus quota," is the same as the domestic supply curve shifted $q'_D - q'_S$ units to the right. The domestic demand curve is the same as in Figure 11.3. Equilibrium in the domestic market for computers occurs at point F in Figure 11.5, at the intersection of the supply curve including the quota and the domestic demand curve. The figure shows that, relative to the initial situation with free trade, the quota (1) raises the domestic price of computers above the world price, to the level marked p_T in Figure 11.5; (2) reduces domestic purchases of computers from q_D to q'_D ; (3) increases domestic production of computers from q_S to q'_S ; and (4) reduces imports to $q'_D - q'_S$, consistent with a quota. Like a tariff, the quota helps domestic producers by increasing their sales and the price they receive for their output, while hurting domestic consumers by forcing them to pay a higher price.

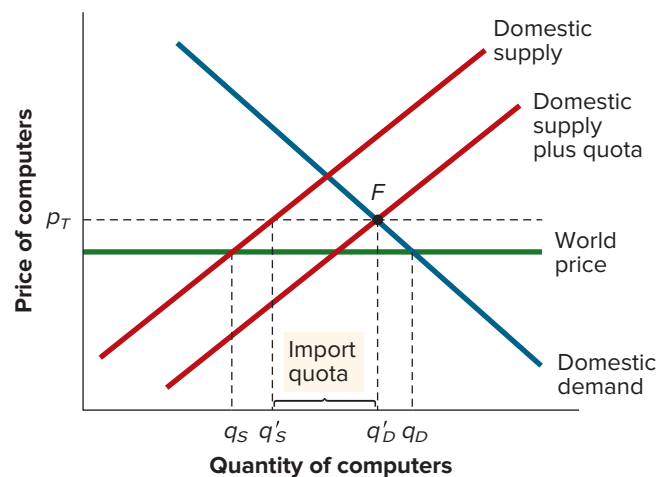


FIGURE 11.5

The Market for Computers after the Imposition of an Import Quota.

The figure shows the effects of the imposition of a quota that permits only $q'_D - q'_S$ computers to be imported. The total supply of computers to the domestic economy equals the domestic supply curve shifted to the right by $q'_D - q'_S$ units (the fixed amount of imports). Market equilibrium occurs at point F . The effects of the quota on the domestic market are identical to those of the tariff analyzed in Figure 11.3. The domestic price rises to p_T , domestic production of computers rises from q_S to q'_S , domestic purchases of computers fall from q_D to q'_D , and computer imports fall from $q_D - q_S$ to $q'_D - q'_S$. The quota differs from the tariff in that under a quota system the government collects no revenue.

Interestingly, under our assumption that the quota is set to permit the same level of imports as the tariff, the effects on the domestic market of the tariff (Figure 11.3) and the quota (Figure 11.5) are not only similar, they are *equivalent*. Comparing Figures 11.3 and 11.5, you can see that the two policies have identical effects on the domestic price, domestic purchases, domestic production, and imports.

Although the market effects of a tariff and a quota are the same, there is one important difference between the two policies, which is that a tariff generates revenue for the government, whereas a quota does not. With a quota, the revenue that would have gone to the government goes instead to those firms that hold the import licenses. A holder of an import license can purchase a computer at the world price and resell it in the domestic market at price p_T , pocketing the difference. Thus with a tariff the government collects the difference between the world price and the domestic market price of the good; with a quota, private firms or individuals collect that difference. Why then would the government ever impose a quota rather than a tariff? One possibility is that the distribution of import licenses is a means of rewarding the government's political supporters. Sometimes, international political concerns may also play a role (see The Economic Naturalist 11.2 for a possible example).

EXAMPLE 11.4 Effects of an Import Quota

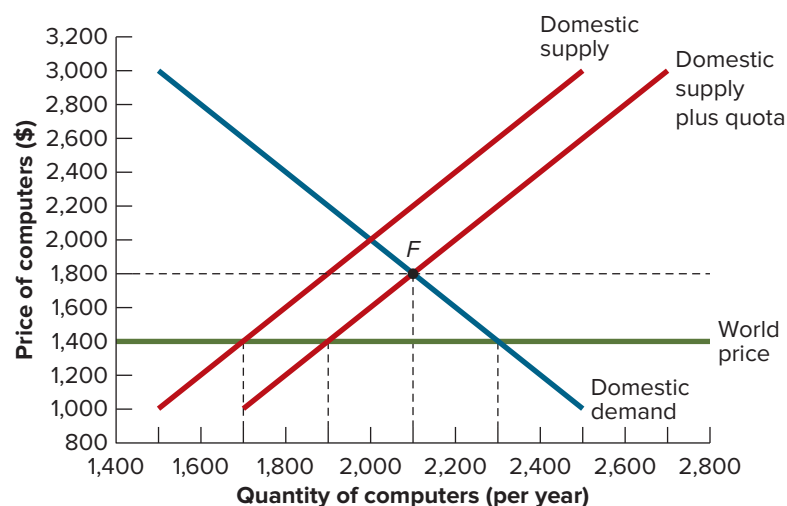
What are the effects of an import quota on trade?

Suppose the supply of and demand for computers in Costa Rica, as well as the world price of computers, are as given in Figure 11.6. Suppose that the government imposes a quota of 200 on the number of computers that can be imported. What effect would this have on the domestic market for computers (relative to the free-trade alternative)?

FIGURE 11.6

The Market for Computers in Costa Rica after the Imposition of an Import Quota.

A quota of 200 computers per year raises the price of computers by \$400 and reduces imports by 400 computers per year.



After the quota is imposed, the equilibrium in the domestic market occurs at point F in the figure. As the figure shows, relative to free trade, the domestic price increases to \$1,800 per computer, domestic purchases of computers decrease to 2,100, domestic production of computers increases to 1,900, and imports decrease to 200 (the difference between the 2,100 computers demanded and the 1,900 computers domestically produced).

Note that the domestic price, domestic production, and domestic demand are the same in Examples 11.3 and 11.4. Thus, under the assumptions we made, the tariff and the quota have the same effects on the domestic market for computers. The only difference between the two policies is that with a quota, the government does not get the tariff revenue it got in Example 11.3. That revenue goes instead to the holders of import licenses, who can buy computers on the world market at \$1,400 and sell them in the domestic market at \$1,800.

The Economic Naturalist 11.2



Why did the U.S. start a trade war with China?

Starting in 2018 and continuing during 2019, the U.S. imposed several rounds of tariffs on commodities and products imported from China. The tariffs, ranging from 10 to 25 percent, were applied to imports worth hundreds of billions of dollars in total. In retaliation, China imposed tariffs of 5 to 25 percent on more than a hundred billion dollars' worth of goods imported from the United States.

As our analysis in this chapter predicts, studies show that the U.S. tariffs were largely reflected in higher prices for American consumers and for American producers who buy the affected goods.² The price of these goods increased not only when they were imported, but also, through reduced competition, when they were produced in the United States. The Chinese tariffs also hurt American producers who export to China. The overall costs of the trade war to Americans were thus substantial. Why did the United States start it?



Songquan Deng/Shutterstock

Why did the U.S. start a trade war with China?

One stated reason for imposing the tariffs was that U.S. trade negotiators were concerned that China was violating U.S. intellectual property rights, for example, by not enforcing patent laws or by requiring U.S. firms to share technology if they want to do business in China. Loss of intellectual property disadvantages U.S. producers in the long run. Concerns about worker protections, environmental protections, and regulatory differences have also historically influenced trade policy.

Another stated reason for the tariffs was national security. Traditionally, countries have emphasized maintaining food and defense independence through local production. Recently, the COVID-19 pandemic also has highlighted the vulnerability of countries that rely on trading partners for medications and medical equipment.

While the U.S.–China trade war is in principle still ongoing, a first ceasefire was reached in January 2020, when the United States and China signed a Phase One trade deal, and when some of the tariffs were reduced. At around the same time, the two countries turned to fighting the COVID-19 pandemic and its economic fallout—which shifted these governments' focus away from the trade war. While it remains to be seen how trade relations between the United States and China further develop, a range of political and economic issues between the two countries, going beyond those analyzed in this chapter, are likely to affect the negotiations.

Tariffs and quotas are not the only barriers to trade that governments erect. Importers may be subject to unnecessarily complex bureaucratic rules (so-called

²See, for example: Mary Amiti, Stephen J. Redding, and David Weinstein, "The Impact of the 2018 Trade War on U.S. Prices and Welfare," NBER Working Paper No. 25672 (March 2019); and Pablo D. Fajgelbaum, Pınelopi K. Goldberg, Patrick J. Kennedy, and Amit K. Khandelwal, "The Return to Protectionism," NBER Working Paper No. 25638 (October 2019).

red tape barriers), and regulations of goods that are nominally intended to promote health and safety sometimes have the side effect, whether intentionally or unintentionally, of restricting trade. One example is European restrictions on imports of genetically modified foods. Although these regulations were motivated in part by concerns about the safety of such foods, they also help protect Europe's politically powerful farmers from foreign competition.

The Inefficiency of Protectionism

Free trade is efficient because it allows countries to specialize in the production of goods and services in which they have the greatest comparative advantage. Conversely, protectionist policies that limit trade are inefficient—they reduce the total economic pie. Why, then, do governments adopt such policies? The reason is similar to why some city governments impose rent controls (see Chapter 2, *Supply and Demand*). Although rent controls reduce economic welfare overall, some people benefit from them—namely, the tenants whose rents are held artificially below market level. Similarly, as we have seen in this section, tariffs and quotas benefit certain groups. Those who benefit from these restrictions (such as firms facing import competition, and the workers in such firms) may support politicians who promise to enact the restrictions.

The fact that free trade is efficient suggests an alternative to trade restrictions, however. Because eliminating restrictions on trade increases the overall economic pie, in general the winners from free trade will be able to compensate the losers in such a way that everyone becomes better off. Government programs that assist workers displaced by import competition are an example of such compensation. They include programs to expand job training and retraining opportunities, especially for the less-educated, who are more likely to be hurt by competition from abroad; to provide transition assistance for displaced workers, including support for internal migration within the U.S. from regions with declining industries to regions with expanding industries; to mitigate residential and educational segregation and increase the access of those left behind to employment and educational opportunities; and to promote community redevelopment in regions with declining industries.

While by no means an easy task, developing and improving such programs, and making them widely available, could do a lot to help those who lose from trade. Spreading the benefits of free trade—or at least reducing its adverse effects on certain groups—reduces the incentives of those groups to oppose free trade.

Although we have focused on the winners and losers from trade, not all opposition to free trade is motivated by economic interest. For example, many who oppose further opening to trade cite environmental concerns. Protecting the environment is an important and laudable goal, but restricting trade may not be the most effective means of achieving that goal. Restricting trade lowers world income, reducing the resources available to deal with environmental problems. Furthermore, much of the income loss arising from barriers to trade is absorbed by poor nations trying to develop their economies. For this reason, leaders of developing countries are among the strongest advocates of free trade.



The Economic Naturalist 11.3

What is fast track authority?

In practice, trade agreements among countries are very complex. For example, agreements usually spell out in great detail the goods and services for which tariffs are being reduced or quotas are being expanded. Trade negotiators must also take into account barriers to trade other than explicit tariffs or quotas, such as rules that require a country's government to buy only from domestic suppliers.

Because trade negotiations can be so complex, having each country's legislature vote on each item in a proposed trade agreement is not practical.

In the United States, the solution to this problem has been for Congress to vote to give the president *fast track authority*. Under this authority, the executive branch is given discretion to negotiate the terms of a proposed trade agreement. Congress then has the opportunity to vote the agreement up or down, but it cannot amend the proposal or accept only certain parts of it.

Fast track authority has been successfully used by presidents of both parties to negotiate trade agreements. However, the granting of fast track authority itself can be contentious, reflecting political concerns about trade and globalization. In 2015, for example, Democrats strongly resisted President Obama's request for fast track authority to negotiate a trade agreement with a number of countries, and in 2018 some Congress members resisted a similar request by President Trump.

RECAP ↑

PROTECTIONIST POLICIES: TARIFFS AND QUOTAS

- The view that free trade is injurious and should be restricted is called protectionism.
- The two most common types of trade barriers are tariffs, or taxes on imported goods, and quotas, legal limits on the quantity that can be imported. A tariff raises the domestic price to the world price plus the tariff. The result is increased domestic production, reduced domestic consumption, and fewer imports. A quota has effects on the domestic market that are similar to those of a tariff. The main difference is that under a quota, the government does not collect tariff revenue.
- Trade barriers are inefficient; they reduce the overall size of the economic pie. Thus, in general, the winners from free trade should be able to compensate the losers in such a way that everyone becomes better off. Government programs to help workers displaced by import competition are an example of such compensation. They are not easy to develop and implement. Without effective compensation, however, opposition to trade by those who are hurt by it may lead to the inefficiency brought by protectionism.

SUMMARY

- According to the principle of comparative advantage, the best economic outcomes occur when each nation specializes in the goods and services at which it is relatively most productive and then trades with other nations to obtain the goods and services its citizens desire. (LO1)
- In a closed economy, the relative price of a good or service is determined at the intersection of the supply curve of domestic producers and the demand curve of domestic consumers. In an open economy, the relative price of a good or service equals the *world price*—the price determined by supply and demand in the world economy. If the price of a good or service in a closed economy is greater than the world price and the country opens its market to trade, it will become a net importer of that good or service. But if the closed-economy price is below the world price and the country opens itself to trade, it will become a net exporter of that good or service. (LO2)
- Although free trade is beneficial to the economy as a whole, some groups—such as domestic producers of imported goods, and workers in imported-good

sectors—are hurt by free trade. Groups that are hurt by trade may support political candidates who promise to impose *protectionist* measures, such as tariffs or quotas. A *tariff* is a tax on an imported good that has the effect of raising the domestic price of the good. A higher domestic price increases domestic supply, reduces domestic demand, and reduces imports of the good. A *quota*, which is a legal limit on the amount of a good that may be imported, has the same effects as a

tariff, except that the government collects no tax revenue. (The equivalent amount of revenue goes instead to those firms with the legal authority to import goods.) Because free trade is efficient, the winners from free trade should be able to compensate the losers so that everyone becomes better off. Thus policies to assist those who are harmed by trade, such as assistance and retraining for workers idled by imports, are usually preferable to trade restrictions. (LO3)

KEY TERMS

absolute advantage
closed economy
comparative advantage

open economy
protectionism
quota

tariff
world price

REVIEW QUESTIONS

1. Explain what “having a comparative advantage” at producing a particular good or service means. What does “having an absolute advantage” at producing a good or service mean? (LO1)
2. A small, open economy is equally productive in producing coffee and tea. What will this economy produce if the world price of coffee is twice that of tea? Half that of tea? What will the country produce if the world price of coffee happens to equal the world price of tea? (LO2)
3. True or false: If a country is more productive in every sector than a neighboring country, then there is no benefit in trading with the neighboring country. Explain. (LO2)
4. Suppose France has a comparative advantage in cheese production and Britain has a comparative advantage in bicycle manufacturing. How would you expect French bicycle manufacturers and British dairy farmers to react to a proposal to reduce trade barriers between Britain and France? (LO1, LO2)
5. Show graphically the effects of a tariff on imported automobiles on the domestic market for automobiles. Who is hurt by the tariff and why? Who benefits and why? (LO3)
6. Show graphically the effects of a quota on imported automobiles on the domestic market for automobiles. Who does the quota hurt and who benefits? Explain. (LO3)

PROBLEMS



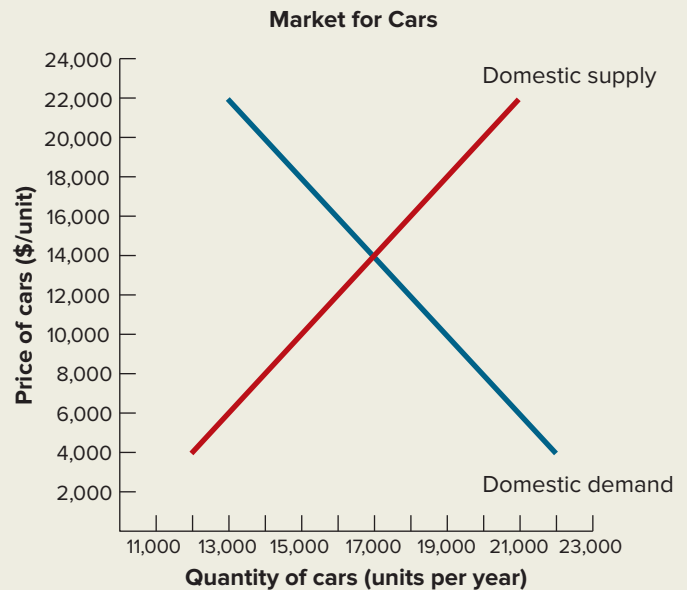
1. Ted can wax a car in 20 minutes or wash a car in 60 minutes. Ishana can wax a car in 15 minutes or wash a car in 30 minutes. What is each person’s opportunity cost of washing a car? Who has a comparative advantage in washing cars? (LO1)
2. Ted can wax 4 cars per day or wash 12 cars. Ishana can wax 3 cars per day or wash 6 cars. What is each person’s opportunity cost of washing a car? Who has a comparative advantage in washing cars? (LO1)
3. Isa and Antonio are auto mechanics. Isa takes 4 hours to replace a clutch and 2 hours to replace a set of brakes. Antonio takes 6 hours to replace a clutch and 2 hours to replace a set of brakes. State whether anyone has an absolute advantage at either task and, for each task, identify who has a comparative advantage. (LO1)
4. An economy has two workers, Bella and Edward. Per day of work, Bella can pick 100 apples or 25 bananas, and Edward can pick 50 apples or

50 bananas. Bella and Edward each work 200 days per year. (LO1)

- Which worker has an absolute advantage in apples? Which has a comparative advantage? Calculate each worker's opportunity cost of picking an additional apple.
 - Find the maximum number of each type of fruit that can be picked annually in this economy, assuming that none of the other type of fruit is picked. What is the most of each type that can be picked if each worker fully specializes according to his or her comparative advantage?
5. Suppose that a U.S. worker can produce 1,000 pairs of shoes or 10 industrial robots per year. For simplicity, assume there are no costs other than labor costs and firms earn zero profits. Initially, the U.S. economy is closed. The domestic price of shoes is \$30 a pair, so that a U.S. worker can earn \$30,000 annually by working in the shoe industry. The domestic price of a robot is \$3,000, so that a U.S. worker can also earn \$30,000 annually working in the robot industry.

Now suppose that the U.S. opens trade with the rest of the world. Foreign workers can produce 500 pairs of shoes or 1 robot per year. The world price of shoes after the U.S. opens its markets is \$10 a pair, and the world price of robots is \$5,000. (LO2, LO3)

- What do foreign workers earn annually, in dollars?
 - When it opens to trade, which good will the United States import and which will it export?
 - Find the real income of U.S. workers after the opening to trade, measured in (1) the number of pairs of shoes annual worker income will buy and (2) the number of robots annual worker income will buy. Compare to the situation before the opening of trade. Does trading in goods produced by "cheap foreign labor" hurt U.S. workers?
 - How might your conclusion in part c be modified in the short term, if it is costly for workers to change industries? What policy response might help with this problem?
6. The demand and supply for automobiles in a certain country is given in the following graph. (LO2, LO3)
- Assuming that the economy is closed, find the equilibrium price and production of automobiles.
 - The economy opens to trade. The world price of automobiles is \$8,000. Find the domestic quantities demanded and supplied and the quantity of imports or exports. Who will favor the opening of the automobile market to trade, and who will oppose it?



- The government imposes a tariff of \$2,000 per car. Find the effects on domestic quantities demanded and supplied.
 - As a result of the tariff, what will happen to the quantity of imports or exports, and what is the revenue raised by the tariff? Who will favor the imposition of the tariff, and who will oppose it?
7. Suppose the domestic demand and supply for automobiles is as given in Problem 6. (LO2, LO3)
- The economy opens to trade. The world price of automobiles is \$10,000. Find the domestic quantities demanded and supplied and the quantity of imports or exports.
 - Now assume that the government imposes a quota on automobile imports of 2,000 cars. What will happen to the quantity of imports or exports?
 - Who will favor the imposition of the quota, and who will oppose it?
8. You are the president of Islandia, a small island nation that enjoys a comparative advantage in tourism. Trade representatives from the United States, which enjoys a comparative advantage in manufactured goods, have proposed a free-trade agreement between the two countries. Manufacturing workers have opposed the agreement, arguing that Islandia should maintain its steep tariff on American manufactured goods. In an election, the union representing these workers has more than enough votes to prevail over the union representing tourism workers. If you are determined to keep your job, how should you respond to the American proposal? (LO3)

ANSWERS TO SELF-TESTS

11.1

	Productivity in web design	Productivity in bicycle repair
Monique	2 web page updates per hour	1 repair per hour
Miguel	3 web page updates per hour	3 repairs per hour

The entries in the table tell us that Miguel has an absolute advantage over Monique in both activities. While Miguel, the mechanic, can update 3 web pages per hour, Monique, the web designer, can update only 2. Miguel's absolute advantage over Monique is even greater in the task of fixing bikes—3 repairs per hour versus Monique's 1.

But as in the second example in this chapter, the fact that Miguel is a better web designer than Monique does not imply that Miguel should update his own web page. Miguel's opportunity cost of updating a web page is 1 bicycle repair, whereas Monique must give up only half a bicycle repair to

update a web page. Monique has a comparative advantage over Miguel at web design and Miguel has a comparative advantage over Monique at bicycle repair. (LO1)

11.2 If the world price of computers is \$1,200, domestic demand for computers is 2,400 computers. Domestic supply is 1,600 computers. The difference between the quantity demanded and the quantity supplied, 800 computers, is imported.

A tariff of \$400 raises the domestic price of computers to \$1,600. Now domestic demand is 2,200 and domestic supply is 1,800. The difference, 400 computers, equals imports. Revenue for the government is (\$400/computer) (400 imported computers) = \$160,000.

If the world price of computers is \$1,800 and there is no tariff, domestic demand is 2,100; domestic supply is 1,900; and imports are 200. A tariff of \$400 raises the world price to \$2,200, which is greater than the domestic price when there is no trade (\$2,000). No computers are imported in this case and no tariff revenue is raised. (LO3)

GLOSSARY

A

- absolute advantage** one person has an absolute advantage over another if he or she takes fewer hours to perform a task than the other person
- accounting profit** the difference between a firm's total revenue and its explicit costs
- adverse selection** the pattern in which insurance tends to be purchased disproportionately by those who are most costly for companies to insure
- allocative function of price** changes in prices direct resources away from overcrowded markets and toward markets that are underserved
- anchoring and adjustment** an estimation technique that begins with an initial approximation (the anchor), which is then modified in accordance with additional available information (the adjustment)
- availability heuristic** a rule of thumb that estimates the frequency of an event by the ease with which it is possible to summon examples from memory
- average benefit** the total benefit of undertaking n units of an activity divided by n
- average cost** the total cost of undertaking n units of an activity divided by n
- average total cost (ATC)** total cost divided by total output

B

- barrier to entry** any force that prevents firms from entering a new market
- basic elements of a game** the players, the strategies available to each player, and the payoffs each player receives for each possible combination of strategies
- buyer's reservation price** the largest dollar amount the buyer would be willing to pay for a good
- buyer's surplus** the difference between the buyer's reservation price and the price he or she actually pays

C

- cartel** a coalition of firms that agree to restrict output for the purpose of earning an economic profit
- cash on the table** an economic metaphor for unexploited gains from exchange
- change in demand** a shift of the entire demand curve
- change in supply** a shift of the entire supply curve
- change in the quantity demanded** a movement along the demand curve that occurs in response to a change in price
- change in the quantity supplied** a movement along the supply curve that occurs in response to a change in price
- closed economy** an economy that does not trade with the rest of the world
- Coase theorem** if, at no cost, people can negotiate the purchase and sale of the right to perform activities that cause externalities, they can always arrive at efficient solutions to the problems caused by externalities

- commitment device** a way of changing incentives so as to make otherwise empty threats or promises credible
- commitment problem** a situation in which people cannot achieve their goals because of an inability to make credible threats or promises
- comparative advantage** one person has a comparative advantage over another if his or her opportunity cost of performing a task is lower than the other person's opportunity cost
- complements** two goods are complements in consumption if an increase in the price of one causes a leftward shift in the demand curve for the other (or if a decrease causes a rightward shift)
- constant (or parameter)** a quantity that is fixed in value
- constant returns to scale** a production process is said to have constant returns to scale if, when all inputs are changed by a given proportion, output changes by the same proportion
- credible promise** a promise to take an action that is in the promiser's interest to keep
- credible threat** a threat to take an action that is in the threatener's interest to carry out
- cross-price elasticity of demand** the percentage by which the quantity demanded of the first good changes in response to a 1 percent change in the price of the second

D

- decision tree (or game tree)** a diagram that describes the possible moves in a game in sequence and lists the payoffs that correspond to each possible combination of moves
- demand curve** a schedule or graph showing the quantity of a good that buyers wish to buy at each price
- dependent variable** a variable in an equation whose value is determined by the value taken by another variable in the equation
- descriptive economic principle** See **positive (descriptive) economic principle**
- dominant strategy** one that yields a higher payoff no matter what the other players in a game choose
- dominated strategy** any other strategy available to a player who has a dominant strategy

E

- earned-income tax credit (EITC)** a policy under which low-income workers receive credits on their federal income tax
- economic efficiency** See **efficiency**
- economic loss** an economic profit that is less than zero
- economic profit (or excess profit)** the difference between a firm's total revenue and the sum of its explicit and implicit costs
- economic rent** that part of the payment for a factor of production that exceeds the owner's reservation price, the price below which the owner would not supply the factor

economic surplus the benefit of taking an action minus its cost

economics the study of how people make choices under conditions of scarcity and of the results of those choices for society

economies of scale *See* **increasing returns to scale**

efficiency (or **economic efficiency**) a condition that occurs when all goods and services are produced and consumed at their respective socially optimal levels

efficient (or **Pareto efficient**) a situation is efficient if no change is possible that will help some people without harming others

elastic the demand for a good is elastic with respect to price if its price elasticity of demand is greater than 1

equation a mathematical expression that describes the relationship between two or more variables

equilibrium a balanced or unchanging situation in which all forces at work within a system are canceled by others

equilibrium price and equilibrium quantity the price and quantity at the intersection of the supply and demand curves for the good

excess demand (or **shortage**) the amount by which quantity demanded exceeds quantity supplied when the price of a good lies below the equilibrium price

excess profit *See* **economic profit**

excess supply (or **surplus**) the amount by which quantity supplied exceeds quantity demanded when the price of the good exceeds the equilibrium price

explicit costs the actual payments a firm makes to its factors of production and other suppliers

external benefit (or **positive externality**) a benefit of an activity received by people other than those who pursue the activity

external cost (or **negative externality**) a cost of an activity that falls on people other than those who pursue the activity

externality an external cost or benefit of an activity

F

factor of production an input used in the production of a good or service

first-dollar insurance coverage insurance that pays all expenses generated by the insured activity

fixed cost the sum of all payments made to the firm's fixed factors of production

fixed factor of production an input whose quantity cannot be altered in the short run

fungibility the property of an entity whose individual units are interchangeable, as money in separate accounts

G

game tree *See* **decision tree**

H

health maintenance organization (HMO) a group of physicians that provides health services to individuals and families for a fixed annual fee

homo economicus the narrowly self-interested, well-informed, highly disciplined, and cognitively formidable actor often assumed in traditional economic models

hurdle method of price discrimination the practice by which a seller offers a discount to all buyers who overcome some obstacle

I

imperfectly competitive firm (or **price setter**) a firm that has at least some control over the market price of its product

implicit costs the opportunity costs of the resources supplied by the firm's owners

in-kind transfer a payment made not in the form of cash, but in the form of a good or service

income effect the change in the quantity demanded of a good that results because a change in the price of a good changes the buyer's purchasing power

income elasticity of demand the percentage by which a good's quantity demanded changes in response to a 1 percent change in income

increasing returns to scale (or **economies of scale**) a production process is said to have increasing returns to scale if, when all inputs are changed by a given proportion, output changes by more than that proportion

independent variable a variable in an equation whose value determines the value taken by another variable in the equation

inelastic the demand for a good is inelastic with respect to price if its price elasticity of demand is less than 1

inferior good a good whose demand curve shifts leftward when the incomes of buyers increase and rightward when the incomes of buyers decrease

invisible hand theory Adam Smith's theory that the actions of independent, self-interested buyers and sellers will often result in the most efficient allocation of resources

J

judgmental and decision heuristics rules of thumb that reduce computation costs

L

law of demand people do less of what they want to do as the cost of doing it rises

law of diminishing returns a property of the relationship between the amount of a good or service produced and the amount of a variable factor required to produce it; the law says that when some factors of production are fixed, increased production of the good eventually requires ever-larger increases in the variable factor

long run a period of time of sufficient length that all the firm's factors of production are variable

loss aversion the tendency to experience losses as more painful than the pleasures that result from gains of the same magnitude

M

macroeconomics the study of the performance of national economies and the policies that governments use to try to improve that performance

marginal benefit the increase in total benefit that results from carrying out one additional unit of an activity

marginal cost the increase in total cost that results from carrying out one additional unit of an activity

marginal revenue the change in a firm's total revenue that results from a one-unit change in output

market the market for any good consists of all buyers or sellers of that good

market equilibrium occurs in a market when all buyers and sellers are satisfied with their respective quantities at the market price

market power a firm's ability to raise the price of a good without losing all its sales

means-tested a benefit program whose benefit level declines as the recipient earns additional income

microeconomics the study of individual choice under scarcity and its implications for the behavior of prices and quantities in individual markets

monopolistic competition an industry structure in which a large number of firms produce slightly differentiated products that are reasonably close substitutes for one another

N

Nash equilibrium any combination of strategy choices in which each player's choice is his or her best choice, given the other players' choices

natural monopoly a monopoly that results from economies of scale (increasing returns to scale)

negative externality *See* **external cost**

negative income tax (NIT) a system under which the government would grant every citizen a cash payment each year, financed by an additional tax on earned income

nominal price the absolute price of a good in dollar terms

nonpositional good a good whose value does not depend heavily on how it compares with other goods in the same category

normal good a good whose demand curve shifts rightward when the incomes of buyers increase and leftward when the incomes of buyers decrease

normal profit the opportunity cost of the resources supplied by the firm's owners, equal to accounting profit minus economic profit

normative economic principle one that says how people should behave

O

oligopoly an industry structure in which a small number of large firms produce products that are either close or perfect substitutes

open economy an economy that trades with other countries

opportunity cost the value of what must be forgone to undertake an activity

P

parameter *See* **constant**

Pareto efficient *See* **efficient**

payoff matrix a table that describes the payoffs in a game for each possible combination of strategies

perfect hurdle a threshold that completely segregates buyers whose reservation prices lie above it from others whose reservation prices lie below it, imposing no cost on those who jump the hurdle

perfectly competitive market a market in which no individual supplier has significant influence on the market price of the product

perfectly discriminating monopolist a firm that charges each buyer exactly his or her reservation price

perfectly elastic demand demand is perfectly elastic with respect to price if price elasticity of demand is infinite

perfectly elastic supply supply is perfectly elastic with respect to price if elasticity of supply is infinite

perfectly inelastic demand demand is perfectly inelastic with respect to price if price elasticity of demand is zero

perfectly inelastic supply supply is perfectly inelastic with respect to price if elasticity is zero

Personal Responsibility and Work Opportunity Reconciliation Act the 1996 federal law that transferred responsibility for welfare programs from the federal level to the state level and placed a five-year lifetime limit on payment of AFDC benefits to any given recipient

positional arms control agreement an agreement in which contestants attempt to limit mutually offsetting investments in performance enhancement

positional arms race a series of mutually offsetting investments in performance enhancement that is stimulated by a positional externality

positional externality this occurs when an increase in one person's performance reduces the expected reward of another's in situations in which reward depends on relative performance

positional good a good whose value depends relatively heavily on how it compares with other goods in the same category

positive (or descriptive) economic principle one that predicts how people will behave

positive externality *See* **external benefit**

poverty threshold the level of income below which the federal government classifies a family as poor

precautionary saving saving for protection against unexpected setbacks such as the loss of a job or a medical emergency

price ceiling a maximum allowable price, specified by law

price discrimination the practice of charging different buyers different prices for essentially the same good or service

price elasticity of demand the percentage change in the quantity demanded of a good or service that results from a 1 percent change in its price

price elasticity of supply the percentage change in quantity supplied that occurs in response to a 1 percent change in price

price setter *See* **imperfectly competitive firm**

price taker a firm that has no influence over the price at which it sells its product

prisoner's dilemma a game in which each player has a dominant strategy, and when each plays it, the resulting payoffs are smaller than if each had played a dominated strategy

profit the total revenue a firm receives from the sale of its product minus all costs—explicit and implicit—incurred in producing it

profit-maximizing firm a firm whose primary goal is to maximize the difference between its total revenues and total costs

protectionism the view that free trade is injurious and should be restricted

pure monopoly the only supplier of a unique product with no close substitutes

Q

quota a legal limit on the quantity of a good that may be imported

R

rational person someone with well-defined goals who tries to fulfill those goals as best he or she can

rationing function of price changes in prices distribute scarce goods to those consumers who value them most highly

real price the dollar price of a good relative to the average dollar price of all other goods

regression to the mean the phenomenon that unusual events are likely to be followed by more nearly normal ones

repeated prisoner's dilemma a standard prisoner's dilemma that confronts the same players repeatedly

representativeness heuristic a rule of thumb according to which the likelihood of something belonging to a given category increases with the extent to which it shares characteristics with stereotypical members of that category

rise *See slope*

run *See slope*

S

satisficing a decision-making strategy that aims for adequate results because optimal results may necessitate excessive expenditure of resources

seller's reservation price the smallest dollar amount for which a seller would be willing to sell an additional unit, generally equal to marginal cost

seller's surplus the difference between the price received by the seller and his or her reservation price

short run a period of time sufficiently short that at least some of the firm's factors of production are fixed

shortage *See excess demand*

slope in a straight line, the ratio of the vertical distance the straight line travels between any two points (*rise*) to the corresponding horizontal distance (*run*)

socially optimal quantity the quantity of a good that results in the maximum possible economic surplus from producing and consuming the good

status quo bias the general resistance to change, often stemming from *loss aversion*

substitutes two goods are substitutes in consumption if an increase in the price of one causes a rightward shift in the demand curve for the other (or if a decrease causes a leftward shift)

substitution effect the change in the quantity demanded of a good that results because buyers switch to or from substitutes when the price of the good changes

sunk cost a cost that is beyond recovery at the moment a decision must be made

supply curve a graph or schedule showing the quantity of a good that sellers wish to sell at each price

surplus *See excess supply*

T

tariff a tax imposed on an imported good

tit-for-tat a strategy for the repeated prisoner's dilemma in which players cooperate on the first move and then mimic their partner's last move on each successive move

total cost the sum of all payments made to the firm's fixed and variable factors of production

total expenditure (or total revenue) the dollar amount that consumers spend on a product ($P \times Q$) is equal to the dollar amount that sellers receive

total revenue *See total expenditure*

total surplus the difference between the buyer's reservation price and the seller's reservation price

tragedy of the commons the tendency for a resource that has no price to be used until its marginal benefit falls to zero

U

unit elastic the demand for a good is unit elastic with respect to price if its price elasticity of demand equals 1

V

variable a quantity that is free to take a range of different values

variable cost the sum of all payments made to the firm's variable factors of production

variable factor of production an input whose quantity can be altered in the short run

vertical intercept in a straight line, the value taken by the dependent variable when the independent variable equals zero

W

Weber-Fechner law the relationship according to which the perceived change in any stimulus varies according to the size of the change measured as a proportion of the original stimulus

workers' compensation a government insurance system that provides benefits to workers who are injured on the job

world price the price at which a good or service is traded on international markets

INDEX

Page numbers followed by n refer to notes.

- Absolute advantage, 262, 264
- Absolute dollar amounts, 6–7
- Abstract models, 5
- ACA (Affordable Care Act of 2010), 206–207, 247–248
- Accounting profit, 120–123
- Acid rain, 251–252
- Active learning, 14
- “Addiction and Backsliding: A Theory of Compulsive Consumption” (Winston), 203n
- Adultery, 204
- “Advancing Safety around the World” (Coulter), 209n
- “The Adventure of Silver Blaze” (Doyle), 7n
- Adverse-selection problem, 246–247
- Advertisements
 - in airline industry, 170–173
 - cigarette, 178–179
 - magazine, 236
 - perfume, 236
 - print vs. digital, 46
- AFDC (Aid to Families with Dependent Children), 254
- Affordable Care Act of 2010 (ACA), 206–207, 247–248
- Agrarian societies, 33
- Agricultural commodities
 - corn market, 54, 121–122, 124–125
 - seasonal price movements in, 54
 - weather conditions and, 51
- Aid to Families with Dependent Children (AFDC), 254
- Ainslie, George, 203n
- Airline industry
 - advertisements in, 170–173
 - business vs. vacation travelers and, 164
 - frequent-flyer coupons and, 8
 - price discrimination in, 164
 - product clustering in, 185
 - regulation of, 127
 - seasonal price movements in, 54
- Air-quality regulations, 224
- Algebraic expression, 63–64, 152
- Allocation of resources, 33, 34, 54, 218–219
- Allocative function of price, 123, 126
- All-you-can-eat restaurants, 9
- American Airlines, 170–173
- Amiti, Mary, 275n
- Anabolic steroids, 232–233
- Anaconda Plan, 261
- Analyze This* (film), 169, 182
- Anchoring and adjustment, 199
- “Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias” (Kahneman et al.), 205n
- Apple, 48
- Arbitration agreements, 234
- Arms control agreements, 234–236
- Arms races, 208–209, 234
- AT&T, 143
- ATC (average total cost), 147–148
- Auctioning, 250–252
- Automated teller machines (ATMs), 16
- Automobile industry. *see also* Gasoline
 - engine size in U.S. vs. Britain, 70
 - four-cylinder vs. eight-cylinder cars, 69–70
 - fuel-efficient cars, 70
 - hybrid sports cars, 180–182
 - mufflers and, 223–224
 - navigation systems in, 15–16
 - oligopoly and, 143
 - price discrimination in, 164
 - price volatility in, 112–114
 - productivity growth in, 91
 - sport utility vehicles, 70
 - traffic laws and, 224
- Autor, David, 269, 269n
- Availability heuristic, 196–197
- Average benefits, 10, 12
- Average costs, 10, 12
- Average total cost (ATC), 147–148
- Axelrod, Robert, 177
- Babashoff, Shirley, 198–199
- Bad decisions. *see* Decision pitfalls
- Barriers to entry/exit, 126–127
- Barriers to trade, 270–276
- Baseball, 234
- Basic elements of games, 170
- Basketball, 114–115, 234
- Baskin-Robbins, 71
- A Beautiful Mind* (film), 172n
- Behavioral economics, 195–214. *see also* Rational choice model
 - anchoring and adjustment in, 199
 - availability heuristic and, 196–197
 - cognitive errors and, 195, 201, 204, 211
 - impulse-control problems and, 202–204
 - judgmental and decision heuristics, 196–200
 - loss aversion and, 205–207
 - misinterpretation of contextual clues in, 200–202
 - Obamacare and, 206–207
 - parable of Buridan’s ass in, 200
 - psychophysics of perception and, 200
 - regression to the mean and, 198–199
 - relative position and, 207–214
 - rent-convenience trade-off and, 201
 - representative heuristic and, 197
 - satisficing and, 195

- Behavioral economics—*Cont.*
 status quo bias and, 206, 207
 Weber-Fechner law and, 200
- Benefits. *see also* Cost-Benefit Principle
 average, 10, 12
 external (positive externalities), 217–219, 226
 marginal, 9–12, 56, 57, 105
- Bennett, Tony, 169, 180, 182
- Beyoncé, 66
- Bezos, Jeff, 3, 68–69
- Bias, status quo, 206, 207, 253
- Big-ticket items, 75
- Blame, 198
- Blockade runners, 261
- BMW automobiles, 70
- Body piercing, 235
- Book publishing industry, 126, 164
- Border patrols, 73
- Bottle manufacturing, 97–103
- Braille dots, on drive-up ATM keypads, 16
- Brinkman, Philip, 206n
- Budget share, 75
- Buffett, Warren, 131
- Buridan's ass, parable of, 200
- Bush, George W., 210
- Business vs. vacation travelers, 164
- Buyer's reservation price, 36, 55, 66
- Buyer's surplus, 55
- California, water shortages in, 67–68
- Cameron, James, 66
- Campaign spending limits, 234
- Canada, health insurance in, 206
- Capital costs, 98, 100, 101
- Capitalist economies, 33–34
- Carbon dioxide (CO₂) emissions, 252–253
- Cars. *see* Automobile industry
- Cartel, defined, 175
- Cartel game, 175–177
- Cash on the table, 55–56, 141, 219
- Cash rebates, 163
- Cash transfers, 254
- Cement industry, 143
- Centralized allocation and control, 33
- Centrally planned economies, 33
- Chamberlin, Edward, 143n
- Changes. *see also* Shifts
 in demand, 44–45
 in price, 35, 108
 in quantity demanded, 44–45
 in quantity supplied, 45
 in supply, 45
- Changing incentives, 187
- Character judgments, 189
- Chickens, economic value to humans, 231
- China
 communist economy in, 33
 trade shock and, 269
 U.S. trade war with, 275
- "The China Shock: Learning from Labor Market Adjustment to Large Changes in Trade" (Autor et al.), 269n
- "Choice and Delay of Reinforcement" (Chung & Herrnstein), 203n
- Choosing the Right Pond* (Frank), 207n
- Chung, Shin-Ho, 203n
- Cigarette smoking, 77, 143, 178–179
- Civil War (1861–1865), 261
- Class size, 1–3
- Clean Air Act Amendments of 1995, 251
- Climate change, 252
- Closed economy, 265–268
- Clustering, 184–186
- Coase, Ronald, 220
- Coase theorem, 219–223
- Coates, Dan, 206n
- Coca-Cola, 186
- Cognitive errors, 195, 201, 204, 211
- Collective action problems, 210, 211
- Collective investment decisions, 228
- Colleges and universities, 1–3
- Commitment device, 186
- Commitment problem, 186–190
- Communist economies, 33, 44
- Comparative advantage, 262–266
- Complements, 45–47, 86
- Computers. *see also* Software industry
 keyboard purchases, 3–4
 laptop purchases, 6–7
 market for, 266–268, 270–272
 operating systems for, 96, 126, 146
 quotas on imports, 273, 274
 tariffs on imports, 270–272
- Confederate States of America, 261
- Constant returns to scale, 146
- Constants, defined, 21
- Consumer surplus, 160, 163
- Consumption, seasonal price movements and, 54
- Consumption possibilities, 268
- "Context-Dependent Preferences" (Tversky & Simonson), 201n
- Contextual clues, misinterpretation of, 200–202
- Convenience stores, 142, 184–185
- Copyrights, 126, 141, 145
- Corn market, 54, 121–122, 124–125
- Coronavirus pandemic (2020), 57, 275
- Corporate earnings forecasters, 131–132
- Cosmetic surgery, 236
- Cost-Benefit Principle
 as abstract model, 5
 application of, 3–5
 buyer's reservation price and, 36
 buy/sell decisions and, 92
 carbon taxes and, 253
 class size and, 2–3
 decision pitfalls and, 6–12
 defined, 2
 economic surplus and, 4, 222
 health care and, 242–244

- hurdle method of price discrimination and, 161
- law of demand and, 66
- marginal costs and benefits in, 9–12
- as normative economic principle, 13
- opportunity cost and, 4–5
- as positive economic principle, 13
- price elasticity of demand and, 75
- profit maximization and, 101, 152
- recycling soft drink containers and, 92–93
- shared living expenses and, 221–222
- socially optimal quantity and, 56
- splitting economic surplus and, 223
- status quo bias and, 206
- third-party payment system and, 243–244
- unequal rent amounts and, 223
- wants (tastes) and, 48
- Costs. *see also* Marginal costs; Opportunity costs; Prices
 - average, 10, 12
 - average total, 147–148
 - capital, 98, 100, 101
 - explicit, 4, 120, 123
 - external (negative externalities), 217–219, 225–226
 - fixed, 98, 147–150
 - of health care, 243–248
 - implicit, 4, 7–8, 12, 120, 123
 - start-up, 147–150
 - sunk, 9, 12
 - total, 99, 101, 147–148
 - variable, 101, 147
- Cost-saving innovations, 127–128
- CO₂ (carbon dioxide) emissions, 252–253
- Coulter, Scott, 209n
- COVID-19 pandemic (2020), 57, 275
- Credible threats and promises, 182–183
- Crimes of passion, 204
- Criticism, 198
- Cross-price elasticity of demand, 86–87
- Cuba, communist economy in, 33
- Curry, Steph, 114, 138
- Decision making
 - buy/sell decisions, 92
 - economic models in, 5
 - game theory for, 170–173
 - with incomplete information, 195–196
 - pitfalls in, 6–12
 - trial and error in, 5
- Decision pitfalls
 - failing to think at margins, 8–12
 - ignoring implicit costs, 7–8
 - measuring in proportions vs. absolute dollar amounts, 6–7
- Decision trees, 181–183
- Demand, 65–87. *see also* Demand curve;
 - Price elasticity of demand; Supply and demand
 - change in, 44–45
 - cross-price elasticity of, 86–87
 - elastic, 74, 81
 - excess, 39–42, 133
 - graphing, 40
 - income effect and, 70–71
 - income elasticity of, 86–87
 - inelastic, 74, 81
 - law of, 66
 - needs vs. wants and, 67
 - origins of, 66–67
 - seasonal price movements and, 54
 - social forces influencing, 67
 - substitution effect and, 68–70
 - unit elastic, 74, 81
 - wants in relation to, 66–67
- Demand curve
 - change in demand and, 44–45
 - change in quantity demanded and, 44–45
 - complements and, 45–47
 - defined, 35, 91
 - equations for, 63
 - future expectations and, 48
 - horizontal addition, 71–72
 - horizontal interpretation of, 36
 - for imperfectly competitive firms, 144
 - income and, 47, 48
 - individual, 71–72
 - inferior goods and, 48
 - market, 71–72, 82, 104
 - monetary and nonmonetary costs on, 65
 - normal goods and, 48
 - for perfectly competitive firms, 96–97, 144
 - perfectly elastic and inelastic, 81–82
 - shifts in, 45–48, 51–53
 - slope of, 35–36, 144
 - straight-line, 80–81, 85
 - substitutes and, 46–47
 - vertical interpretation of, 36, 133
- Demand elasticity. *see* Price elasticity of demand
- “The Demand for Unobservable and Other Nonpositional Goods” (Frank), 207n
- DeNiro, Robert, 169
- Departures from rational choice with regret, 201, 204, 211
- Departures from traditional rational choice without regret, 211
- Dependent variables, 21–22
- Deposits, 42
- Descriptive economic principle, 13
- Developing countries, 276
- Digital vs. print advertisements, 46
- Diminishing returns, law of, 98, 101, 103
- Discount pricing, 156–165. *see also* Price discrimination
- Discrimination, in housing market, 42
- Dishonesty, 188, 189
- Displaced workers, 276
- Dominant strategy, 171–173
- Dominated strategy, 171
- “Don’t Burn Your Bridge before You Come to It: Seven Types of Ambiguity in Precommitment” (Elster), 203n
- Dorn, David, 269, 269n
- Doyle, Arthur Conan, 7, 7n

- Earned-income tax credit (EITC), 137, 256
- Economic efficiency. *see* Efficiency
- Economic loss, 122, 143, 154–155
- Economic models, 5
- Economic naturalism
- auto manufacturers and navigation systems, 15–16
 - California water shortages, 67–68
 - car engine size in U.S. vs. Britain, 70
 - cartel agreements, 175–177
 - China trade shock, 269
 - cigarette advertising on television, 178–179
 - cigarette taxes and teenage smoking, 77
 - convenience stores, 184–185
 - corporate earnings forecasters, 131–132
 - defined, 14
 - drive-up ATM keypad buttons with Braille dots, 16
 - fast track authority, 276–277
 - federal pay raises and rent increases, 47
 - football players and anabolic steroids, 232–233
 - four-cylinder vs. eight-cylinder cars, 69–70
 - free speech laws, 225
 - fruit growing in public places, 230
 - gasoline price volatility, 112–114
 - hardware manufacturers and free software, 15
 - health maintenance organizations vs. conventional
 - health insurance, 245–246
 - housing market and wealthy individuals, 68–69
 - Intel microprocessors, 150
 - luxury tax on yachts, 77–78
 - Obamacare, 206–207
 - privatization of Social Security, 210
 - recycling aluminum vs. glass beverage
 - containers, 104
 - Sears “Scratch ‘n’ Dent Sale,” 164
 - seasonal price movements, 54
 - shared milkshakes, 231
 - shouting at parties, 179–180
 - Sports Illustrated* jinx, 198–199
 - student discounts, 157
 - subsidizing planting of trees, 226
 - supermarket checkout lines, 127
 - term-paper revisions, 51
 - toilet paper shortages, 57
 - U.S. trade war with China, 275
 - voting behavior, 211–213
 - waiting lines in poor neighborhoods, 71
 - weight-loss camps, 203
- Economic pie, 137, 155, 220, 276
- Economic profit
- accounting profit vs., 120–123
 - of cigarette makers, 178, 179
 - cost-saving innovations and, 128
 - defined, 120
 - economic rent vs., 129–130
 - monopolistic competition and, 142–143
 - negative (loss), 122, 143, 154–155
 - new entrants into market and, 125
 - substitutes for exclusive inputs and, 146
 - zero-profit tendency, 126
- Economic rent, 129–130
- Economics. *see also* Behavioral economics
- defined, 2
 - of health care, 242
 - macroeconomics, 13
 - microeconomics, 13
 - normative vs. positive, 13
- The Economics of Imperfect Competition* (Robinson), 143n
- Economic surplus
- benefits to society, 135
 - Cost-Benefit Principle and, 4, 222
 - defined, 4
 - efficiency and, 220
 - with income transfers, 137
 - market equilibrium price and, 132–134
 - perfect discrimination and, 160
 - price controls and, 136–137
 - property rights and, 230
 - shared living expenses and, 222
 - socially optimal quantity and, 56
 - splitting, 223
 - subsidies and, 136–137
- Economic systems
- capitalist, 33–34
 - centrally planned, 33
 - communist, 33, 44
 - free-market, 33–34
 - mixed, 33
 - monopolistic competition, 142–143, 184–186
 - natural monopoly, 146
 - oligopoly, 143, 175–177
 - pure monopoly, 142, 143
- “An Economic Theory of Self-Control” (Thaler & Shefrin), 203n
- Economic trade-offs, 1
- Economies of scale, 146–150
- “The Economist Who Realized How Crazy We Are” (Lewis), 205n
- Economy
- closed, 265–268
 - network, 146
 - open, 265–269
- Education, in colleges and universities, 1–3
- “The Effect of the Luxury Excise Tax on the Sale of Luxury Boats” (Zimmerman), 77n
- Efficiency
- Coase theorem and, 220
 - defined, 56
 - economic pie and, 155, 220
 - economic surplus and, 220
 - equilibrium and, 54–58
 - free trade and, 269, 276
 - “good” compared to, 134
 - importance of, 135
 - inefficiency, 155, 219, 276
 - market equilibrium and, 132–135
 - social, 159, 220–221
 - social justice and, 135

- Efficiency loss, 164
- Efficient, defined, 132
- EITC (earned-income tax credit), 137, 256
- Elastic demand, 74, 81
- Elasticity. *see* Price elasticity of demand;
Price elasticity of supply
- Elephants, extinctions threats for, 231
- Elster, Jon, 203n
- Employment. *see also* Labor market
output–employment relationship, 97–100
public service, 43, 256, 257
- Entertainment industry
mobility of performers in, 111–112
movie stars in, 115
price discrimination in, 164
price elasticity and movie tickets, 82–84
product clustering in, 185
student discounts in, 157
- Entrapment laws, 204
- Entry into market, 125–127
- Environmental issues. *see also* Pollution
acid rain, 251–252
carbon dioxide emissions, 252–253
climate change, 252
international trade and, 276
price incentives for regulation, 248–253
- Equations
constructing, 21–22, 26–27
defined, 21
demand curve, 63
deriving from graphs, 23–24
graphing, 22–23
simultaneous, 28–29, 64
supply curve, 63
transforming tables into, 26–27
from verbal descriptions, 21–22
- Equilibrium
defined, 38
efficiency and, 54–58
free trade and, 269
games in, 172–173
market, 38–44
Nash, 172–173
opportunities for gain and, 130
social optimum vs., 130–135
unexploited opportunities and, 130
- Equilibrium price, 38–41, 45–54
- Equilibrium quantity, 38–41, 45–53
- Europe, restrictions on imports of genetically
modified foods, 276. *see also specific countries*
- Excess demand, 39–42, 133
- Excess profit. *see* Economic profit
- Excess supply, 39, 40, 133
- Exchange, cash on the table in, 55–56
- Exit from market, 126–127
- Expectations about the future, 48, 51, 108
- Expenditures, total, 82–86
- Explicit costs, 4, 120, 123
- Exponential discounting, 202n
- Exports, 269
- External benefits. *see* Positive externalities
- External costs. *see* Negative externalities
- Externalities, 217–236
Coase theorem and, 219–223
compensatory taxes and subsidies, 226
defined, 217
inefficiency and, 219
invisible hand theory and, 217–218, 228
laws and regulations, 223–225
negative, 217–219, 225–226
optimal amount of negative externalities
not zero, 225–226
positional, 232–236
positive, 217–219, 226
property rights and tragedy of the commons, 227–232
remedies for, 223–226
resource allocation and, 218–219
- Factors of production, 97–98
- Fajgelbaum, Pablo D., 275n
- Fashion norms, 235
- Fast track authority, 276–277
- Faulty reasoning. *see* Decision pitfalls
- Federal pay raises and rent increases, 47
- Finder's fees, 42
- First Amendment, 225
- First-dollar coverage, 244–245
- First impressions, in job interviews, 67
- Fixed costs, 98, 147–150
- Fixed factor of production, 98
- Food distribution systems, 31–32
- Football, 232–234
- Ford Excursion, 70
- Forecasting corporate earnings, 131–132
- "The Framing of Decisions and the Psychology
of Choice" (Tversky & Kahneman), 196n
- France, 206, 265
- Franchises, 145–146
- Frank, Robert H., 207n
- Frederick, Shane, 203n
- Free entry and exit, 126–127
- Free-market economies, 33–34
- Free speech laws, 225
- Free trade, 268–269, 276. *see also* International trade
- Freeway speed, invisible hand and, 127
- Frequent-flyer coupons, 8
- Fruit growing in public places, 230
- Fuel. *see* Gasoline
- Fuel-efficient cars, 70
- Fungibility, 196
- Future expectations, 48, 51, 108
- Gambling, 204
- Game theory, 169–190. *see also* Prisoner's dilemma
basic elements of games, 170
cartel game, 175–177
commitment problem in, 186–190
credible threats and promises in, 182–183

- Game theory—*Cont.*
 decision (game) trees in, 181–183
 for decision making, 170–173
 dominant/dominated strategy in, 171–173
 incentives and, 172–173, 182, 186–190
 Nash equilibrium in, 172–173
 payoff matrix in, 170–174, 177, 233
 remote-office game, 183, 186, 188
 sequential games, 184
 timing issues in, 180–186
- Game trees, 181–183
- Gasoline
 local retailing, 142, 144
 price controls for, 42
 shortages of, 42, 69
 size of automobiles and, 69–70
 subsidies for, 136–137
 taxation on, 70, 241
 volatility of prices, 112–114
- Gates, Bill, 129
- Genetically modified foods, 276
- Gilovich, Thomas, 198–199, 199n
- Glass bottle manufacturing, 97–103
- Global warming, 252
- Goldberg, Pinelopi K., 275n
- Goods. *see also* Imports
 big-ticket items, 75
 exports, 269
 inferior, 48
 market demand curve for, 82
 market price of, 34–35
 nonpositional, 208, 209, 211
 normal, 48
 positional, 208, 209
 private, 211
 public, 211
 socially optimal quantity of, 56
 total daily expenditure on, 82
- Government. *see* Laws and regulations; Taxation
- Government assistance for displaced workers, 276
- Government licenses, 145–146
- Graf, Steffi, 232
- Graphs. *see also* Slope
 deriving equation of straight line from, 23–24
 equation of straight line, 22–23
 equilibrium price and quantity, 40
 marginal revenue, 151–152
 price elasticity of demand, 78–82
 price elasticity of supply, 109
 profit maximization, 101–103
 rent controls, 42
 supply and demand, 40
 transforming tables into, 26–27
 unregulated housing market, 41
- Great Britain. *see* United Kingdom
- Hair transplantation, 236
- Hanson, Gordon, 269, 269n
- Hardware manufacturers, inclusion of free software by, 15
- Health care
 costs of, 243–248
 economics of, 242
 immunizations and, 57, 75, 106, 242
- Health insurance, 206–207, 243–248
 “Health Insurance and the Demand for Medical Care”
 (Manning et al.), 244n
- Health maintenance organizations (HMOs), 245–246
- Herrnstein, Richard, 203n
- Heuristics, judgmental and decision, 196–200
- Hirsch, Fred, 208, 208n
- HMOs (health maintenance organizations), 245–246
- Holmes, Sherlock (fictional character), 7
- Home appliances, 75, 164
- Homo economicus, 195, 196
- Honda Civic, 70
- Honesty, 188, 189
- Horizontal addition, 71–72
- Horizontal interpretation of demand curve, 36
- Horizontal interpretation of supply curve, 37
- Hotelling, Harold, 185, 185n
- Household Food Consumption Survey, 255
- Housing market
 buying decisions of wealthy individuals, 68–69
 centralized allocation and control in, 33
 discrimination in, 42
 finder’s fees and key deposits in, 42
 new houses and carpenter wage rates, 50
 rental housing, 32–34, 41–43
 unregulated, 41
 zoning laws and, 224, 230
- How We Know What Isn’t So: The Fallibility of Human Reason in Everyday Life* (Gilovich), 199n
- Hurdle method of price discrimination, 160–164
- Hybrid sports cars, 180–182
- Hyperbolic discounting, 203n
- Ice cream, waiting lines for, 65, 71
- Illicit drugs, 67, 73
- Immunizations, 57, 75, 106, 242
- “The Impact of the 2018 Trade War on U.S. Prices and Welfare” (Amiti et al.), 275n
- Imperfectly competitive firms
 defined, 96–97, 141
 demand curve for, 144
 forms of, 142–144
 monopolistic competition, 142–143, 184–186
 monopoly, 142–143, 146, 155–156
 oligopoly, 143, 175–177
 perfectly competitive firms compared to, 144
 as price setters, 96–97, 141
 profit maximization for, 150–155
- Imperfect price discrimination, 160
- Implicit costs, 4, 7–8, 12, 120, 123
- Imports
 free trade and, 268–269
 licenses for, 272, 274, 275
 quotas on, 272–275
 tariffs on, 270–272, 274, 275

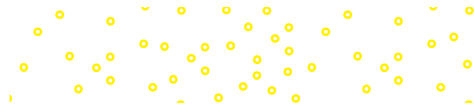
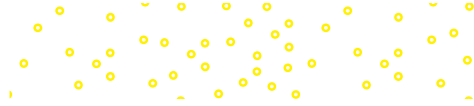
- Impulse-control problems, 202–204
- Incentives
- for behavior change, 124–125
 - of buyers and sellers, 37–40, 56
 - environmental regulation and, 248–253
 - game theory and, 172–173, 182, 186–190
 - income transfers and, 137
 - for market entry, 143
 - price, 248–253
 - psychological, 188–190
 - substitutes and, 75
 - vaccinations and, 106
 - workplace safety and, 210
- Income. *see also* Wages
- demand curve and, 47, 48
 - individual, 227
 - subsidies to, 43
- Income effect, 35, 70–71
- Income elasticity of demand, 86–87
- Income redistribution, 253–257
- Income transfers, 137
- Increasing opportunity cost, 49
- Increasing returns to scale, 146
- Independent variables, 21–22
- Individual demand curve, 71–72
- Individual income, 227
- Individual supply curve, 94–95
- Inefficiency, 155, 219, 276. *see also* Efficiency
- Inelastic demand, 74, 81
- Inferior goods, 48
- Ingenuity of sellers, 48
- In-kind transfers, 254
- Instagram, 146
- Insurance, health, 206–207, 243–248
- Intel microprocessors, 150
- Intergovernmental Panel on Climate Change (IPCC), 252
- Internal Revenue Service, 255
- International property rights, 231, 232
- International trade, 261–277. *see also* Imports
- barriers to, 270–276
 - during Civil War, 261
 - comparative advantage as basis for, 262–266
 - consumption possibilities and, 268
 - environmental issues and, 276
 - exports and, 269
 - fast track authority and, 276–277
 - free trade, 268–269, 276
 - in open vs. closed economy, 265–269
 - protectionism in, 270–277
 - quotas and, 270, 272–275
 - red tape barriers and, 275
 - restrictions on, 270–276
 - supply and demand in, 266–270
 - tariffs and, 270–275
 - trade shocks and, 269
 - Trans-Pacific Partnership and, 261–262
 - winners and losers from, 268–269
 - world prices and, 267–269
- Intertemporal choice, 202–203
- Interviews, first impressions at, 67
- “The Intimate Contest for Self-Command” (Schelling), 203n
- Intuit Corp., 15
- Investments, 211, 228
- Invisible hand theory, 123–128
- cost-saving innovations and, 127–128
 - defined, 123
 - externalities and, 217–218, 228
 - free entry and exit in, 126–127
 - freeway speed and, 127
 - functions of price in, 123
 - monopoly and, 155–156
 - profit and loss in, 123–126
 - recap/summary, 128
 - Smith on, 119, 123, 217
 - supermarket checkout lines and, 127
- IPCC (Intergovernmental Panel on Climate Change), 252
- Iron law of politics, 206
- Ithaca (New York), ethnic diversity in, 119
- Jack-of-all-trades, 33
- James, LeBron, 67
- Janoff-Bulman, Ronnie, 206n
- Jevons, Stanley, 34
- Job interviews, first impressions at, 67
- Johnson, Dwayne, 115
- Judgmental and decision heuristics, 196–200
- “Judgment under Uncertainty: Heuristics and Biases” (Tversky & Kahneman), 196n, 199n
- Kahneman, Daniel, 6n, 196, 196–197n, 199n, 205n
- Keeler, E. B., 244n
- Kennedy, Patrick J., 275n
- Key deposits, 42
- Khandelwal, Amit K., 275n
- Kindergarten, starting dates for, 234–235
- Knetsch, Jack L., 205n
- Labor market. *see also* Employment; Wages
- displaced workers and, 276
 - free trade and, 268, 269
- Land sales in Manhattan, 110
- Law of demand, 66
- Law of diminishing marginal utility, 225
- Law of diminishing returns, 98, 101, 103
- Law of supply, 103–104
- Laws and regulations
- entrapment laws, 204
 - environmental, 248–253
 - externalities and, 223–225
 - noise ordinances, 224
 - traffic laws, 224
 - workplace safety, 208–211
 - zoning laws, 224, 230
- Legal issues. *see* Laws and regulations
- Lemonade, elasticity of supply of, 110–111
- Length of workweek, 209–210

- Lewis, Michael, 205n
 Liebowitz, A., 244n
 Lincoln, Abraham, 261
 Littering, 104–106
 Living standards, 33
 Location, monopolistic competition and, 184–186
 Loewenstein, George, 203n
 London Philharmonic Orchestra, 91
 Long run
 cost-saving innovation and economic profit in, 128
 defined, 97
 law of diminishing returns and, 103
 price elasticity of demand in, 75
 price elasticity of supply in, 112, 115
 zero-profit tendency and, 126
 Loss aversion, 205–207
 Losses
 economic, 122, 143, 154–155
 efficiency, 164
 invisible hand theory on, 123–126
 Lost-\$10-ticket vs. lost-\$10-cash problem, 196
 “Lottery Winners and Accident Victims:
 Is Happiness Relative?” (Brinkman
 et al.), 206n
 Luxury tax on yachts, 77–78
 Lying, 189

 Macroeconomics, defined, 13
 Madrian, Brigitte, 206, 206n
 Magazine advertisements, 236
 Magic cards, 141, 142
 Major League Baseball (MLB), 234
 Manhattan. *see* New York City
 Manning, W. G., 244–245, 244n
 Marginal benefits, 9–12, 56, 57, 105
 Marginal cost curve, 101–104
 Marginal costs
 Cost-Benefit Principle and, 9–12
 defined, 9, 101
 law of diminishing returns and, 101
 reduction of, 50
 socially optimal quantity and, 56
 vaccinations and, 57
 Marginal revenue, 150–155
 Marginal revenue curve, 151–152
 Marginal utility, 225
 Market demand curve, 71–72, 82, 104
 Market equilibrium
 efficiency and, 132–135
 size of economic pie and, 137
 supply and demand and, 38–44
 unexploited opportunities and, 130
 Market equilibrium price, 132–134
 Market power, 145–147
 Market prices, 34–35
 Markets. *see also* Housing market; Labor market
 for allocation of resources, 33, 34
 buyers and sellers in, 34–37
 defined, 34
 entry into, 125–127
 exit from, 126–127
 perfectly competitive, 95–107
 Market supply curve, 94–95, 104, 125–126
 Marquis, M. S., 244n
 Marshall, Alfred, 35, 45
 Marvel Studio films and action figures, 48
 Marx, Karl, 34
 Means-tested benefit programs, 254–255
 Medicaid, 243, 245, 254
 Medical insurance, 206–207, 243–248
 Medicare, 243, 245
 Mediterranean Sea, pollution in, 231
 “Mental Accounting and Consumer Choice”
 (Thaler), 196n
 Microeconomics, 13
 Microsoft, 96–97, 126, 146
 Military-arms-control agreements, 186
 Military arms races, 208–209
 Milkshakes, shared, 231
 Minimum wage, 255–257
 Mixed economies, 33
 MLB (Major League Baseball), 234
 Mobsters, 186
 Models, economic, 5
 Monet, Claude, 34
 Monopolistic competition, 142–143, 184–186
 Monopoly, 142–143, 146, 155–156
 Moral sentiments, 188
 Motor vehicles. *see* Automobile industry
 Movie industry. *see* Entertainment industry
 Multiplayer prisoner’s dilemmas, 175
 Municipal noise ordinances, 224

 Nash, John, 172, 172n
 Nash equilibrium, 172–173
 National Basketball Association (NBA), 114–115, 234
 National Football League (NFL), 232–234
 Natural monopoly, 146
 Natural selection, 14
 Naval blockades, 261
 NBA (National Basketball Association),
 114–115, 234
 Needs vs. wants, 67
 Negative externalities, 217–219, 225–226
 Negative income tax (NIT), 255–257
 Nerd norms, 235
 Network economies, 146
 New entrants into market, 125–127
 Newhouse, J. P., 244n
 New York City
 centralized allocation and control in, 33
 finder’s fees and key deposits in, 42
 food distribution system in, 31–32
 housing prices in, 68–69
 land sales in Manhattan, 110
 rental housing market in, 32–34, 41–42
 Times Square neighborhood of, 224
 zoning laws in, 224

- NFL (National Football League), 232–234
 9/11 terrorist attacks (2001), 114
 NIT (negative income tax), 255–257
 Noise ordinances, 224
 Nominal price, 69
 Nonpositional goods, 208, 209, 211
 No-passing zones, 224
 Normal goods, 48
 Normal profit, 121, 123–125
 Normative economic principle, 13
 North Korea, communist economy in, 33
 Nude photographs, 236
- Obama, Barack, 246, 247, 277
 Obamacare (Affordable Care Act), 206–207, 247–248
 Occupational Health and Safety Administration, 209
 O'Donoghue, Ted, 203n
 Oil prices, 114
 Oligopoly, 143, 175–177
Omerta (underworld code), 186
 One World Trade Center, 145, 146
 OPEC (Organization of the Petroleum Exporting Countries), 114, 177
 Open economy, 265–269
 Operating systems for computers, 96, 126, 146
 Opportunity costs
 comparative advantage and, 263–265
 Cost-Benefit Principle and, 4–5
 defined, 4
 increasing, 49
 profit maximization and, 158–159
 of pursuing any given occupation, 91
 supply and, 54, 92–93
 term-paper revisions and, 51
 Organization of the Petroleum Exporting Countries (OPEC), 114, 177
 Output
 price discrimination and, 157–160
 profit-maximizing level of, 98–100, 152–153
 Output–employment relationship, 97–100
- Pandemic, COVID-19 (2020), 57, 275
 Parable of Buridan's ass, 200
 Parameters, defined, 21
 Pareto, Vilfredo, 132
 Pareto efficiency, 132. *see also* Efficiency
 Paris Agreement (2015), 252
 Parties, shouting at, 179–180
 Patents, 145, 146
 Payoff matrix, 170–174, 177, 233
 Payroll-deduction savings plans, 203, 206
 Peer influences on demand, 67
 Pepsi, 186
 Perception, psychophysics of, 200
 Perfect hurdles, 161–162
 Perfectly competitive firms
 cost concepts for, 98–103
 demand curve for, 96–97, 144
 imperfectly competitive firms compared to, 144
 as price takers, 95, 127, 141
 production in short run for, 97–98
 profit-maximizing output level for, 98–100
 shutdown condition and, 101
 supply curve for, 104
 Perfectly competitive markets, 95–107
 Perfectly elastic demand curve, 81, 82
 Perfectly elastic supply curve, 110–111
 Perfectly inelastic demand curve, 81–82
 Perfectly inelastic supply curve, 110
 Perfect price discrimination, 160
 Perfume advertisements, 236
 Permits for pollution, 250–252
 Personal Responsibility and Work Opportunity Reconciliation Act of 1996, 254
 Pharmaceutical industry, 145
 Philip Morris, 178–179
 Pigou, A. C., 226
 Pizza
 demand curve for, 35–36
 elasticity of demand for, 74
 equilibrium price and quantity of, 38–41
 market for, 34
 price controls on, 43–44
 supply curve for, 36–37
 Planting of trees, subsidizing of, 226
 Pollution
 Coase theorem and, 219–221
 multinational, 231
 permits for, 250–252
 regulation of, 224
 socially optimal level of, 56, 105, 225–226
 taxation of, 248–250
 Poor decisions. *see* Decision pitfalls
 Positional arms control agreements, 234–236
 Positional arms races, 208, 234
 Positional externalities, 232–236
 arbitration agreements, 234
 arms races and arms control agreements, 234–236
 campaign spending limits, 234
 defined, 234
 football players and steroid use, 232–233
 kindergarten starting dates, 234–235
 payoffs dependent on relative performance, 232–233
 roster limits, 234
 social norms, 235–236
 Positional goods, 208, 209
 Positive economic principle, 13
 Positive externalities, 217–219, 226
 Poverty, 255
 Poverty threshold, 255
 Power, market, 145–147
 “The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior” (Madrian & Shea), 206n
 Praise, 198
 Preferences. *see* Wants
 Presidential elections, 234
 Price ceilings, 43

- 
- Price changes, 35, 108
- Price controls
- gasoline and, 42
 - on home heating oil, 135
 - market equilibrium and, 41–44
 - pizza and, 43–44
 - price subsidies, 136–137
 - rent and, 33, 41–43
- Price discrimination, 157–165
- deep discounts for scratched/damaged goods, 164
 - defined, 157
 - examples of, 163–164
 - hurdle method of, 160–164
 - imperfect, 160
 - output and, 157–160
 - perfect, 160
 - public interest and, 163
 - recap/summary, 165
- Price elasticity of demand, 73–87
- applications of, 77–78
 - big-ticket items and, 75
 - budget share and, 75
 - calculation of, 74, 82
 - Cost-Benefit Principle and, 75
 - cross-price elasticity of, 86–87
 - defined, 73–74
 - determinants of, 75–76
 - elastic and inelastic, 74, 81
 - estimates for selected products, 76
 - graphical interpretation of, 78–82
 - in long run vs. short run, 75
 - perfectly elastic/inelastic, 81–82
 - slope in relation to, 80
 - straight-line demand curve and, 80–81, 85
 - substitution possibilities and, 75
 - time considerations and, 75
 - total expenditure and, 82–86
 - unit elastic, 74, 81
- Price elasticity of supply, 108–115
- defined, 108
 - determinants of, 111–114
 - flexibility of inputs and, 111
 - graphical interpretation of, 109
 - in long run vs. short run, 112
 - mobility of inputs and, 111–112
 - perfectly elastic/inelastic, 110–111
 - substitute inputs and, 112
 - time considerations and, 112
 - unique and essential inputs, 114–115
- Price incentives, 248–253
- Prices. *see also* Costs; Price controls
- allocative function of, 123, 126
 - changes in, 35, 108
 - equilibrium, 38–41, 45–54
 - input, 50, 51, 107
 - market, 34–35
 - market equilibrium, 132–134
 - nominal, 69
 - oil, 114
 - profit maximization and, 99
 - rationing function of, 123
 - real, 69–70
 - reservation, 36, 37, 55, 66
 - seasonal movement of, 54
 - volatility of, 112–114
 - world, 267–269
- Price setters, 96–97, 141. *see also* Imperfectly competitive firms
- Price subsidies, 136–137. *see also* Subsidies
- Price takers, 95, 127, 141. *see also* Perfectly competitive firms
- “Priming the pump” strategy, 15
- Print vs. digital advertisements, 46
- Prisoner’s dilemma, 174–180
- cartel game, 175–177
 - character judgments and, 189
 - cigarette advertising on television and, 178–179
 - defined, 174
 - football players and steroid use, 233
 - multiplayer, 175
 - original scenario for, 174
 - recap/summary, 180
 - repeated, 177–178
 - shouting at parties and, 179–180
 - tit-for-tat strategy in, 177–178
- Private goods, 211
- Private ownership, 229–232
- Private universities, 1, 3
- Producer surplus, 163
- Product clustering, 185–186
- Product compatibility, 126
- Production, 97–98, 101–104
- Productivity, 91, 263–265
- Product space, 185–186
- Profit. *see also* Economic profit
- accounting, 120–123
 - defined, 95
 - invisible hand theory on, 123–126
 - normal, 121, 123–125
- Profit maximization
- graphical approach to, 101–103
 - for imperfectly competitive firms, 150–155
 - opportunity cost and, 158–159
 - output level for, 98–100, 152–153
 - in perfectly competitive markets, 95–107
 - seller’s goal for, 160
- Profit-maximizing firms, defined, 95
- Property rights
- international, 231, 232
 - private ownership and, 229–232
 - tragedy of the commons and, 227–232
 - unpriced resources and, 227–229
- Proportions, costs and benefits as, 6–7
- Protectionism, 270–277. *see also* International trade
- Psychological incentives, 188–190
- Psychophysics of perception, 200
- 

- Public goods, 211
- Public investments, 211
- Public service employment, 43, 256, 257
- Publishing industry, 126, 164
- Pure monopoly, 142, 143
- Quantity
- change in quantity demanded, 44–45
 - change in quantity supplied, 45
 - equilibrium, 38–41, 45–53
 - socially optimal, 56
- Quicken software, 15
- Quotas, 270, 272–275
- R&D (research and development), 149, 226, 265
- Rational choice model
- departures from without regret, 211
 - departures from with regret, 201, 204, 211
 - fungibility of wealth and, 196
 - intertemporal choice and, 202–203
 - lost-\$10-ticket vs. lost-\$10-cash problem, 196
 - rent-convenience trade-off and, 201
- Rational person, defined, 3
- Rationing function of price, 123
- Real price, 69–70
- Real-world examples. *see* Economic naturalism
- Rebates, 157, 161–163
- Reconstructive surgery, 236
- Recycling services
- aluminum containers vs. glass containers, 104
 - container-recycling services, 92–93
 - individual supply curve for, 94
 - market supply curve for, 94–95
 - socially optimal level of recycling glass containers, 105
- Red Cross blood services, 95
- Redding, Stephen J., 275n
- Red tape barriers, 275
- Regression to the mean, 198–199
- Regulations. *see* Laws and regulations
- Relative position, 207–214
- Remote-office game, 183, 186, 188
- Renoir, Pierre-Auguste, 34, 35
- Rental housing market, 32–34, 41–43
- Rent controls, 33, 41–43
- Rent-convenience trade-off, 201
- “The Rentocracy: At the Intersection of Supply and Demand” (Tierney), 32n
- Repeated prisoner’s dilemma, 177–178
- Representative heuristic, 197
- Research and development (R&D), 149, 226, 265
- Reservation price, 36, 37, 55, 66
- Resource allocation, 33, 34, 54, 218–219
- Restaurants
- all-you-can-eat, 9
 - commitment problem and, 186
 - tipping at, 186, 187, 189
- “The Return to Protectionism” (Fajgelbaum et al.), 275n
- Revenue, 82–86, 150–155
- Right-of-way rules, 224
- Rise, in slope of lines, 23
- Robinson, Joan, 143n
- Roster limits, 234
- Run, in slope of lines, 23
- Sales, temporary, 163
- Salt, price elasticity of demand and, 75
- Satisficers, 195
- Saving
- impulse-control problems and, 203, 204
 - middle-income families and, 210
 - payroll-deduction plans, 203, 206
- Scale economies, 146–150
- Scarcity, 2, 3, 241, 263
- Schelling, Thomas, 203n
- School lunch subsidies, 254
- Schools, colleges and universities, 1–3
- Sears “Scratch ‘n’ Dent Sale,” 164
- Seles, Monica, 232
- Self-control, 203
- Self-interest, 119, 131, 189, 217
- Selfishness, 189
- Seller’s reservation price, 37, 55
- Seller’s surplus, 55
- Sense of justice, 189
- September 11 terrorist attacks (2001), 114
- Sequential games, 184
- Services industry, 91
- Shared living expenses, 221–222
- Shared milkshakes, 231
- Shea, Dennis, 206, 206n
- Shefrin, H., 203n
- Shifts. *see also* Changes
- in demand curve, 45–48, 51–53
 - in supply curve, 48–53, 107–108
- Shortages
- defined, 39
 - as excess demand, 39
 - of gasoline, 42, 69
 - of toilet paper, 57
- Short run
- cost-saving innovation and economic profit in, 128
 - defined, 97
 - marginal cost curve in, 103
 - price elasticity of demand in, 75
 - price elasticity of supply in, 112
 - production in, 97–98
 - shutdown condition in, 101
- Shouting at parties, 179–180
- Shutdown condition, 101
- Silver Blaze* (Doyle), 7
- Simon, Herbert, 195
- Simonson, Itamar, 201n
- Simultaneous equations, 28–29, 64
- Skateboard market and price of fiberglass, 49
- Slice of economic pie. *see* Economic pie

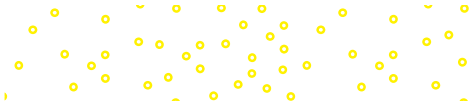
- Slope
 changes in, 25–26
 defined, 23
 of demand curve, 35–36, 144
 of marginal cost curve, 103
 price elasticity of demand and, 80
 rise and run of, 23
 of supply curve, 36–37, 94
- “Smart for one but dumb for all” behaviors, 56, 57, 131
- Smith, Adam. *see also* Invisible hand theory
 on invisible hand, 119, 123, 217
 on market price of goods, 34
 on self-interest, 119, 131, 217
The Wealth of Nations by, 34, 119, 131n
- Smoking, 77, 143, 178–179
- SNAP (Supplemental Nutrition Assistance Program), 254
- Snow tires, 224
- Social efficiency, 159, 220–221
- Social justice and efficiency, 135
- Social Limits to Growth* (Hirsch), 208n
- Socially optimal quantity, 56
- Social norms, 235–236
- Social optimum
 equilibrium vs., 130–135
 health care expenditures and, 242
 pollution levels and, 105, 225–226
- Social Security, 210
- Software industry. *see also* Computers
 economies of scale in, 147
 hardware manufacturers and free software, 15
 microprocessors, 150
 operating systems, 96, 126, 146
 product compatibility in, 126
 start-up costs in, 147
- SO₂ (sulfur dioxide) emissions, 251–252
- Soviet Union, centrally planned economy in, 33
- SpaceX, 10–11
- Spatial location, 184
- Specialization, 33, 263–265
- “Specious Reward: A Behavioral Theory of Impulsiveness and Impulse Control” (Ainslie), 203n
- Speed limits, 224
- Splitting economic surplus, 223
- Sports
 baseball, 234
 basketball, 114–115, 234
 football, 232–234
 talented players in, 114
 tennis, 45–46, 232
- Sports Illustrated* jinx, 198–199
- Sport utility vehicles (SUVs), 70
- Spreadsheets, 146
- “Stability and Competition” (Hotelling), 185n
- Standard of living, 33
- Start-up costs, 147–150
- State universities, 1
- Status quo bias, 206, 207, 253
- Steroids, 232–233
- Stock analysts, 131–132
- Straight-line demand curve, 80–81, 85
- Strategic behavior. *see* Game theory
- Student discounts, 157
- “Subjective Probability: A Judgment of Representativeness” (Kahneman & Tversky), 197n
- Subsidies
 externalities and, 226
 gasoline, 136–137
 to income, 43
 for planting of trees, 226
 price-related, 136–137
 research and development, 226
 school lunches and, 254
- Substitutes, 46–47, 75, 86, 112, 146
- Substitution effect, 35, 68–70
- Sulfur dioxide (SO₂) emissions, 251–252
- Sunk costs, 9, 12
- Supermarket checkout lines, invisible hand at, 127
- Supplemental Nutrition Assistance Program (SNAP), 254
- Supply. *see also* Price elasticity of supply; Supply and demand;
 Supply curve
 change in, 45
 competitive, 104
 excess, 39, 40, 133
 graphing, 40
 law of, 103–104
 opportunity cost and, 54, 92–93
 seasonal price movements and, 54
 ultimate bottleneck in, 114–115
- Supply and demand, 31–58. *see also* Demand; Supply
 algebraic framework for, 63–64
 in centrally planned economies, 33
 efficiency and equilibrium in, 54–58
 in food distribution systems, 31–32
 graphing, 40
 income effect and, 35
 inferior goods and, 48
 international trade and, 266–270
 in market-based economies, 33–34
 market equilibrium and, 38–44
 normal goods and, 48
 price and quantity changes related to, 44–54
 rent controls and, 33, 41–43
 substitution effect and, 35
- Supply curve
 change in quantity supplied and, 45
 change in supply and, 45
 defined, 36, 91
 equations for, 63
 future expectations and, 51, 108
 horizontal interpretation of, 37
 individual, 94–95
 input prices and, 50, 51, 107
 as marginal cost curve, 104
 market, 94–95, 104, 125–126
 market entry/exit and, 125–126
 number of suppliers and, 107–108
 for perfectly competitive firms, 104
 perfectly elastic and inelastic, 110–111

- price changes of other products, 108
- shifts in, 48–53, 107–108
- slope of, 36–37, 94
- technology changes and, 50–51, 107
- vertical interpretation of, 37, 133
- weather conditions and, 51
- Supply elasticity. *see* Price elasticity of supply
- Supportive management styles, 198
- Surplus
 - buyer's, 55
 - consumer, 160, 163
 - defined, 39
 - economic, 4
 - excess supply as, 39
 - producer, 163
 - seller's, 55
 - total, 55, 56
- Surplus-enhancing transactions, 133, 134
- SUVs (sport utility vehicles), 70
- Sympathy, 189
- Systematic cognitive errors, 201, 204

- Tables, constructing equations and graphs
 - from, 26–27
- Talent, 114, 129–130
- TANF (Temporary Assistance for Needy Families), 254
- Tariffs, 270–275
- Taste, norms of, 235–236
- Tastes. *see* Wants
- Tattoos, 235
- Taxation
 - of carbon emissions, 252–253
 - on cigarettes, 77
 - compensatory, 226
 - earned-income tax credit, 137, 256
 - externalities and, 226
 - on gasoline, 70, 241
 - luxury tax on yachts, 77–78
 - negative income tax, 255–257
 - of pollution, 248–250
 - voting behavior and, 211–213
- Technology, supply curve and changes in, 50–51, 107
- Teenage smoking, 77
- Telephone companies, 143
- Temporal location, 184
- Temporary Assistance for Needy Families (TANF), 254
- Temporary sales, 163
- Tennis, 45–46, 232
- “Ten of the Greatest Optical Illusions” (Wiseman), 212n
- Term-paper revisions, 51
- Thaler, Richard, 9n, 196, 196n, 200n, 203n, 205n
- The Theory of Monopolistic Competition* (Chamberlin), 143n
- Theory of the invisible hand. *see* Invisible hand theory
- Thinking Fast and Slow* (Kahneman), 6n
- Thinking-like-an-economist mindset, 127
- Third-party payment system, 243–244
- Tierney, John, 32n
- Timber harvesting on remote public land, 231
- Time. *see also* Long run; Short run
 - price elasticity of demand and, 75
 - price elasticity of supply and, 112
- “Time Discounting and Time Preference: A Critical Review” (Frederick et al.), 203n
- Timing, in game theory, 180–186
- Tipping at restaurants, 186, 187, 189
- Titanic* disaster, 66–67
- Tit-for-tat strategy, 177–178
- Tjoa, Bill, 16
- T-Mobile, 143
- Tobacco industry, 77, 143, 178–179
- Toilet paper shortages, 57
- Total cost curve, 147
- Total costs, 99, 101, 147–148
- Total expenditure (revenue), 82–86
- Total surplus, 55, 56
- “Toward a Positive Theory of Consumer Choice” (Thaler), 9n, 200n
- Toyota Prius, 70
- TPP (Trans-Pacific Partnership), 261–262
- Trade. *see* International trade
- Trade-offs, 1–4, 13
- Trade restrictions, 270–276
- Trade shocks, 269
- Trade wars, 275
- Traffic laws, 224
- Tragedy of the commons, 229–232
- Transfer payments, 254
- Trans-Pacific Partnership (TPP), 261–262
- Trial and error, in decision making, 5
- Trump, Donald, 277
- Trustworthiness, 189
- TurboTax* software, 15
- Tversky, Amos, 6n, 196, 196–197n, 199n, 201n

- Underworld code (*Omerta*), 186
- Unequal rent amounts, 223
- Unique and essential inputs, 114–115
- United Airlines, 170–173
- United Kingdom, 70, 206
- United States
 - comparative advantage for, 265
 - fast track authority in, 277
 - health insurance in, 206–207, 247–248
 - trade war with China, 275
- Unit elastic demand, 74, 81
- Universities, 1–3
- Unpriced resources, 227–229
- Unregulated housing market, 41
- Utility, marginal, 225

- Vacation vs. business travelers, 164
- Vaccinations. *see* Immunizations
- Vanity, norms against, 236
- Variable costs, 101, 147
- Variable factor of production, 98
- Variables, in equations, 21–22
- Vehicles. *see* Automobile industry

- 
- Verbal descriptions, construction of equations
from, 21–22
- Verizon, 143
- Vertical intercept, 23–25
- Vertical interpretation of demand curve, 36, 133
- Vertical interpretation of supply curve, 37, 133
- Video game producers, 148–149
- Voting behavior, 211–213
- Wages. *see also* Income
minimum, 255–257
profit maximization and, 99–100
- Waiting lines, 65, 71
- Wants (preferences)
commitment problem and, 189
Cost-Benefit Principle and, 48
as determinant of reservation price, 66
as driver of demand, 66–67
needs vs., 67
- Warner Brothers, 169, 180, 182
- Wealth, fungibility of, 196
- The Wealth of Nations* (Smith), 34, 119, 131n
- Wearstler, Kelly, 262, 263
- Weather conditions, 51
- Weber-Fechner law, 200
- Wedding celebrations, 211
- Weight-loss camps, 203
- Weinstein, David, 275n
- Whales harvesting in international waters, 231
- Wheat, 96
- Windows operating system, 96, 126, 146
- Winston, Gordon, 203n
- Wireless phone service, 143
- Wiseman, Richard, 212, 212n
- Word processors, 146
- Work and workers. *see* Employment; Labor market
- Workplace safety regulation, 208–211
- Workweek, length of, 209–210
- World prices, 267–269
- World Trade Organization, 253
- World War II (1939–1945), 213
- Yachts, luxury tax on, 77–78
- Yosemite Hospitality, LLC, 145–146
- Zero-profit tendency, 126
- Zimmerman, Dennis, 77n
- Zoning laws, 224, 230
- 