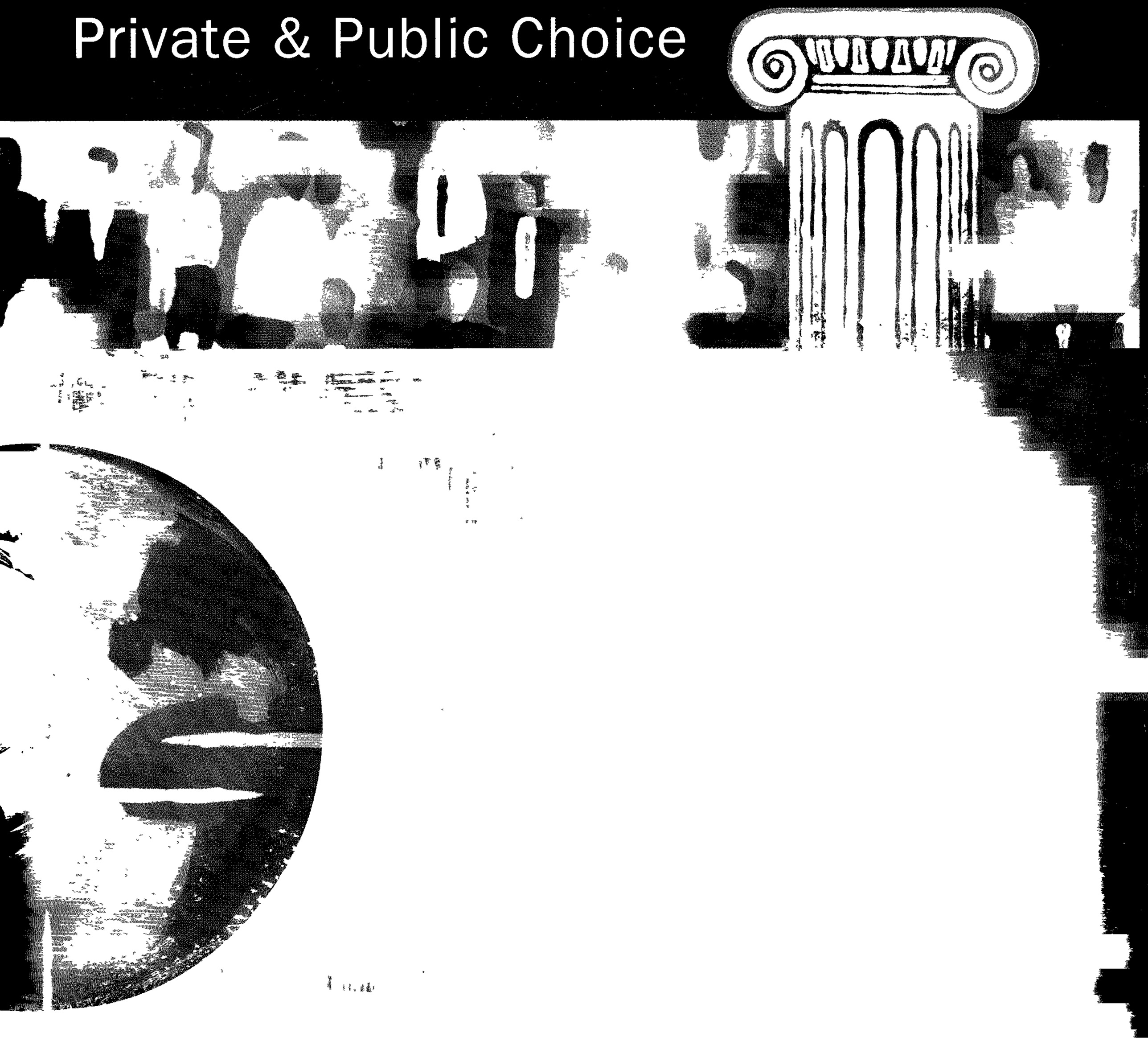
11th edition

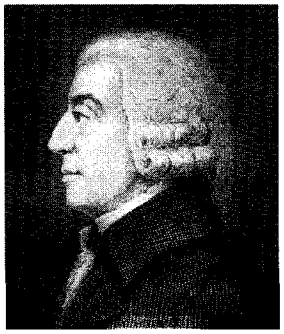
MICROECONOMICS

Private & Public Choice



The Evolution of Economics as a Science

1776 ADAM SMITH (1723–1790)



© HULTON- DEUTSCH COLLECTION / CORBIS

Smith's book *An Inquiry into the Nature and Causes of the Wealth of Nations* provided the first comprehensive analysis of wealth and prosperity and introduced "the invisible hand" principle. It also explained that the wealth of a nation was determined by its production of goods and services, not by its gold and silver.

1817 DAVID RICARDO (1772–1823)

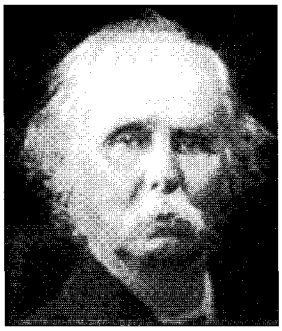
© BETTMANN / CORBIS

In his book *On the Principles of Political Economy and Taxation,* Ricardo developed the law of comparative advantage and used it to explain why trade leads to mutual gains.

1871 WILLIAM STANLEY JEVONS (1835–1882)

Along with Carl Menger and Leon Walras, Jevons (in *The Theory of Political Economy*) introduced (1) the idea that the value of goods is determined subjectively rather than by the labor required for production, and (2) the law of diminishing marginal utility. Independently, the same concepts were developed by Menger in *Grundsätze* (1871) and Walras in *Elements of Pure Economics* (1874). These two concepts are still an integral part of modern analysis.

1890 ALFRED MARSHALL (1842–1924)



In his book *The Principles of Economics*, Marshall introduced and developed many of the key concepts of modern microeconomics, including concepts like supply and demand, equilibrium, short run and long run, elasticity, and consumer and producer surplus. The book went through eight editions between 1890 and 1920.

1936 JOHN MAYNARD KEYNES (1883–1946)

© BETTMANN/ CORBIS

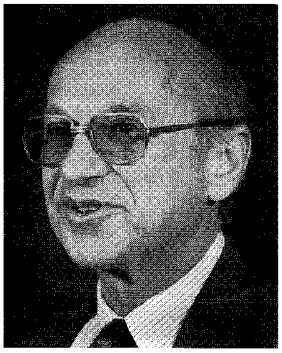
In his book *The General Theory of Employment, Interest, and Money*, Keynes developed the framework for modern macroeconomics. He also developed an explanation for the widespread unemployment of the Great Depression, and he elevated the importance of fiscal policy.

1940s FRIEDRICH VON HAYEK (1899–1992)

© HULTON- DEUTSCH COLLECTION/ CORBIS

In two vitally important publications, *The Road to Serfdom* (1944) and "The Use of Knowledge in Society," an article in the *American Economic Review* in 1945, Hayek explained the role of knowledge in economics, enhanced our understanding of the market process, and highlighted the fatal defects of centrally planned economies.

1960s and 1970s MILTON FRIEDMAN (1912–)



© BETTMANN / CORBIS

Friedman's work elevated the importance of monetary policy and convinced many that monetary instability was the major cause of both business fluctuations and inflation. His book *A Monetary History of the United States, 1867–1960* (with Anna Schwartz) was a particularly important publication.

1970s and 1980s ROBERT LUCAS (1937–)

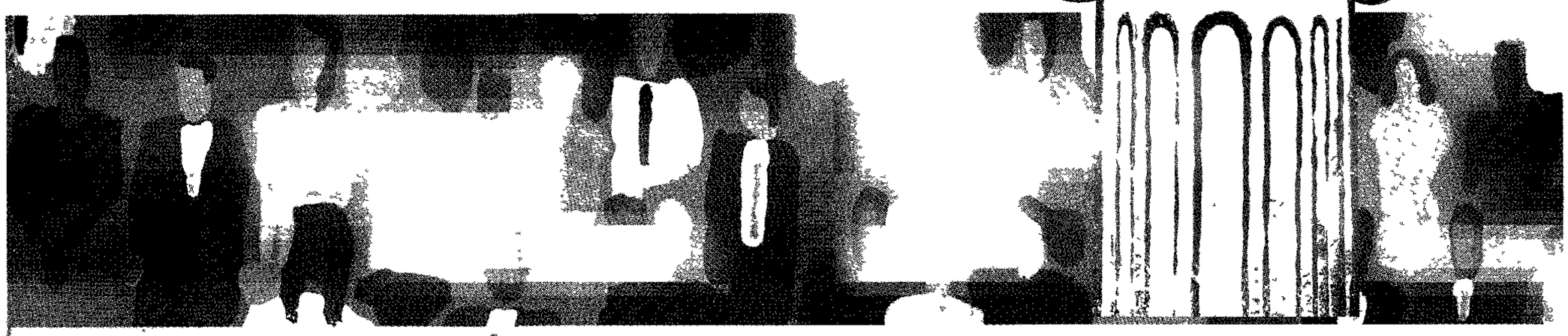
© RALF-FINN HESTOFT / CORBIS

The role people's expectations play in the macroeconomy dramatically altered prior economics analysis. Although several economists made major contributions in this area, Lucas is generally recognized as the leading contributor.

11th edition

MICROECONOMICS

Private & Public Choice



JAMES D. GWARTNEY
Florida State University

•

RICHARD L. STROUP
Montana State University

•

RUSSELL S. SOBEL
West Virginia University

•

DAVID A. MACPHERSON
Florida State University

THOMSON
SOUTH WESTERN

Austra a Canada Mex co S ngapore Spa n Un ted K ngdom Un ted States

THOMSON
SOUTH-WESTERN

Microeconomics: Private and Public Choice, Eleventh Edition

James D. Gwartney, R chard L. Stroup, Russe S. Sobe , Dav d A. Macpherson

VP/Editorial Director:
Jack W. Calhoun

Editor-in-Chief:
Alex von Rosenberg

Developmental Editor:
Amy Ray

Marketing Manager:
John Carey

Production Editor:
Robert Dreas

Manager of Technology, Editorial:
V cky True

Sr. Technology Project Editor:
Peggy Buskey

Web Coordinator:
Karen Schaffer

Manufacturing Coordinator:
Sandee M lewski

Production House:
Pre Press Company, Inc.

Printer:
Cour er Kendallv lle, Inc.
Kenda v e, nd ana

Art Director:
Michae Stratton

Internal and Cover Designer:
Nick Gliebe/Design Matters
Cincinnat , Oh o

Cover and Internal Illustrations:
© Jon Allen/Illustration Works

Photography Manager:
John H

Photo Researcher:
Jan Se de

COPYR GHT © 2006
Thomson South Western, a part of The Thomson Corporat on. Thomson, the Star ogo, and South Western are trade marks used here n under cense.

Pr nted n the Un ted States of Amer ca
1 2 3 4 5 08 07 06 05

Student Ed t on: SBN 0 324 32036 1
Instructor's Ed t on: ISBN 0 324 33570 9

ALL R GHTS RESERVED.

No part of th s work covered by the copyr ght hereon may be reproduced or used n any form or by any means graph c, e ectron c, or mechan ca , n clud ng photocc py ng, record ng, tap ng, Web d str but on or 'nformat on storage and retrieva systems, or in any other manner w thout the wr tten per m ss on of the pub sher.

For perm ss on to use mater a from th s text or proc uct, subm t a request on ne at http://www.tho msonr ghts.com.

L'brary of Congress Contro Number: 2005921042

For more nformat on about our products, contact us at:

Thomson Learn'ng Academ c Resource Center

1 800 423 0563

Thomson Higher Education
5191 Natorp Bou evard
Mason, OH 45040
USA

Asia (including India)
Thomson Learn ng
5 Shenton Way
#01 01 uIC Building
S ngapore 068808

Australia/New Zealand
Thomson Learn ng Austra a
102 Dodds Street
Southbank, Victoria 3006
Austra a

Canada
Thomson Ne son
1120 B rchmount Road
Toronto, Ontario
M1K 5G4
Canada

Latin America
Thomson Learn ng
Seneca, 53
Co on'a Po anco
11560 Mex co
D.F.Mex co

UK/Europe/Middle East/Africa
Thomson Learn ng
H gh Holborn House
50/51 Bedford Row
London WC1R 4LR

United Kingdom
Spa n (nc ud ng Portuga)
Thomson Paran nfo
Calle Magallanes, 25
28015 Madr d, Spa n

Preface

Our main goal in this edition was to make the book both more understandable and exciting to read. Readability has been a past strength, but we worked hard to make this edition more concise and the most student friendly edition ever. Even though we did not eliminate coverage of any topic, this edition still has about 7 percent fewer words than the previous one. In this case, we believe shorter is better.

The aim of this text is to: (a) make economics understandable, (b) illustrate the power and relevance of economics to our daily lives, and (c) explain why both individuals and nations prosper. Throughout, we seek to communicate basic and, in some cases, fairly complicated ideas in a clear and understandable manner. Clarity is a major objective, but simplicity is not substituted for in depth analysis. We do not accept the view that economics must be either difficult or "watered down." Instead, we believe that a clear and understandable writing style reinforced with examples, illustrations, and visual aids can make the subject come alive in a way that will capture the interest of even the beginning student.

THE ORGANIZATION OF TEXT: FLEXIBLE COVERAGE OF PUBLIC CHOICE, THE KEYNESIAN MODEL, AND SPECIAL TOPICS

The organization is designed to provide instructors with maximum flexibility. Those using the full length text for a two semester course can cover either microeconomics or macroeconomics first. Beginning with the ninth edition, the text was divided into core chapters and a concluding special topics section. The twenty seven core chapters cover all of the material taught in most principles courses, and they are presented in the usual manner. Examples and data from the real world are used to reinforce the analysis. In addition, the Applying the Basics section includes fourteen relatively short discussions on high profile topics like the growing U.S. budget deficit, reforming Social Security, and the economics of health care. Each topic is designed to be covered during a single class period. This organization has been quite popular, largely because it makes it easier for instructors to tailor their course to fit their own preferences.

Those teaching a microeconomics course who like to stress the importance of public choice will probably want to cover the first six chapters prior to beginning the core microeconomics section. Others will prefer to cover only the first four chapters and then move immediately to the core microeconomics material. The book is designed for both of these options. Instructors who like to integrate applications extensively into their micro course can pick and choose among special topics covering issues like Social Security, investing in the stock market, healthcare, school choice, income differences between men and women, and property rights and the quality of the environment.

Those teaching a macroeconomics course integrating public choice will probably want to cover Chapters 5 and 6 prior to the core macroeconomics material. Others may want to move directly from Chapter 4 (or Chapter 3) to the core macroeconomics material. The macroeconomics chapters have been written so there will be no problems with either option. Similarly, instructors who want to omit the Keynesian 45 degree aggregate expenditure model (Chapter 11) can do so without having to worry that this will cause problems for students in subsequent chapters. Special Topic 1 provides facts and figures about government spending and taxation for those who highlight this topic. Our own teaching experience indicates that the special topics on budget deficits and the national debt, European unemployment, and the "Irish miracle" can enrich a macroeconomics course.

CHANGES IN THIS EDITION

New in Macroeconomics

Chapter 8. New Application box entitled "Would Personal Savings Accounts Make Unemployment Compensation Work Better?"

Chapter 12. A new section on the politics of fiscal policy and the related box entitled "Have Supply Side Economists Found a Way to Soak the Rich?" were added.

Chapter 15. Material on the Phillips Curve, both the early views and the current expectations view, was integrated into the chapter.

Special Topic 7, "Institutions, Policies, and the Irish Miracle," is new.

New in Microeconomics

Chapter 20. New box entitled, "Cooking the Books: How the Market Responds to Criminal Behavior."

Substantial revisions to Special Topic 4 focusing on how to best invest in the stock market.

New feature entitled, "Can Imports from Canada Reduce the Drug Prices of Americans?" in the special topic relating to health care.

Substantial updates and revisions to the special topic on the environment and natural resource use.

New in Both Microeconomics and Macroeconomics: Economics in the Movies

Economics pervades our culture including our entertainment. Both the macroeconomics and microeconomics editions of the book contain a new boxed feature in the text called "Economics in the Movies." These boxes featured in the text include various scenes from popular movies that reflect economic concepts. A number of instructors, including the authors, now use clips from popular movies to stimulate student interest and drive home the importance of these concepts. The instructor's manual provides more ideas about how this can be done effectively

Additional Text Features

Economics: Private and Public Choice retains several features that make the presentation of economics both more interesting and more understandable:

- **Keys to Economic Prosperity** Students often fail to appreciate the organizational and institutional factors that provide the foundation for economic progress. In order to help remedy this situation, we have incorporated a "Key to Economic Prosperity" feature that highlights how important factors like the gains from trade, secure property rights, competition, and free trade are to economic prosperity. In all, twelve of the most important factors that underlie modern economic prosperity are highlighted at appropriate places throughout the text and are also listed inside the back cover.
- **Applications in Economics** "Applications in Economics" boxes apply economic theory to real world issues and controversies.
- **Measures of Economic Activity** The "Measures of Economic Activity" boxes explain how important economic indicators such as the unemployment rate and the index of leading indicators are assembled and what they mean.
- **Outstanding Economist** Numerous boxes in the text highlight the lives of many economists and the contributions they have made.

- **Myths of Economics** These boxed articles dispel commonly held fallacies of economic reasoning. Because they are tomorrow's leaders, we believe that all students should be aware of common economic misperceptions that tend to hamper a nation's economic progress.
- **Chapter Focus Questions and Closing Key Point Summaries** Each chapter begins with four or five questions that summarize the focus of the chapter. At the end of each chapter, the Key Points section provides the student with a concise statement of the material covered in the chapter (the chapter learning objectives). These two features help students better integrate the material into the broader economic picture.
- **Critical Analysis Questions** Each chapter concludes with a set of discussion questions and problems designed to test the student's ability to analyze economic issues and to apply economic theory to real world events. Appendix B at the end of the text contains suggested answers for approximately half of these questions.

SUPPLEMENTARY MATERIALS

For the Student

Coursebooks

The Coursebooks for this edition were prepared by coauthor Professor Russell Sobel and are now available in not two, but three versions, covering all three courses: Economics, Microeconomics, and Macroeconomics. The Coursebooks are more than study guides. Each includes numerous multiple choice, true/false, and discussion questions to help students self test their knowledge of each chapter. Answers and short explanations for most questions are provided in the back of the Coursebooks. Each chapter also contains problem and project exercises designed to improve the student's knowledge of the mechanics. Like the textbook, the Coursebooks are designed to help students develop the economic way of thinking.

Gwartney Xtra!

This Web site, which comes free with the purchase of a new book, offers a robust set of online multimedia learning tools, including Master the Learning Objectives, The Graphing Workshop, CNN Video Clips, and Xtra! Quizzing. Thirty five questions similar to the multiple choice questions of the Test Bank are available for each chapter. These self testing questions provide students with ample opportunity to practice and obtain feedback prior to examinations. Students with used textbooks can purchase Xtra! by going online to **http://gwartneyxtra.swlearning.com**.

***Microeconomics Alive!* and *Macroeconomics Alive!* CD-ROMs.** These CDs help students learn economics the high tech, high fun way. These study CDs, compatible with both Microsoft Windows and Macintosh operating systems, provide animated lessons, interactive graphing exercises, and real life simulations. Learn economics for the first time or review the concepts in your economics course—either way, these award winning CD ROMs provide you with the perfect tools. Find out more at **http://econalive.swlearning.com**.

InfoTrac

Students shouldn't forget to take advantage of their subscriptions to InfoTrac, which comes free with their textbooks. InfoTrac gives them anytime, anywhere online access to a data base of full text articles from hundreds of scholarly and popular periodicals using fast and easy search tools. InfoTrac is updated daily, with articles dating back as far as four years.

Support Web Site

Valuable resources can be found on the text's Internet support site found at the text's on line support site at **http://gwartney.swlearning.com**. Students will find an interactive study center as well as online quizzing.

For the Instructor

We believe that many of the features incorporated into this text will help you become a better teacher. We have incorporated the Keys to Economic Prosperity series, movie clips, homework assignments, and diverse use of the online quiz questions into our own classes. We have also found that students are quite interested in the issues covered in the Special Topics section of the book. We feel sure that many of these features will help make your class more interesting to students. Of course, the full set of supplements accompanies the text. They include the following:

Test Banks

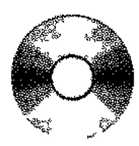
The Test Banks for the eleventh edition were prepared by the author team. The authors have worked hard to update and improve the test banks for this edition. The two Test Banks contain approximately 7,000 questions multiple choice and short answer. Within each chapter, the questions correspond to the major headings of the text. The first ten questions of each chapter are suitable for use as a comprehensive quiz covering the material of the chapter. The multiple choice questions from the Coursebook and the online practice quizzes are also included in special sections of the Test Bank. Instructors who would like to motivate their students to study the Coursebook and online quizzes can easily use these questions and incorporate them into their quizzes and exams.

Computerized Test Banks

The computerized Test Banks for this edition have been enhanced significantly. *Exam View* Computerized Testing Software contains all of the questions in the printed Test Bank. *Examview* is an easy to use test creation software compatible with both Microsoft Windows and Macintosh. Instructors can add or edit questions, instructions, and answers and select questions by previewing them on the screen, selecting them randomly, or selecting them by number. Instructors can also create and administer quizzes online, whether over the Internet, a local area network (LAN), or a wide area network (WAN).

PowerPoint

We believe that our PowerPoint presentation, prepared by Chuck Skipton, is the best you will find in the principles market. It provides chapter by chapter lecture notes with fully animated, hyperlinked slides of the textbook's exhibits. The dynamic slides and accompanying captions make it easy for instructors to present (and students to follow) sequential changes. The dynamic graphics are also used to highlight various relationships among economic variables. In order to facilitate discussion and interaction, questions are strategically interspersed throughout each chapter to help students develop the economic way of thinking. Instructions explaining how professors can easily add, delete, and modify slides in order to tailor make the presentation to their liking are included. If instructors want to make the PowerPoint presentation available to students, they can place it on their Web site (or the site for their course). The slides are available in two sets: Figure Slides and also Lecture Slides. These sets are available for download at the support Web site: **http://gwartney.swlearning.com**.



Instructor's Manual and Instructor's Resource CD

The *Instructor's Manual* was prepared by author David Macpherson. Instructions and information on how to use and modify the PowerPoint material is contained in the front of the manual. Also included at the front of the manual is information about *ExamView,* the computerized testing software that accompanies the book. The remainder of the manual is divided up by chapters, and each chapter is divided into three parts. The first part consists of a detailed chapter outline in lecture note form It is designed to help instructors organize their notes to match the 11th edition of the book. Instructor's can easily prepare detailed, personalized notes by revising the computerized version of the lecture notes on the *Instructor's Resource CD.* The second part of each chapter contains teaching tips, sources of supplementary materials, and other helpful information. Part 3 of each chapter consists of in class economic games and experiments. Contributed in part by Professor Charles Stull of Kalamazoo College, the games are an enormously popular feature with instructors. We hope you will try them

The *Instructor's Resource CD* contains the key supplements designed to aid instructors, including the content from the instructor's manual, test banks, and PowerPoint lecture and exhibit slides for overhead use. The CD also includes an "Integrated Resource Guide," which provides a quick reference to all of South Western's resources for teaching a principles course in economics as well as suggestions for how to incorporate the resources into all the phases of instruction.

Support Web Site for Instructors

This password protected Web site includes some of the same essential resources that can be found on the *Instructor's Resource CD,* including instructor's manuals and test banks in Microsoft Word, and the PowerPoint lecture and exhibit slides. To get access to the site to download these supplements, register online at **http://gwartney.swlearning.com**.

WebTutor Toolbox™ and Web Tutor Advantage™ on Blackboard and WebCT

The *WebTutor Toolbox* uses the Internet to turn everyone in your class into a front row student. It offers interactive study guide features such as quizzes, concept reviews, flashcards, discussion forums, and more. Instructor tools are also provided to facilitate communication between students and faculty. Preloaded with content, *WebTutor ToolBox* pairs all the content of the book's support Web site with all the sophisticated course management functionality of either course management platform.

More than just an interactive study guide, *WebTutor Advantage* delivers innovative learning aids that actively engage students. Benefits include automatic and immediate feedback from quizzes; interactive, multimedia rich explanations of concepts, such as flash animated graphing tutorials and graphing exercises that use an online graph drawing tool; streaming video applications; online exercises; flashcards; and interaction and involvement through online discussion forums. Powerful instructor tools are also provided to facilitate communication and collaboration between students and faculty. More information on WebTutor can be found at **http://webtutor.thomsonlearning.com**. (Other platform choices are available upon request. Please visit the WebTutor Web site for details.)

Principles of Economics Videotape

Principles of Economics, a forty minute videotape giving students an insightful overview of ten common economic principles: Trade offs, Opportunity Cost, Marginal Thinking, Incentives, Trade, Markets, Government's Role, Productivity, Inflation, and the Phillips Curve. The video shows viewers how to apply these principles to their daily lives. It is filled with interviews from some of the leading economists, and includes profiles of real students facing economic choices.

Turner Learning/CNN Economics Video

Professors can bring the real world into the classroom by using the *Turner Learning/CNN Economics Video.* This video, which includes current news stories of economic interest, is produced in cooperation with Turner Learning, Inc. Contact your Thomson Learning sales representative for ordering information.

TEXT-SUPPLEMENT VALUE PACKAGES

Economics: Private and Public Choice can be packaged with a plethora of supplements useful to you and your students. In addition to the Coursebooks, the textbooks can be packaged with the following:

Homework Xpress! Your Homework Management Solution

http://homeworkxpress.swlearning.com

- **InfoTrac® with InfoMarks™**

When you adopt *Economics: Private and Public Choice,* 11th edition, you and your students will gain anytime, anywhere access to reliable resources with InfoTrac College Edition, the

online library. This fully searchable database offers more than twenty years' worth of full text articles (not abstracts) from almost 5,000 diverse sources, such as top academic journals, newsletters, and up to the minute periodicals, including *Time, Newsweek, Science, Forbes,* and *USA Today.*

The Wall Street Journal

Economics: Private and Public Choice, 11th edition, makes it easy for students to apply economic concepts to this authoritative publication, and for you to bring the most up to date, real world events into your classroom. For a nominal additional cost, *Economics: Private and Public Choice,* 11th edition, can be packaged with a card entitling students to a fifteen week subscription to both the print and online versions of the *Wall Street Journal.*

Favorite Ways to Learn Economics

Authors David Anderson of Centre College and Jim Chasey of Homewood Flossmoore High School use experiments to bring economic education to life. This is a growing trend, and for good reason. It works! Favorite Ways to Learn experiments and problem sets reinforce the key principles of microeconomics and macroeconomics covered in most college Advanced Placement courses. Instructors will see an improvement in their students' comprehension. This supplement comes in both student and instructor versions.

How to Think Like an Economist

Most economics instructors believe that a primary goal of this course is to teach students how economists think. There's more to thinking like an economist than knowing the concepts and technical tools of analysis. *How to Think Like an Economist* gives students unique insight into the economist's mind, allowing them to see how interesting issues are approached from an economic perspective. This soft cover guide can be bundled free with new copies of *Economics: Private and Public Choice,* 11th edition.

Economic HITS on the Web with InfoTrac

This resource booklet supports students' research efforts on the World Wide Web. This manual includes an introduction to the Web, material on finding information and documenting Internet sources for research, and numerous economic activities along with the hottest economic URLs.

Your Thomson Learning sales representative can help you with the many text supplement value packages available to you and your students. Completely customizable textbooks are also available via *TextChoice,* Thomson Learning's online digital content. *TextChoice* is the fastest, easiest way for instructors to create their own learning materials. You can select content from hundreds of our best selling titles, choose material from one of our databases or add your own material.

Harvard Business Cases

Thomson is now an official distributor of Harvard Business School Publishing materials, giving you access to the richest and most respected cases and article content at competitive prices. This comprehensive collection includes more than 9,500 Harvard Business School case studies and background notes, as well as selected case items from Babson College, Business Enterprise Trust, Design Management Institute, IESE, IMD, Richard Ivey School of Business, Stanford University, and the University of Hong Kong. For more information, contact your local Thomson South Western Sales Representative.

Gale Business & Company Resource Center

Another expansive resource available is our Gale Business & Company Resource Center (BCRC). Our exclusive and robust online resource center allows students to conduct detailed business research and analysis from their own desks anytime and anywhere they have an Internet connection. Through the BCRC, students gain access to a vast assortment of global business information, including competitive intelligence, career and investment opportunities, business rankings, company histories, and much more. For more information on this valuable resource, visit **http://access.gale.com/thomsonlearning/**.

A NOTE TO INSTRUCTORS

As we try to improve the book from one edition to the next, we rely heavily on our experiences as teachers. But our experience using the book is minuscule compared with that of the hundreds of instructors who use it nationwide. If you encounter problems or have suggestions for improving the book, we urge you to let us know by writing to us in care of Thomson Higher Education, 5191 Natorp Blvd., Mason, OH 45040. Such letters are invaluable, and we are glad to receive both praise and suggestions for improvement. Many such suggestions have found their way into this new book.

A NOTE TO STUDENTS

This textbook contains several features we think will help you "maximize" (a good economic term) the returns from your study efforts. Here are some of the things that will help you and a few tips for making the most of them:

Each chapter begins with a series of focus questions that communicate the central issues of the chapter. Before you read the chapter, briefly think about the focus questions, why they are important, and how they relate to the material in prior chapters.

The textbook is organized in the form of an outline. The headings within the text (highlighted with a color background) are the major points of the outline. Minor headings are subpoints under the major headings. In addition, important subpoints within sections are often set off and numbered. Bold italicized type is used to highlight material that is particularly important. Sometimes "thumbnail sketches" are included to recap material and help the reader keep the important points mentally organized. Careful use of the headings, highlighted material, and the thumbnail sketches will help you master the material.

A "Key Points" summary appears at the end of each chapter. Use the summary as a checklist to determine whether you understand the major points of the chapter.

A review of the exhibits and illustrative pictures will also provide you with a summary of the key points of each chapter. The accompanying captions briefly describe the economic phenomena illustrated by the exhibits.

The key terms introduced in each chapter are defined in the margins. As you study the chapter, go over the marginal definition of each key term as it is introduced. Later, you may also find it useful to review the marginal definitions. If you have forgotten the meaning of a term introduced earlier, consult the glossary at the end of the book.

The boxed features go into more depth on various topics without disrupting the flow of the text. In general, the topics of the boxed features have been chosen because they are a good application of the theory described in the book and because students tend to be interested in them. Reading the boxed features will supplement the text and enhance your understanding of important economic concepts.

The critical analysis questions at the end of each chapter are intended to test your understanding of the economic way of thinking. Solving these questions and problems will greatly enhance your knowledge of the material. Answers to approximately half of these questions are provided in Appendix B.

If you need more practice, be sure to obtain a Coursebook and solve the questions and problems for each chapter. The Coursebook also contains the answers to the multiple choice questions and a brief explanation of why an answer is correct (and other choices incorrect). In most cases, if you master the concepts of the test items in the Coursebook, you will do well on the quizzes and examinations of your instructor. For extra help, in addition to the Coursebook, visit the book's student support Web site at **http://gwartney.swlearning.com** and take advantage of the Gwartney Xtra! Web site at **http://gwartneyxtra.swlearning.com**. Among other things, this Web site includes thirty five multiple choice practice questions for each chapter of the text.

ACKNOWLEDGMENTS

A project of this type is a team effort. Through the years, numerous people have assisted us in various ways. In this edition, the contributions of three people stand out. The first is developmental editor Amy Ray. She made substantial revisions to every chapter and worked as hard as we did to simplify the language and improve the readability of the text. She never gave up trying to get us to write like popular journalists rather than college professors. We really appreciate both her skillful revisions and personal commitment to the project. The contributions of Amy Gwartney and Jane Shaw Stroup were also invaluable. They edited several chapters prior to their submission to the publisher and, between the two of them, proofed every page of the text. Any future financial gain that they may derive will be well deserved.

We are also very much indebted to the excellent team of professionals at Thomson Business and Professional Publishing, including Peter Adams, acquisitions editor, for his help and support of our efforts; Bob Dreas, production editor, who coordinated the copy editing, proofreading, and indexing and kept the book on schedule; Jan Seidel, art and literary rights editor, who helped us locate and obtain permissions for the many photos and movie stills; Peggy Buskey, senior technology editor, who orchestrated the protection of the numerous electronic supplements, and John Carey, marketing manager, who worked hard to inform the marketplace about the advantages of the book.

We would also like to express our appreciation to Chuck Skipton of the University of Tampa for his contribution to what we believe is the very best set of PowerPoint slides accompanying an introductory economics text We are also appreciative of the contributions of Matthew Brown, Joseph Calhoun, Kerry King, Lynn MacDonald, Dirk Mateer, Samona Bociuba, and Linda Ghent to the Test Bank, online quizzes, and other supplementary materials. Robert Lawson of Capital University assisted us with the preparation of several exhibits. The text still bears an imprint of the contributions of Woody Studenmund of Occidental College and Gary Galles of Pepperdine University, who assisted us in numerous ways with past editions. We also want to express our appreciation to Amy Gwartney, Jane Shaw Stroup, Terri Sebel, and Karen Macpherson for their patience, support, and encouragement throughout the project.

Many instructors made important contributions to the 11th edition by providing us with insightful critical reviews. The following reviewers helped us improve the textbook immensely, and we thank them for their honest feedback:

Donald Boudreaux, George Mason University; Matthew Brown, Santa Clara University; Barry Haworth, University of Louisville; Jim Hubert, Seattle Central Community College; Katherine Huger, Charleston Southern University; Derek Kellenberg, Georgia Institute of Technology; Cory Krupp, Duke University; Robert Lawson, Capital University; James Payne, Eastern Kentucky University; Jennifer Platania, Elon University; Edward Stringham, San Jose State University; and Christopher Westley, Jacksonville State University.

We have often revised material in light of suggestions made by reviewers, users, friends, and even a few competitors. In this regard, we would also like to express our appreciation to the following people for their reviews and helpful suggestions of recent editions:

Douglas Agbetsiafa, Indiana University, South Bend; James C. W. Ahiakpor, California State University, Hayward; Ali T. Akarca, University of Illinois at Chicago; Stephen A. Baker, Capital University; Alana Bhatia, University of Colorado at Boulder; Edward J. Bierhanzl, Florida A&M University; Charles A. Booth, University of Alabama at Birmingham; Ford J. Brown, University of Minnesota Morris; Dennis Brennen, Harper College; James Bryan, Manhattanville College; Darcy R. Carr, Coastal Carolina University; Mike Cohick, Collin County Community College; David S. Collins, Virginia Highlands Community College; Jim F. Couch, University of North Alabama; Steven R. Cunningham, University of Connecticut; George W. Dollar, Clearwater Christian College; Jeff Edwards, Collin County Community College; Robert C. Eyler, Sonoma State University; James R. Fain, Oklahoma State University; Kathryn Finn, Western Washington University; Andrew W. Foshee, McNeese State University; Marsha Goldfarb University of Maryland, Balti

more County; David Harris, Northwood University; Ronald Helgens, Golden Gate University; Robert E. Herman, Nassau Community College/SUNY; William D. Hermann, Golden Gate University, San Francisco; Brad Hobbs, Florida Gulf Coast University; Woodrow W. Hughes, Jr., Converse College; Rob H. Kamery, Christian Brothers University; Frederic R. Kolb, University of Wisconsin, Eau Claire; Barbara Kouskoulas, Lawrence Technological University; David W. Kreutzer, James Madison University; George Kuljurgis, Oakland University; Randy W. LaHote, Washtenaw Community College; Tsung Hui Lai, Liberty University; Bob Lawson, Capital University; Don R. Leet, California State University, Fresno; George P. Lephardt, Milwaukee School of Engineering; Joe LeVesque, Northwood University; John McArthur, Wofford College; Ed Mills, Kendell College; David M. Mitchell, Oklahoma State University; Hadley T. Mitchell, Taylor University; Glen A. Moots, Northwood University; John R. Neal, Lake Sumter Community College; Lloyd Orr, Indiana University, Bloomington; Judd W. Patton, Bellevue College; Robert Reinke, University of South Dakota; Robert C. Rencher, Jr., Liberty University; Dan Rickman, Oklahoma State University; Karin L. Russell, Keiser College; Lewis F. Schlossinger, Community College of Aurora; Thomas W. Secrest, USC Coastal Carolina; Ben S. Shippen, Jr., Mercer University; Ken Somppi, Southern Union State Community College; William A. Steiden, Jefferson Community College; Richard D. C. Trainer, Warsaw School of Economics; Scott Ward, East Texas Baptist University; Tom Lee Waterston, Northwood University; Jim Wharton, Northwood University; Edward Wolpert, University of Central Florida; Janice Yee, Wartburg College; and Anthony Zambelli, Cuyamaca College.

About the Authors

James D. Gwartney holds the Gus A. Stavros Eminent Scholar Chair at Florida State University, where he directs the Stavros Center for the Advancement of Free Enterprise and Economic Education. He served as Chief Economist of the Joint Economic Committee of the U.S. Congress during 1999 2000. He is the coauthor of *Economic Freedom of the World,* an annual report on the institutions and policies of more than 120 countries that is published by a worldwide network of institutes. His publications have appeared in both professional journals and popular media such as the *Wall Street Journal* and the *New York Times.* His Ph.D. in economics is from the University of Washington. A member of the Mont Pelerin Society, he was invited by the incoming Putin Administration in March 2000 to make presentations and have discussions with leading Russian economists concerning the future of the Russian economy. In 2004 he was the recipient of the Adam Smith Award of the Association of Private Enterprise Education for his contribution to the advancement of free market ideals.

Richard L. Stroup is Professor of Economics and Interim Department Head at Montana State University, as well as Senior Associate at the Political Economy Research Center. With a Ph.D. in economics from the University of Washington, Professor Stroup has served as director of the Office of Policy Analysis in the U.S. Department of the Interior and has been published widely in professional journals. He is a contributing editor of numerous books on the economics of resources and the environment and authored *Economics* (Cato, 2003), which won the Sir Antony Fisher International Memorial Award presented by the Atlas Economic Research Foundation. Professor Stroup has lectured throughout the United States and abroad to professional and general audiences.

Russell S. Sobel is a professor of economics and director of the Entrepreneurship Center at West Virginia University. He received his Ph.D. in economics from Florida State University in 1994. Professor Sobel regularly teaches courses in both principles of economics and public choice economics. His enthusiastic teaching style has earned him many university teaching awards. Professor Sobel regularly gives lectures at economic education outreach programs on basic economic principles. Professor Sobel has published over thirty academic articles in refereed economics journals. His research focuses on applications of economics to public policy and on how economic freedom promotes entrepreneurship in an economy.

David A. Macpherson is the Rod and Hope Brim Eminent Scholar in Economics and is the director of the Pepper Institute on Aging and Public Policy at Florida State University, where he has received two university wide awards for teaching excellence. Professor Macpherson has written more than forty articles in leading economics journals, including the *Journal of Labor Economics, Industrial and Labor Relations Review,* and *Review of Economics and Statistics.* He is the coauthor of the annual *Union Membership and Earnings Data Book: Compilations from the Current Population Survey*, and is coauthor of *Contemporary Labor Economics*. Professor Macpherson also cowrote the book *Pensions and Productivity*. His current research interests include pensions, discrimination, labor unions, and the minimum wage. He received his undergraduate degree and Ph.D. from the Pennsylvania State University

Brief Contents

Part 1 : The Economic Way of Thinking **1**

Chapter 1	The Econom c Approach	4
Chapter 2	Some Tools of the Econom st	29

Part 2 : Markets and Government **54**

Chapter 3	Supp y, Demand, and the Market Process	56
Chapter 4	Supp y and Demand: App cat ons and Extens ons	83
Chapter 5	D ff'cult Cases for the Market and the Ro e of Government	107
Chapter 6	The Econom cs of Co ect ve Dec s on Mak'ng	124

Part 3 : Core Microeconomics **145**

Chapter 7	Consumer Cho ce and E asticity	148
Chapter 8	Costs and the Supply of Goods	170
Chapter 9	Pr ce Takers and the Compet t ve Process	195
Chapter 10	Pr ce Searcher Markets with Low Entry Barr'ers	216
Chapter 11	Pr ce Searcher Markets w th H gh Entry Barr ers	235
Chapter 12	The Supp y of and Demand for Product ve Resources	258
Chapter 13	Earn ngs, Product'v'ty, and the Job Market	278
Chapter 14	nvestment, the Cap ta Market, and the Wea th of Nat ons	294
Chapter 15	Income Inequality and Poverty	312
Chapter 16	Ga'ning from International Trade	329

Part 4 : Applying the Basics: Special Topics **353**

Specia Topic 1:	Government Spend ng and Taxat'on	356
Special Top c 2:	The Internet: How Is It Changing the Economy?	367
Spec a Top c 3:	The Econom cs of Soc a Secur ty	375
Spec a Topic 4:	The Stock Market: Its Funct on, Performance, and Potent al as an Investment	386
Spec a Top c 5:	The Econom cs of Hea th Care	397
Spec a Top c 6:	School Choice: Can It Improve the Quality of Education?	407
Special Topic 7:	Is Discr minat on Responsible for the Earnings D fference between Men and Women?	415
Spec al Top c 8:	Do Labor Un ons Increase the Wages of Workers?	421
Spec al Top c 9:	How Does Government Regu at on Affect Your L fe?	432
Spec'a Top c 10:	Resource Management· How Property Rights and Markets Rep ace Conf ct w th Cooperat on	443
Spec a Topic 11:	D ff cu t Env ronmenta Cases and the Ro e of Government	453

Append ces

A	General Bus ness and Economic Indicators	463
B	Answers to Se ected Critica Analysis Quest ons	471
Glossary		487
Index		493

Table of Contents

Part 1 The Economic Way of Thinking **1**

Chapter 1 The Economic Approach **4**

What Is Economics About? 5
Outstanding Economist
The Importance of Adam Smith, the Father of Economic Science **6**
Economics at the Movies
Ferris Bueller's Day Off (1986) **6**
The Economic Way of Thinking 10
Positive and Normative Economics 17
Pitfalls to Avoid in Economic Thinking 17
Economics as a Career 20
Looking Ahead **00**
Key Points **20**
Critical Analysis Questions **21**
Addendum
Understanding Graphs **23**

Chapter 2 Some Tools of the Economist **29**

What Shall We Give Up? 30
Trade Creates Value 32
Outstanding Economist
Thomas Sowell **32**
Economics at the Movies
Wall Street (1987) **33**
The Importance of Property Rights 35
Key to Prosperity
Private Ownership **35**
Applications in Economics
Protecting Endangered Species and the Environment with Private-Property Rights **39**
Production Possibilities Curve 40
Trade, Output, and Living Standards 45
Key to Prosperity
Gains from Trade **45**
Human Ingenuity and the Creation of Wealth 47
Key to Prosperity
Human Ingenuity **47**
Economic Organization 48
Looking Ahead **00**
Key Points **50**
Critical Analysis Questions **51**
Addendum
Comparative Advantage, Specialization, and Gains from Trade **52**

Part 2 Markets and Government **54**

Chapter 3 Supply, Demand, and the Market Process **56**

Consumer Choice and the Law of Demand 58
Changes in Demand Versus Changes in Quantity Demanded 61
Thumbnail Sketch
Factors That Cause Changes in Demand and Quantity Demanded **65**

Producer Cho ce and the Law of Supply 65
Key to Prosperity
Profits and Losses 66
Outstanding Economist
Alfred Marshall 68
Changes in Supply Versus Changes n Quantity Supplied 69
Thumbnail Sketch
Factors That Cause Changes in Supply and Quantity Supplied 69
How Market Pr ces are Determ ned: Supply and Demand Interact 71
Economics at the Movies
Pretty Woman (1990) 74
How Markets Respond to Changes n Demand and Supply 75
Thumbnail Sketch
How Changes in Demand and Supply Affect Market Price and Quantity 77
Invis'ble Hand Princ ple 78
Key to Prosperity
Invisible Hand Principle 78
Key Points 81
Critical Analysis Questions 81

Chapter 4 Supply and Demand: Applications and Extensions 83
The L nk between Resource and Product Markets 84
The Econcm'cs of Pr ce Contro s 86
Applications in Economics
The Imposition of Price Ceilings During Hurricane Hugo 87
Black Markets and the Importance of the Lega Structure 92
The mpact of a Tax 94
Tax Rates, Tax Revenues, and the Laffer Curve 99
Applications in Economics
The Laffer Curve and Mountain-Climbing Deaths 102
The Impact of a Subs'dy 102
Key Points 105
Critical Analysis Questions 105

Chapter 5 Difficult Cases for the Market, and the Role of Government 107
A C oser Look at Econom'c Eff c ency 108
If t's Worth Do ng, It's Worth Do ng mperfect y 109
Economics at the Movies
Along Came Polly (2004) 110
Thinking about the Economic Role of Government 110
Potent'al Shortcomings of the Market 112
Applications in Economics
Capturing External Benefits: The Case of Walt Disney World 117
Pu 'ng Th ngs Together 121
Key Points 122
Critical Analysis Questions 122

Chapter 6 The Economics of Collective Decision Making 124
The S ze and Growth of the U.S. Government 125
The D fferences and Similar t'es between Governments and Markets 127
Applications in Economics
Perspectives on the Cost of Political Competition: What Does It Cost to Get Elected? 127
Po 't cal Dec'sion Mak ng: An Overv ew 129
Outstanding Economist
James Buchanan 130
When the Po t cal Process Works We 132
When the Pol tical Process Works Poorly 134
Applications in Economics
Sweet Subsides to Sugar Growers: A case study of Special-Interest Effect 135

Economic Organization: Who Produces: Who Pays, and Why It Matters 140
The Economic Way of Thinking about Government 142
Could Constitutional Changes Help Promote Prosperity? 142
Thumbnail Sketch
What Weakens the Case for Market-Sector Allocation Versus Public-Sector Intervention, and Vice Versa **142**
Key Points **143**
Critical Analysis Questions **144**

Part 3 Core Microeconomics **145**

Chapter 7 Consumer Choice and Elasticity 148
The Fundamentals of Consumer Choice 149
Marginal Utility, Consumer Choice, and the Demand Curve of an Individual 150
Market Demand Reflects the Demand of Individual Consumers 154
Elasticity of Demand 154
How Demand Elasticity and Price Changes Affect Total Expenditures (or Revenues) on a Product 160
Income Elasticity 161
Price Elasticity of Supply 162
Key Points **163**
Critical Analysis Questions **163**
Addendum
Consumer Choice and Indifference Curves **164**

Chapter 8 Costs and the Supply of Goods 170
The Organization of the Business Firm 171
Costs, Competition, and the Corporation 174
Applications in Economics
Cooking the Books: How the Market Responds to Criminal Behavior **176**
The Economic Role of Costs 177
Short Run and Long Run Time Periods 179
Applications in Economics
Economic and Accounting Costs: A Hypothetical Example **179**
Categories of Costs 180
Output and Costs in the Short Run 180
Thumbnail Sketch
Compact Glossary on Cost **181**
Output and Costs in the Long Run 185
What Factors Cause Cost Curves to Shift? 189
The Economic Way of Thinking about Costs 190
Myths of Economics
A Good Business Decision Maker Will Never Sell a Product for Less than Its Production Costs 191
Key Points **192**
Critical Analysis Questions **193**

Chapter 9 Price Takers and the Competitive Process **195**
Price Takers and Price Searchers 196
Applications in Economics
The Aalsmeer Flower Auction: An Illustration of a Competitive Market **196**
What Are the Characteristics of Price Taker Markets? 197
Applications in Economics
Experimental Economics: the Significance of Competition 200
How Does the Price Taker Maximize Profit? 198
The Firm's Short Run Supply Curve 203
The Short Run Market Supply Curve 204
Price and Output in Price Takers Markets 204
Applications in Economics
Coffee Production in a Price-Taker Market 207
The Role of Profits and Losses 210

Competition Promotes Prosper'ty 211
Key to Prosperity
Competition 211
Outstanding Economist
Friedrich A. von Hayek 212
Key Points 213
Critical Analysis Questions 213

Chapter 10 Price-Searcher Markets with Low Entry Barriers 216
Compet t ve Price Searcher Markets 217
Economics at the Movies
You've Got Mail (1998) 220
Dynam'c Compet tion, Innovat'on, and Bus'ness Failures 221
Contestable Markets and the Competitive Process 222
Complex Dec s on Mak'ng and the Entrepreneur 223
Applications in Economics
Entrepreneurs at the Helm of Some Bizarre Occupations 224
Applications in Economics
Five Entrepreneurs Who Have Changed Our Lives 225
Entrepreneurship and Econom c Progress 227
Key to Prosperity
Entrepreneurship 227
Eva uating Compet tive Pr ce Searcher Markets 228
A Spec'a Case: Pr'ce D scrim'nat on 230
Key Points 232
Critical Analysis Questions 233

Chapter 11 Price-Searcher Markets with High Entry Barriers 235
Why Are Entry Barr'ers Sometimes H'gh? 236
The Character stics of a Monopo y 237
The Characteristics of an Oligopo y 241
Price and Output under Oligopoly 242
Market Power and Prof't The Ear y B'rd Catches the Worm 246
Applications in Economics
Oligopolistic Decision-Making Game Theory, and the Prisoner's Dilemma 247
Defects of Markets with H'gh Entry Barr'ers 248
Economics at the Movies
Beautiful Mind (2001) 248
Po cy A ternat ves When Entry Barr ers Are H gh 249
The Competit've Process n the Rea Wor d 253
Key Points 255
Critical Analysis Questions 256

Chapter 12 The Supply of and Demand for Productive Resources 258
Human and Nonhuman Resources 259
Outstanding Economist
Gary Becker 261
The Demand For Resources 261
Marginal Product vity and the Firm's H rirg Decis on 265
The Supp y of Resources 269
Supp y, Demand, and Resource Pr ces 272
The Coordinating Function of Resource Prices 274
Key Points 275
Critical Analysis Questions 276

Chapter 13 Earnings, Productivity, and the Job Market 278
Why Do Earn ngs Differ? 279
Applications in Economics
A College Degree as a Job Market Signal: Why You Should Take More Math 281
Thumbr ail Sketch
What are the Sources of Earnings Differentials? 284

Applications in Economics
America's Millionaires **285**
The Econom cs of Emp oyment D scr mination 285
The L nk between Product vity and Earn ngs 288
Key to Economic Progress
Link between Productivity and Earnings **288**
Myths of Economics
Automation Is the Major Cause of Unemployment **290**
Applications in Economics
Are Lifetime Jobs Disappearing? **291**
Key Points **292**
Critical Analysis Questions **292**

Chapter 14 Investment, the Capital Market, and the Wealth of Nations **294**
Why People nvest 295
Economics at the Movies
A Knight's Tale (2001) **296**
nterest Rates 296
The Present Va ue of Future Income and Costs 299
Present Va ue, Prof tab ity, and Investment 301
Applications in Economics
Would You Like to Become a Millionaire **302**
Invest ng n Human Cap ta 303
Uncerta nty, Entrepreneursh p, and Prof t 304
Outstanding Economist
Joseph Schumpeter **305**
The Cap ta Market and the Wea th of Nations 307
Key to Economic Progress
Innovation and the Capital Market **307**
Key Points **309**
Critical Analysis Questions **310**

Chapter 15 Income Inequality and Poverty **312**
How Much Income Inequa ty Exists n the United States? 313
rcome Mobi ty and nequality n Econom c Status 317
Measures of Economic Activity
How Is the Poverty Rate Calculated? **318**
Poverty n the Un ted States 319
Applications in Economics
Competition for Transfers and the Net Gain of Recipients **323**
ncome Inequal ty: *Some Concluding Thoughts* 324
Key Points **326**
Critical Analysis Questions **326**

Chapter 16 Gaining from International Trade **329**
The Trade Sector of the Un ted States 330
Ga ns From Specia zat on and Trade 331
Economics at the Movies
Cast Away (2000) **332**
Key to Prosperity
International Trade **336**
Supp y, Demand, and nternat onal Trade 337
The Econom cs of Trade Restrict ons 339
Why Do Nat ons Adopt Trade Restrictions? 342
Applications in Economics
Do More Open Economies Perform Better? **343**
Trade Barr ers and Popu ar Trade Fa ac es 345
The Changing Nature of Global Trade 347
Key Points **349**
Critical Analysis Questions **350**

Part 4 Applying the Basics: Special Topics 353

Special Topic 1 Government Spending and Taxation 356

Government Expenditures 357
Taxes and the Finance of Government 359
Size of Government: the U S Versus other Countries 364
Key Points 365
Critical Analysis Questions 365

Special Topic 2 The Internet: How Is It Changing the Economy? 367

Use of the Internet 368
Economic Gains From the Internet 368
Key Sectors of Internet Growth 370
More Efficient Consumer Markets 371
More Efficient Input Worksheets and Production Processes 372
Labor Markets and the Internet: Faster and Better Employee Employer Matches 372
Concluding Thought 373
Key Points 374
Critical Analysis Questions 374

Special Topic 3 The Economics of Social Security 375

Why Is Social Security Headed For Problems? 377
Will the Trust Fund Lighten the Future Tax Burden? 378
The Real Problem of the Current System 379
Who Is Helped and Who Is Hurt By Social Security? 380
Personal Retirement Accounts and Social Security Reform? 382
Key Points 384
Critical Analysis Questions 385

Special Topic 4 The Stock Market: Its Function, Performance, and Potential as an Investment Opportunity 386

The Economic Functions of the Stock Market 387
Stock Market Performance: the Historical Record 388
The Interest Rate, the Value of Future Income, and Stock Prices 389
The Random Walk Theory of the Stock Market 390
How the Ordinary Investor Can Beat the Experts 391
Key Points 394
Critical Analysis Questions 395

Special Topic 5 The Economics of Health Care 397

The Structure of the Health Care Industry 398
Third Party Payments and Health Care Inflation 400
Is the U.S. Headed for a Health Care Crisis? 402
Applications in Economics
***Could Imports from Canada Reduce the Drug Prices of Americans?* 403**
Averting the Crisis: How to Move Toward a Consumer Driven System 403
Key Points 406
Critical Analysis Questions 406

Special Topic 6 School Choice: Can It Improve the Quality of Education? 407

Educational Spending and Student Performance 408
Economics and the Structure of the Educational System 409
Alternative Ways of Increasing Competition and Expanding the Options of Consumers 411
The Effect of Structural Change 413
Conclusion 414
Key Points 414
Critical Analysis Questions 414

Special Topic 7 Is Discrimination Responsible for the Earnings Difference between Men and Women? 415

Emp oyment D'scrim'nation and the Earnings of Women 416
Marita Status and the Earn'ngs of Women 417
The Chang'ng Career Object ves of Women 417
Impl cat ons for the Future 420
Key Points 420
Critical Analysis Questions 420

Special Topic 8 Do Labor Unions Increase the Wages of Workers? 421

Un'on Membersh p as a Share of the Workforce 422
How Do Un'ons Influence Wages? 424
What G ves a Un on Strength? 426
The Wages of Un on and Nonunion Emp oyees 428
The mpact of un ons on the Wages of A Workers 430
Key Points 430
Critical Analysis Questions 431

Special Topic 9 How Does Government Regulation Affect Your Life? 432

Regulat on of Business 433
The Political Economy of Regulation 434
Applications in Economics
Limousines and Restraints on Entry 436
The Costs of Regu at on 437
Applications in Economics
Superfund: a Highly Inefficient Cleanup Program 438
Future D rect ons For Regulatory Pol cy 441
Key Points 442
Critical Analysis Questions 442

Special Topic 10 Resource Management: How Property Rights and Markets Replace Conflict with Cooperation 443

Property Rights: the Key to Environmenta Quality 444
Applications in Economics
Looking Ahead at the Flying D Ranch 448
How Property R ghts Affect Environmental Qua ty and L'fe Expectancy 448
Markets, Econom zing, and Resource Conservat on 448
Applications in Economics
Managing Forests for Nature, Too 450
Key Points 452
Critical Ana lysis Questions 452

Special Topic 11 Difficult Environmental Cases and the Role of Government 453

Government Regu at'on and the Environment 454
Climate Change and the Uncertainty Problem 454
Market ke Schemes: Reduc ng the Cost of Specif c Regu at ons 456
Applications in Economics
Fighting the Risks of Global Warming: Mitigation or Adaptation? 457
Property R ghts as a Too For Government Pol cy 458
Weaken ng Property R ghts: the Endangered Spec es Act 459
Government Ownersh p of Resources and Prov s on of Serv ces 461
Conclus'on 462
Key Points 462
Critical Analysis Questions 462

Appendix A General Business and Economic Indicators 463

Appendix B Answers to Selected Critical Analysis Questions 471

Glossary 487

Index 493

RELATIONSHIP BETWEEN MAIN EDITION AND THE MACRO/MICRO EDITIONS

Micro	Macro	Full-Length Text	
Part 1	**Part 1**	**Part 1**	**The Economic Way of Thinking**
1	1	1	The Economic Approach
2	2	2	Some Tools for the Economist
Part 2	**Part 2**	**Part 2**	**Markets and Governments**
3	3	3	Supply, Demand, and the Market Process
4	4	4	Supply and Demand: Applications and Extensions
5	5	5	Difficult Cases for the Market, and the Role of Government
6	6	6	The Economics of Collective Decision Making
	Part 3	**Part 3**	**Core Macroeconomics**
	7	7	Taking the Nation's Economic Pulse
	8	8	Economic Fluctuations, Unemployment, and Inflation
	9	9	An Introduction to Basic Macroeconomic Markets
	10	10	Working with Our Basic Aggregate Demand and Aggregate Supply Model
	11	11	Keynes and the Evolution of Macroeconomics
	12	12	Fiscal Policy
	13	13	Money and the Banking System
	14	14	Modern Macroeconomics and Monetary Policy
	15	15	Stabilization Policy, Output, and Employment
	16	16	Economic Growth and the Wealth of Nations
	Part 4	**Part 4**	**International Economics**
16	17	17	Gaining from International Trade
	18	18	International Finance and the Foreign Exchange Market
Part 3		**Part 5**	**Core Microeconomics**
7		19	Consumer Choice and Elasticity
8		20	Costs and the Supply of Goods
9		21	Price Takers and the Competitive Process
10		22	Price Searcher Markets with Low Entry Barriers
11		23	Price Searcher Markets with High Entry Barriers
12		24	The Supply of and Demand for Productive Resources
13		25	Earnings, Productivity, and the Job Market
14		26	Investment, the Capital Market, and the Wealth of Nations
15		27	Income Inequality and Poverty

Micro Part 4	Macro Part 5	Full-Length Text Part 6	Applying the Basics: Special Topics in Economics
1	1	1	Government Spending and Taxation
2	2	2	The Internet: How Is It Changing the Economy?
3	**3**	**3**	The Economics of Social Security
4	4	4	The Stock Market: Its Function, Performance, and Potential as an Investment Opportunity
	5	***5***	The Federal Budget and the National Debt
	6	***6***	Labor Market Policies and Unemployment: A Cross-Country Analysis
	7	***7***	Institutions, Policies, and the Irish Miracle
5		**8**	The Economics of Health Care
6		9	School Choice: Can It Improve the Quality of Education in America?
7		10	Is Discrimination Responsible for the Earnings Differences between Men and Women?
8		11	Do Labor Unions Increase the Wages of Workers?
9		12	How Does Government Regulation Affect Your Life?
10		13	Resource Management: How Property Rights and Markets Replace Conflict with Cooperation
11		14	Difficult Environmental Cases and the Role of Government

The Economic Way of Thinking

PART 1

Economics is about how people choose. The choices we make influence our lives and those of others. Your future will be influenced by the choices you make with regard to education, job opportunities, savings, and investment. Furthermore, changes in technology, demographics, communications, and transportation are constantly altering the attractiveness of various options and the opportunities available to us. The economic way of thinking is all about how incentives alter the choices people make. It can help you make better choices and enhance your understanding of our dynamic world.

CHAPTER 1

The Economic Approach

Economist, n.–A scoundrel whose faulty vision sees things as they really are, not as they ought to be.

after Ambrose Bierce

Chapter Focus

- **What is scarcity, and why is it important even in relatively wealthy economies?**
- **How does scarcity differ from poverty? Why does scarcity necessitate rationing and cause competition?**
- **What is the economic way of thinking? What is different about the way economists look at choices and human decision making?**
- **What is the difference between positive and normative economics?**

Welcome to the world of economics. You've heard about economics in the news. Maybe you think economics has to do with the stock market or the fine print you've read in the business section of your daily newspaper. You will soon see, however. that economics is much more than that. In fact, a field trip to the fruits and vegetables section at your local grocery store could well be filled with more economics lessons than a trip to the New York Stock Exchange.

In a nutshell, economics is the study of human behavior, with a particular focus on human decision making. In economics you will learn a new and powerful way of thinking that might lead you to question some of your current views and to look at things in a different way. As the satirical definition of an economist in the chapter-opening quote suggests, economic analysis provides valuable insights about how the world really works. These insights, however, often conflict with commonly held beliefs about the way things "ought" to work.

You may have heard some of the following statements: Gas prices are so high that the government should regulate them. The government should mandate air bags in cars to increase public health and safety. The minimum wage should be increased to help the poor. Tariffs should be imposed on foreign imports to save jobs in the United States. Technology, outsourcing, and robotics lead to unemployment and hurt our standard of living.

In this course, you will gain an understanding of these issues that may well alter the way you think about them. You will also develop new insights into how and why people (including yourself) make choices. This course will better enable you to argue political and economic issues with your friends at parties. It may even help you impress your date. On a more serious note, though, the better the decisions you make in your lifetime, the better off you will be. The same goes for a society as a whole. Who knows—you may become so good at economics you discover how to improve the lives of many people around the world, in addition to your own. You could even become the next great economist of our time.

The origins of economics date back to Adam Smith, a Scottish moral philosopher, who expressed the first economic ideas in his breakthrough book, *An Inquiry into the Nature and Causes of the Wealth of Nations,* published in 1776. As the title of his book suggests, Smith sought to explain why people in some nations were wealthier than those in others. This very question is still a central issue in economics. It is so important that throughout this book we will use a special "Keys to Economic Prosperity" symbol in the margin to highlight sections that focus on this topic. **A** listing of the major keys to prosperity is presented inside the front cover of the book. These keys and accompanying discussions will help you understand what factors enable economies, and their citizens, to grow wealthier and prosper.

WHAT IS ECONOMICS ABOUT?

While economics is about the choices *individuals* make, we often group together to form collective organizations, such as corporations, labor unions, nonprofit clubs, and governments. Individual choices, however, still underlie and direct the decisions made within these organizations. Thus, even when we study collective organizations like governments, we will still focus our analysis on the choices and decisions made by individuals within those organizations. We begin our journey into economics by discussing the constraints we face as individuals that force us to make choices.

[Economics is] the science which studies human behavior as a relationship between ends and scarce means which have alternative uses.

—Lionel Robbins[1]

Scarcity Means Having to Make Choices

Would you like some new clothes, a nicer car, and a larger apartment? How about better grades and more time to watch television, go skiing, and travel? Do you dream of driving

[1] Lionel Robbins, *An Essay on the Nature and Significance of Economic Science (1932).*

OUTSTANDING ECONOMIST

The [illegible] of Adam Smith the Father [illegible] Science

Economics is a relatively young science. The [illegible] tion of economics was laid in 1776, when Adam Smith (1723–1790) published *An Inquiry into the Nature and Causes of the Wealth of Nations.* Smith presented what was at that time a revolutionary view. He argued that the wealth of a nation does not lie in gold and silver, but rather in the goods and services produced and consumed by people. According to Smith, coordination, order, and efficiency would result without the planning and direction of a central authority.

Adam Smith was a lecturer at the University of Glasgow, in his native Scotland. Before economics, morals and ethics were actually his concern. His first book was *The Theory of Moral Sentiments.* For Smith, self-interest and sympathy for others were complementary. However, he did not believe that charity alone would provide the essentials for a good life. He stressed that free exchange and competitive markets would harness self-interest as a creative force. Smith believed that individuals *pursuing their own Interests* would be directed by the "invisible hand" of market prices toward the production of those goods that were most advantageous to society.

Ideas have consequences. Smith's ideas greatly influenced not only Europeans but also those who mapped out the structure of the U.S. government. Smith's notion of the "invisible hand" of the market has since become accepted as crucial to the prosperity of nations.[1]

[1]For an excellent biographical sketch of Adam Smith, see David Henderson, ed., *The Fortune Encyclopedia of Economics* (New York: Warner Books, 1993), 836–838. The entire text of this useful encyclopedia is now available online, free of charge, at http://www.econlib.org.

your brand-new Porsche into the driveway of your oceanfront house? As individuals, our desire for goods is virtually unlimited. We may want all of these things. Unfortunately, both as individuals and as a society we face a constraint called **scarcity** that prevents us from being able to completely fulfill our desires.

Scarcity
Fundamental concept of economics that indicates that there is less of a good freely available from nature than people would like.

Scarcity is present whenever there is less of a good or resource freely available from nature than people would like. There are some things that are not scarce — seawater comes

ECONOMICS AT THE MOVIES

PARAMOUNT/THE KOBAL COLLECTION

Ferris Bueller's Day Off (1986)

In one scene in *Ferris Bueller's Day Off,* Ben Stein plays an economics teacher lecturing about macroeconomics. His students are bored and falling asleep. Although some parts of economics might not be as fun as others, it's a misconception that economics is boring. On the contrary, economics will enlighten you about how people make decisions and the way the world works. It will also help you make better decisions yourself, which will make you better off.

Our "Economics at the Movies" features have been inspired by G Dirk Mateer, the author of Economics in the Movies (Thomson South-Western Publishing, 2005)

EXHIBIT 1
A General Listing of Scarce Goods and Limited Resources

History is a record of our struggle to transform available, but limited, resources into goods that we would like to have.

SCARCE GOODS	LIMITED RESOURCES
Food (bread, milk, meat, eggs, vegetables, coffee, etc.)	Land (various degrees of fertility)
Clothing (shirts, pants, blouses, shoes, socks, coats, sweaters, etc.)	Natural resources (rivers, trees, minerals, oceans, etc.)
Household goods (tables, chairs, rugs, beds, dressers, television sets, etc.)	Machines and other human-made physical resources
Education	Nonhuman animal resources
National defense	Technology (physical and scientific "recipes" of history)
Leisure time	Human resources (the knowledge, skill, and talent of individual human beings)
Entertainment	
Clean air	
Pleasant environment (trees, lakes, rivers, open spaces, etc.)	
Pleasant working conditions	

to mind; nature has provided as much of it as people want. But almost everything else you can think of—even your time—is scarce. In economics, the word *scarce* has a very specific meaning that differs slightly from the way it is commonly used. Even if large amounts of a good have been produced, it is still scarce as long as there is not as much of it *freely available from nature* as we would all like. For example, even though goods like apples and automobiles are relatively abundant in the United States, they are still scarce because we would like to have more of them than nature has freely provided. In economics, we generally wish to determine only if a good is scarce or not, and refrain from using the term to refer to the relative availability or abundance of a good or resource.

The unlimited nature of our desires, coupled with the limited nature of the goods and resources available to satisfy these desires, requires that we make choices. Should I spend the next hour studying or watching TV? Should I spend my last $20 on a new CD or on a shirt? Should this factory be used to produce clothing or furniture? **Choice,** the act of selecting among alternatives, is the logical consequence of scarcity. When we make choices, we constantly face trade-offs between meeting one desire or another. To meet one need, we must let another go unmet. The basic ideas of *scarcity* and *choice,* along with the *trade-offs* we face, provide the foundation for economic analysis.

Choice
The act of selecting among alternatives.

Resources are the ingredients, or inputs, people use to produce goods and services. Our ability to produce goods and services is limited precisely because of the limited nature of our resources.

Resource
An input used to produce economic goods. Land, labor, skills, natural resources, and capital are examples. Throughout history, people have struggled to transform available, but limited, resources into things they would like to have—economic goods.

Exhibit 1 lists a number of scarce goods and the limited resources that might be used to produce them. There are three general categories of resources. First, there are *human* resources—the productive knowledge, skill, and strength of human beings. Second, there are *physical* resources—things like tools, machines, and buildings that enhance our ability to produce goods. Economists often use the term **capital** when referring to these human-made resources. Third, there are *natural resources*—things like land, mineral deposits, oceans, and rivers. The ingenuity of humans is often required to make these natural resources useful in production. For example, until recently, the yew tree was considered a "trash tree," having no value. Then, scientists discovered that the tree produces taxol, a substance that could be used to fight cancer. Human knowledge and ingenuity made yew trees a valuable resource. As you can see, natural resources are important, but knowing how to use them productively is just as important. As economist Thomas Sowell points out, cavemen had the same natural resources at their disposal that we do today. The huge difference between their standard of living and ours reflects the difference in the knowledge they could bring to bear on those resources versus what we can.[2] Over time, human

Capital
Human-made resources (such as tools, equipment, and structures) used to produce other goods and services. They enhance our ability to produce in the future.

[2]Thomas Sowell, *Knowledge and Decisions* (New York Basic Books, 1980), 47

ingenuity, discovery, improved knowledge, and better technology have enabled us to produce more goods and services from the available resources. Nonetheless, we will never be able to produce enough goods to entirely fulfill human desires. Because scarcity can't be eliminated, people will always face choices. This is what economics is about.

Scarcity and Poverty Are Not the Same

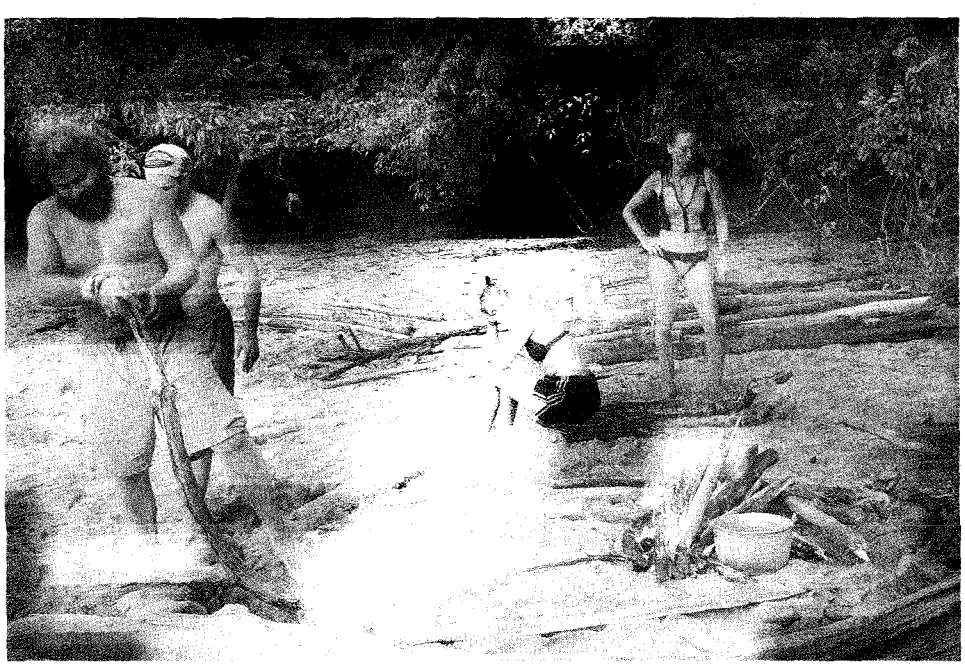
Think for a moment what life was like in 1750. People all over the world struggled 50, 60, and 70 hours a week to obtain the basic necessities of life—food, clothing, and shelter. Manual labor was the major source of income. Animals provided the means of transportation. Tools and machines were primitive by today's standards. As the English philosopher Thomas Hobbes stated in the seventeenth century, life was "solitary, poor, nasty, brutish, and short."[3]

Throughout much of South America, Africa, and Asia, economic conditions today continue to make life difficult. In North America, Western Europe, Oceania, and some parts of Asia, however, economic progress has substantially reduced physical hardship and human drudgery. In these regions, the typical family is more likely to worry about financing their summer vacation than obtaining food and shelter. As anyone who has watched the TV reality show *Survivor* knows, we take for granted many of the items that modern technological advances have allowed us to produce at unbelievably low prices. Contestants on *Survivor* struggle with even basic things like starting a fire, finding shelter, and catching fish. They are thrilled when they win ordinary items like shampoo. rice, and toilet paper. During one episode, a contestant eagerly paid over $125 for a small chocolate bar and spoonful of peanut butter at an auction—and she considered it a great bargain!

Objective
A fact based on observable phenomena that is not influenced by differences in personal opinion.

Subjective
An opinion based on personal preferences and value judgments.

It is important to note, however, that scarcity and poverty are not the same thing. Scarcity is an **objective** concept that describe3 a factual situation in which the limited nature of our resources keeps us from being able to completely fulfill our desires for goods and services. In contrast, poverty is a **subjective** concept that refers to a personal opinion of whether someone meets an arbitrarily defined level of income. This distinction is made even clearer when you realize that different people have vastly different ideas of what it



© CBS/LANDOV

The degree to which modern technology and knowledge allow us to fulfill our desires and ease the grip of scarcity is often taken for granted—as the castaways on the CBS reality series *Survivor* quickly find out when they have to struggle to meet even basic needs, such as food, shelter, and cleaning their bodies and clothes.

[3]Thomas Hobbes, *Leviathan* (1651) Part I, Chapter 13

means to be poor. The average family in the United States that meets the federal government's definition of being "in poverty" would be considered wealthy in most any country in Africa. Even in the United States as recently as the 1950s, a family was considered fairly wealthy if it had central heat and air conditioning, or more than one automobile or television set. In the United States today, the majority of families officially classified as in poverty have many, if not most, of the items that would have been viewed as symbols of wealth only 50 years ago.

The distinction between "needs" and "wants" helps us understand why it is impossible to objectively define poverty. Most people would agree that an absence of poverty means that some basic level of needs has been met. But they would disagree on what constitutes needs versus wants. In the 1920s, less than half of all households in the United States had electricity, and even fewer had a telephone or an automobile. Still, people survived and prospered. Would you consider electricity a need or a want? How about gasoline? How about other items that you generally hear people say they need, like cable television, a computer, and a $100 pair of tennis shoes—are they really needs? Although food and water are necessary for human survival, no one item (such as pizza, steak, a Big Mac, or a $1 bottle of spring water) is essential.

People always want more and better goods for themselves and others they care about—medical care, schooling, and national security are examples. Scarcity is the constraint that prevents us from having as much of *all* goods as we would like, but it is not the same as poverty. Even if every individual were rich, scarcity would still be present.

Scarcity Necessitates Rationing

Scarcity makes **rationing** a necessity. When a good or resource is scarce, some criterion must be used to determine who will receive it and who will go without. The choice of which method is used will, however, have an influence on human behavior. When rationing is done through the government sector, a person's political status and ability to manipulate the political process are the key factors. Powerful interest groups and those in good favor with influential politicians will be the ones who obtain goods and resources. When this method of rationing is used, people will devote time and resources to lobbying and favor seeking with those who have political power, rather than to productive activities.

Rationing
Allocating a limited supply of a good or resource among people who would like to have more of it. When price performs the rationing function, the good or resource is allocated to those willing to give up the *most* "other things" *in* order to get it.

When the criterion is first-come, first-served, goods are allocated to those who are fastest at getting in line or most willing to spend time waiting in line. Many colleges use this method to ration tickets to sporting events, and the result is students waiting in long lines (and sometimes even camping out overnight) to obtain tickets. Imagine how the behavior of students would change if tickets were instead given out to the students with the highest grade point average.

In a market economy, price is generally used to ration goods and resources only to those who are willing and able to pay the prevailing market price. Because only those goods that are scarce require rationing, in a market economy, one easy way to determine whether a good or resource is scarce is to ask if it sells for **a** price. If you have to pay for something, it is scarce.

Scarcity Leads to Competitive Behavior

Competition is a natural outgrowth of scarcity and the desire of human beings to improve their conditions. Competition exists in every economy and every society. It exists both when goods are allocated by price in markets and when they are allocated by other means—political decision making, for example.

How goods are rationed influences what competitive techniques people will use to get them. When the rationing criterion is price, individuals will engage in income-generating activities that enhance their ability to pay the price needed to buy the goods and services they want. Thus, one benefit of using price as a rationing mechanism is that it encourages individuals to engage in the production of goods and services to generate income. In con-

trast, rationing on the basis of first-come, first-served encourages individuals to waste a substantial amount of time unproductively waiting in line, while rationing through the political process encourages individuals to waste time attempting to influence the political process.

Within a market setting, the competition that results from scarcity is an important ingredient in economic progress. Competition among business firms for customers results in newer, better, and less expensive goods and services. Competition between employers for workers results in higher wages, benefits, and better working conditions. Further, competition encourages discovery and innovation. two important sources of growth and higher living standards.

THE ECONOMIC WAY OF THINKING

It [economics] is a method rather than a doctrine, an apparatus of the mind, a technique of thinking which helps its possessor to draw correct conclusions.

—John Maynard Keynes[4]

One does not have to spend much time around economists to recognize that there is an "economic way of thinking." Admittedly, economists, like others, differ widely in their ideological views. A news commentator once remarked that "any half-dozen economists will normally come up with about six different policy prescriptions." Yet, in spite of their philosophical differences, the approach of economists reflects common ground.

That common ground is **economic theory,** developed from basic principles of human behavior. Economic researchers are constantly involved in testing and seeking to verify their theories. When the evidence from the testing is consistent with a theory, eventually that theory will become widely accepted among economists. Economic theory, like a road map or a guidebook, establishes reference points indicating what to look for, and how economic issues are interrelated. To a large degree. the basic economic principles are merely common sense. When applied consistently, however, these commonsense concepts can provide powerful and sometimes surprising insights.

Economic theory
A set of definitions, postulates, and principles assembled in a manner that makes clear the "cause-and-effect" relationships.

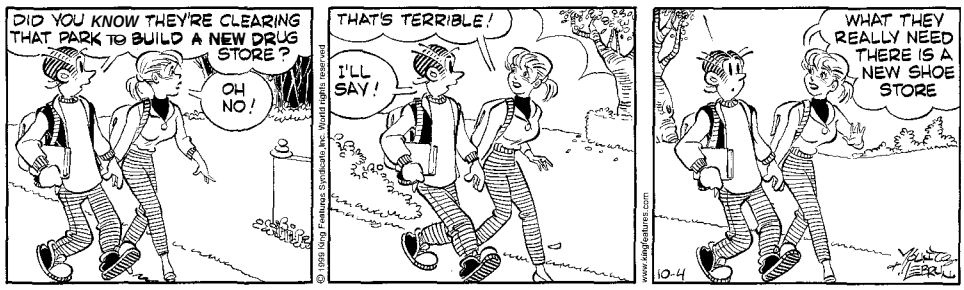
Eight Guideposts to Economic Thinking

The economic way of thinking requires incorporating certain guidelines—somewould say the building blocks of basic economic theory—into your own thought process. Once you incorporate these guidelines, economics can be a relatively easy subject to master. Students who have difficulty with economics have almost always failed to assimilate one or more of these principles. The following are eight principles that characterize the economic way of thinking. We will discuss each of these principles in more depth throughout the book so that you will be sure to understand hoa and when to apply them.

1. The use of scarce resources is costly, so trade-offs must be made. Economists sometimes refer to this as the "there is no such thing as a free lunch" principle. Because resources are scarce, the use of resources to produce one good diverts those resources from the production of other goods. A parcel of undeveloped land could be used for a new hospital or a parking lot, or it could simply be left undeveloped. No option is free of cost—there is always a trade-off. The choice to pursue any one of these options means the others must be sacrificed. The highest valued alternative that must be sacrificed is the **opportunity cost** of the option chosen. For example, if you use one hour of your scarce time to study economics, you will have one hour less time to watch television, read magazines, sleep, work at a job, or study other subjects. Whichever one of these options you would have chosen had you *not* spent the hour studying economics is your highest valued option forgone. If you would have been sleeping, then the opportunity cost of this hour spent studying economics is a forgone hour of sleep. In economics, the opportunity cost of an action is the highest valued option given up when a choice is made.

Opportunity cost
The highest valued alternative that must be sacrificed as a result of choosing an option.

[4]John Maynard Keynes (1883–1946) was an English economist whose writings during the 1920s and 1930s exerted an enormous impact on both economic theory and policy. Keynes established the terminology and the economic framework that are still widely used when economists study problems of unemployment and inflation.



Reprinted with special permission of King Features Syndicate.

When a scarce resource is used to meet one need. other competing needs must be sacrificed. The forgone shoe store is an example of the opportunity cost of building the new drugstore.

It is important to recognize that the use of scarce resources to produce a good is always costly, regardless of who pays for the good or service produced. In many countries, various kinds of schooling are provided free of charge to *students.* However, provision of the schooling is not free *to the community as a whole.* The scarce resources used to produce the schooling—to construct the building, hire teachers, buy equipment, and so on—could have been used instead to produce more recreation, entertainment, housing, medical care, or other goods. The opportunity cost of the schooling is the highest valued option given up because the resources required for its production were instead used for schooling.

By now the central point should be obvious. As we make choices, we are continually faced with trade-offs. Using resources to do one thing leaves fewer resources to do another.

Consider one final example. Mandatory air bags in automobiles save an estimated 400 lives each year. Economic thinking, however, forces us to ask ourselves if the $SO billion spent on air bags could have been used in a better way—perhaps say, for cancer research that could have saved *more* than 400 lives per year. Most people don't like to think of air bags and cancer research as an "either/or" proposition. It's more convenient to ignore these trade-offs. But if we want to get the most out of our resources, we have to consider all of our alternatives. In this case, the appropriate analysis is not lives saved with air bags versus dollars spent on them, but the number of lives that could have been saved (or other things that could have been accomplished) if the $SO billion had been used differently. A candid consideration of hard trade-offs like this is essential to using our resources wisely.

2. Individuals choose purposefully—they try to get the most from their limited resources. People try not to deliberately waste their valuable resources. Instead, they try to choose the options that best advance their personal desires and goals at the least possible cost. This is called **economizing behavior**. Economizing behavior is the result of purposeful, or rational, decision making. When choosing among things of equal benefit, an economizer will select the cheapest option. For example, if a pizza, a lobster dinner, and a sirloin steak are expected to yield identical benefits for Mary (including the enjoyment of eating them), economizing behavior implies that Mary will select the cheapest of the three alternatives, probably the pizza. Similarly, when choosing among alternatives of equal cost, economizing decision makers will select the option that yields the greatest benefit. If the prices of several dinner specials are equal, for example, economizers will choose the one they like the best. Because of economizing behavior, the desires or preferences of individuals are revealed by the choices they make.

Economizing behavior
Choosing the option that offers the greatest benefit at the least possible cost.

Purposeful choosing implies that decision makers have some basis for their evaluation of alternatives. Economists refer to this evaluation as **utility**—the benefit or satisfaction that an individual expects from the choice of a specific alternative. Utility is highly subjective, often differing widely from person to person. The steak dinner that delights one person may be repulsive to another (a vegetarian, for example).

Utility
The subjective benefit or satisfaction a person expects from a choice or course of action.

The idea that people behave rationally to get the greatest benefit at the least possible cost is a powerful tool. It can help us understand their choices. However, we need to realize that a rational choice is not the same thing as a "right" choice. If we want to understand people's choices, we need to understand their own subjective evaluations of their options *us they* ***see*** *them.* As we have said, different people have different preferences. If Joan prefers $50 worth of chocolate to $SO worth of vegetables, buying the chocolate would be the rational choice for her, even though some outside observer might say that Joan is

making a "bad" decision. Similarly, some motorcycle riders choose to ride without a helmet because they believe the enjoyment they get from riding without one is greater than the cost (the risk of injury). When people weigh the benefits they receive from an activity against its cost, they are making a rational choice—even though it might not be the choice you or I would make in the same situation.

3. Incentives matter—choice is influenced in a predictable way by changes in incentives. This is probably the most important guidepost in economic thinking. It is sometimes called the basic postulate of all economics. As the personal benefits from an option increase, a person will be more likely to choose it. On the other hand, as the personal costs associated with an option increase, a person will be less likely to choose it. This guidepost also applies to groups of people, and suggests that making an option more beneficial will predictably cause more of them to choose it. Similarly, making an option more costly will cause fewer of them to choose it.

This basic idea is a powerful tool because its usefulness is practically universal. Incentives affect behavior in virtually all aspects of our lives, ranging from market decisions about what to buy to political choices concerning for whom to vote. If beef prices rise, making beef consumption more expensive relative to other goods, consumers will be less likely to buy it. The "incentives matter" postulate also explains why a person would be unlikely to vote for a political candidate who, if elected, would raise taxes to fund a new government program he or she didn't like very much.

Most errors in economic reasoning occur because people overlook this postulate or fail to apply it consistently. With economic applications generally focusing on people trying to satisfy material desires, casual observers often argue that incentives matter only in cases of human selfishness. This view is false. People are motivated by a variety of goals, some humanitarian and some selfish, and incentives matter equally in both. Even an unselfish individual would be more likely to attempt to rescue a drowning child from a three-foot swimming pool than the rapid currents approaching Niagara Falls. Similarly, people are more likely to give a needy person their hand-me-downs rather than their favorite new clothes.

It is clear that incentives, whether monetary or nonmonetary, matter in human decision making. People will be less likely to walk down a dark alleyway than a well-lit one; they will be more likely to take a job if it has good benefits and working conditions than if

GETTY IMAGES

Because consumers respond to incentives, store owners know they can sell off excess inventory by reducing prices.

it doesn't; and they will be more likely to bend down and pick up a quarter lying on the sidewalk than they will a penny. Even a person who normally bends down to pick up pennies on the sidewalk probably would be less likely to if late for an important appointment, or on a first date.

Just how far can we push the idea that incentives matter? If asked what would happen to the number of funerals performed in your town if the price of funerals rose, how would you respond? The "incentives matter" postulate predicts that the higher cost would reduce the number of funerals. While the same number of people will still die each year, the number of funerals performed will still fall as more people choose to be cremated or buried in cemeteries in other towns. Substitutes are everywhere—even substitutes for funerals.

4. Individuals make decisions at the margin. When making a choice between two alternatives, individuals generally focus on the *difference* in the costs and benefits between alternatives. Economists describe this process as **marginal** decision making, or "thinking at the margin." The last time you went to eat fast food, you probably faced a decision that highlights this type of thinking. Will you get the $1.50 cheeseburger and the $1.00 medium drink, or instead get the $3.00 value meal that has the cheeseburger and drink and also comes with a medium order of fries? Naturally, individual decision making focuses on the difference between the alternatives. The value meal costs 50 cents more (its marginal cost) but will give you one extra food item—the fries (its marginal benefit). Your marginal decision is whether it is worth the extra 50 cents to have the fries. If you pay attention, you'll notice yourself frequently thinking at the margin. Next time you find yourself asking a salesclerk "How much *more* is this one?" when you are choosing between two items, you are doing a marginal analysis.

Marginal
Term used to describe the effects of a change in the current situation. For example, a producer's marginal cost is the cost of producing an additional unit of a product, given the producer's current facility and production rate.

Marginal choices always involve the effects of net additions to or subtractions from current conditions. In fact, the word *additional* is often used as a substitute for *marginal*. For example, a business decision maker might ask, "What is the marginal cost of producing one more, or additional, unit?' Marginal decisions may involve large or small changes. The "one more unit" could be a new factory or a new stapler. It is marginal because it involves additional costs and additional benefits. Given the current situation, what marginal benefits (additional sales revenues, for example) can be expected from the new factory, and what will be the marginal cost of constructing it? What is the marginal benefit versus marginal cost of purchasing a new stapler? The answers to these questions will determine whether building the new factory or buying the new stapler is a good decision.

It is important to distinguish between *average* and *marginal*. A manufacturer's average cost of producing automobiles (which would be the total cost of production divided by the total number of cars the manufacturer produces) may be $25,000, but the marginal cost of producing an additional automobile (or an additional 1,000 automobiles) might be much lower, say, $10,000 per car. Costs associated with research, testing, design, molds, heavy equipment, and similar factors of production must be incurred whether the manufacturer is going to produce 1,000 units, 10,000 units, or 100,000 units. Such costs will clearly contribute to the average cost of an automobile, but they will change very little as additional units are produced. Thus, the marginal cost of additional units may be substantially less than the average cost. Should production be expanded or reduced? That choice should be based on marginal costs, which indicate the *change* in total cost due to the decision.

Confusion between marginal and total benefits or costs can also be a source of error. Almost all of the choices we make are marginal, rather than all-or-nothing decisions. For example, we don't make decisions between eating or wearing clothes—dining well in the nude versus starving in style. Instead, we choose between having a little more food at the cost of a little less clothing, or a little less of something else. So the relevant comparison is not between the total value of food and the total value of clothing but between their marginal values.

People commonly ignore the implications of marginal thinking in their comments, but seldom in their actions. Thus, the concept is far better at explaining how people act than what they say. Students are often overheard telling other students that they shouldn't skip class because they have paid to enroll in it. Of course, the tuition is not a factor relevant at

the margin—it will be the same whether or not the student attends class on that particular day. The only real marginal considerations are what the student will miss that day (a quiz, information for the exam, etc.) versus what he or she could do with the extra time by skipping class. This explains why even students who tell others they paid too much for the class to skip it will ignore the tuition costs when they themselves decide to skip class.

When we confront a decision, the *marginal benefit* and *marginal cost* associated with the choice will determine our decision. Marginal analysis will be used extensively throughout this course. As we develop this concept further, you should pay special attention to understanding how to use it properly.

5. Although information can help us make better choices, its acquisition is costly. Information that helps us make better choices is valuable. However, the time needed to gather it is scarce, making information costly to acquire. As a result, people economize on their search for information just like they do anything else. For example, when you purchase a pair of jeans, you might evaluate the quality and prices of jeans at several different stores. At some point, though, you will decide that additional comparison shopping is simply not worth the trouble. You will make a choice based on the limited information you already have.

The process is similar when individuals search for a restaurant, a new car, or a roommate. They will seek to acquire some information, but at some point, they will decide that the expected benefit derived from gathering still more information is simply not worth the cost. When differences among the alternatives are important to decision makers, they will spend more time and effort gathering information. People are much more likely to read a consumer ratings magazine before purchasing a new automobile than they are before purchasing a new can opener. Because information is costly for people to acquire, limited knowledge and uncertainty about the outcome generally characterize the decision-making process.

6. Beware of the secondary effects: Economic actions often generate indirect as well as direct effects. In addition to direct effects that are quickly visible, people's decisions often generate indirect, or "secondary," effects that may be observable only with time. Failure to consider secondary effects is one of the most common economic errors because these effects are often quite different from initial, or direct, effects. Frederic Bastiat, a nineteenth-century French economist, stated that the difference between a good and a bad economist is that the bad economist considers only the immediate, visible effects, whereas the good economist is also aware of the **secondary effects**. The true cause of these secondary effects might not be seen, even later, except by those using the logic of good economics.

Secondary effects
The indirect impact of an event or policy that may not be easily and immediately observable. In the area of policy, these effects are often both unintended and overlooked.

Perhaps a few simple examples that involve both immediate (direct) and secondary (indirect) effects will help illustrate the point. The immediate effect of an aspirin is a bitter taste in one's mouth. The secondary effect, which is not immediately observable, is relief from a headache. The short-term direct effect of drinking twelve cans of beer might be a warm, jolly feeling. In contrast, the secondary effect is likely to be a sluggish feeling the next morning, and perhaps a pounding headache.

Sometimes, as in the case of the aspirin, the secondary effect—headache relief—is actually an intended consequence of the action. In other cases, however, the secondary effects are unintended. Changes in government policy often alter incentives, indirectly affecting how much people work, earn, invest, consume, and conserve for the future. When a change alters incentives, *unintended consequences* that are quite different from the intended consequences may occur.

Let's consider a couple of examples that illustrate the potential importance of unintended side effects. In an effort to reduce gasoline consumption, the federal government mandates that automobiles be more fuel efficient. Is this regulation a sound policy? It may be, but when evaluating the policy's overall impact, one should not overlook its secondary effects. To achieve the higher fuel efficiency, auto manufacturers will reduce the size and weight of vehicles. As a result, there will be more highway deaths—about 2,000 more per

year—than would otherwise occur because these lighter cars do not offer as much protection for occupants. Furthermore, because the higher mileage standards for cars and light trucks make driving cheaper, people tend to drive more than they otherwise would. This increases congestion and results in a smaller reduction in gasoline consumption than was intended by the regulation. Once you consider the secondary effects, the fuel efficiency regulations are much less beneficial than they might first appear.

Trade restrictions between nations have important secondary effects as well. The proponents of tariffs and quotas on foreign goods almost always ignore the secondary effects of their policies. Import quotas restricting the sale of foreign-produced sugar in the U.S. market, for example, have led to sugar prices that are about three times what they are in the rest of the world. The proponents of this policy—primarily sugar producers—argue that the quotas "save jobs" and increase employment. No doubt, the employment of sugar growers in the United States is higher than it otherwise would be. But what about the secondary effects? The higher sugar prices mean it's more expensive for U.S. firms to produce candy and other products that use a lot of sugar. As a result, many candy producers, including the makers of Life Savers, Jaw Breakers, Red Hots, and Fannie May and Fanny Farmer chocolates, have moved to countries like Canada and Mexico, where sugar can be purchased at its true market price. Thus, employment among sugar-using firms in the United States is reduced. Further, because foreigners sell less sugar in the United States, they have less purchasing power with which to buy products we export to them. This, too, reduces U.S. employment.

Once the secondary effects of trade restrictions like the sugar quota program are taken into consideration, we have no reason to expect that U.S. employment will increase as a result. There may be more jobs in favored industries, but there will be less employment in others. Trade restrictions reshuffle employment rather than increase it. But those who unwittingly fail to consider the secondary effects will miss this point. Clearly, consideration of the secondary effects is an important ingredient of the economic way of thinking.

7. The value of a good or service is subjective. Preferences differ, sometimes dramatically, between individuals. How much is a ticket to see a performance of the Bolshoi Ballet worth? Some people would be willing to pay a very high price, while others might prefer to stay home, even if tickets were free! Circumstances can change from day to day, even for a given individual. Alice, a ballet fan who usually would value the ticket at more than its price of $100, is invited to a party and suddenly becomes uninterested in

"Everybody wants to be sick. I'm using M&M's for pills."

© Bil Keane, Inc. Reprinted with special permission of King Features Syndicate.

Sometimes actions change the incentives people face and they respond accordingly, creating secondary effects that were not intended.

attending the ballet. Now what is the ticket worth? If she knows a friend who would give her $40 for the ticket, it is worth at least that much. If she advertises the ticket on eBay and gets $60 for it, a higher value is created. But if someone who doesn't know of the ticket would have been willing to pay even more, then a potential trade creating even more value is missed. If that particular performance is sold out, perhaps someone in town would be willing to pay $120. One thing is certain: The value of the ticket depends on several things, including who uses it and under what circumstances.

Economics recognizes that people can and do value goods differently. Mike may prefer to have a grass field rather than a parking lot next to his workplace and be willing to bear the cost of walking farther from his car each day. Kim, on the other hand, may prefer the parking lot and the shorter walk. As a science, economics does not place any inherent moral judgment or value on one person's preferences over another's—in economics all individuals' preferences are counted equally. Because the subjective preferences of individuals differ, it is difficult for one person to know how much another will value an item.

Think about how hard it is to know what would make a good gift for even a close friend or family member. Thus, arranging trades, or otherwise moving items to higher valued users and uses, is not a simple task. The entrepreneurial individual, who knows how to locate the right buyers and arranges for goods to flow to their highest valued use, can sometimes create huge increases in value from existing resources. In fact, people moving goods toward those who value them most and combining resources into goods that individuals value more highly is a primary source of economic progress.

Scientific thinking
Developing a theory from basic principles and testing it against events in the real world. Good theories are consistent with and help explain real-world events. Theories that are inconsistent with the real world are invalid and must be rejected.

8. The test of a theory is its ability to predict. Economic thinking is **scientific thinking**. The proof of the pudding is in the eating. How useful an economic theory is depends on how well it predicts the future consequences of economic action. Economists develop economic theories using scientific thinking based on basic principles. The idea is to predict how incentives will affect decision makers and compare the predictions against real-world events. If the events in the real world are consistent with a theory, we say that the theory has *predictive value* and is therefore valid.

If it is impossible to test the theoretical relationships of a discipline, the discipline does not qualify as a science. Because economics deals with human beings who can think and respond in a variety of ways, can economic theories really be tested? The answer to this question is yes, if, on average, human beings respond in predictable and consistent ways to changes in economic conditions. The economist believes that this is the case, even though not all individuals will respond in the specified manner. Economists usually do not try to predict the behavior of a specific individual; instead, they focus on the general behavior of a large number of individuals.

In the 1950s, economists began to do laboratory experiments to test economic theories. Individuals were brought into laboratories to see how they would act in buying and selling situations, under differing rules. For example, cash rewards were given to individuals who, when an auction was conducted, were able to sell at high prices and buy at low prices, thus approximating real-world market incentives. These experiments have verified many of the important propositions of economic theory.

Laboratory experiments, however, cannot duplicate all real economic interactions. How can we test economic theory when controlled experiments are not feasible? This is a problem, but economics is no different from astronomy in this respect. Astronomers can use theories tested in physics laboratories, but they must also deal with the world as it is. They cannot change the course of the stars or planets to see what impact the change would have on the gravitational pull of Earth. Similarly, economists cannot arbitrarily change the prices of cars or unskilled-labor services in real markets just to observe the effects on quantities purchased or levels of employment. However, economic conditions (for example, prices, production costs, technology, and transportation costs), like the location of the planets, do change from time to time. As actual conditions change, an economic theory can be tested by comparing its predictions with real-world outcomes. Just as the universe is the main laboratory of the astronomer, the real-world economy is the primary laboratory of the economist.

POSITIVE AND NORMATIVE ECONOMICS

As a social science, economics is concerned with predicting or determining the impact of changes in economic variables on the actions of human beings. Scientific economics, commonly referred to as **positive economics,** attempts to determine "what is." Positive economic statements involve potentially verifiable or refutable propositions. For example: "If the price of gasoline rises, people will buy less gasoline." We can statistically investigate (and estimate) the relationship between gasoline prices and gallons sold. We can analyze the facts to determine the correctness of a positive economic statement. Remember, a positive economic statement need not be correct, it simply must be testable.

In contrast, **normative economics** is about "what ought to be," given the preferences and philosophical views of the advocate. Value judgments often result in disagreement about normative economic matters. Two people may differ on a policy matter because one is from one political party and the other is from another, or because one wants cheaper food while the other favors organic farming (which is more expensive), and so on. They may even agree about the expected outcome of altering an economic variable (that is, the positive economics of an issue), but disagree as to whether that outcome is desirable.

Unlike positive economic statements, normative economic statements can neither be confirmed nor proven false by scientific testing. "Business firms should not be concerned with profits." "We should have fewer parking lots and more green space on campus." "The price of gasoline is too high." These normative statements cannot be scientifically tested because their validity rests on value judgments.

Normative economic views can sometimes influence our attitude toward positive economic analysis, however. When we agree with the objectives of a policy, it's easy to overlook the warnings of positive economics. Although positive economics does not tell us which policy is best, it can provide evidence about the likely effects of a policy. Sometimes proponents unknowingly support policies that are actually in conflict with their own goals and objectives. Positive economics, based on sound economic logic, can help overcome this potential problem.

Economics can expand our knowledge of how the real world operates, in both the private and the public (government) sectors. If we do not fully understand the implications, including the secondary effects, of alternative actions, we will not be able to choose intelligently. Yet, it is not always easy to use economic thinking to isolate the impact of a change. Let's now consider some pitfalls to avoid in economic thinking.

A positive science may be defined as a body of systematized knowledge concerning what is; a normative or regulative science is a body of systematized knowledge relating to criteria of what ought to be, and concerned therefore with the ideal as distinguished from the actual.

—John Neville Keynes[5]

Positive economics
The scientific study of "what is" among economic relationships.

Normative economics
Judgments about "what ought to be" in economic matters. Normative economic views cannot be proven false because they are based on value judgments.

PITFALLS TO AVOID IN ECONOMIC THINKING

Violation of the *Ceteris Paribus* Condition Can Lead One to Draw the Wrong Conclusion

Economists often qualify their statements with the words *ceteris paribus*. *Ceteris paribus* is a Latin term meaning "other things constant." An example of a *ceteris paribus* statement would be the following: "*Ceteris paribus,* an increase in the price of housing will cause buyers to reduce their purchases of housing." Unfortunately, we live in a dynamic world, so things seldom remain constant. For example, as the price of housing rises, the income of consumers might also increase. Each of these factors—higher housing prices and increasing consumer income—will have an impact on housing purchases. In fact, we would generally expect them to have opposite effects: Higher prices are likely to reduce housing purchases, while higher consumer incomes are likely to increase them. We point out this pitfall because sometimes statistical data (or casual observations) do not support economic theories. In most of these cases, other factors have also changed. The effects observed simply reflect the combined effect of these changes.

Ceteris paribus
A Latin term meaning "other things constant," used when the effect of one change is being described, recognizing that if other things changed. the) also could affect the result. Economists often describe the effects of one change, knowing that in the real world, other things might change and also exert an effect.

[5]John Neville Keynes, *The Scope and Method of Political Economy,* 4th ed. (1917), 34–35

The task of sorting out the effects of two or more variables that change at the same time is difficult. However, with a strong grip on economic theory, some ingenuity, and enough data, it can usually be done. This is, in fact, precisely the day-to-day work of many professional economists.

Good Intentions Do Not Guarantee Desirable Outcomes

There is a tendency to believe that if the proponents of a policy have good intentions, their proposals must be sound. This is not necessarily the case. Proponents may be unaware of some of the adverse secondary effects of their proposals, particularly when they are indirect and observable only over time. Even if their policies would be largely ineffective, politicians may still find it advantageous to call attention to the severity of a problem and propose a program to deal with it. In other cases, proponents of a policy may actually be seeking a goal other than the one they espouse. They may tie their arguments to objectives that are widely supported by the general populace. Thus, the fact that an advocate says a program will help the poor, increase wages, improve health care, expand employment, or achieve some other highly desirable objective, does not necessarily make it so.

Let's begin with a couple of straightforward examples. Federal legislation has been introduced that would require all children, including those under age two, to be fastened in a child safety seat when traveling by air. Proponents argue the legislation will increase the survival rate of children in the case of an airline crash and thereby save lives. Certainly, saving lives is a highly desirable objective, but will this really be the case? *Some* lives will probably be saved. But what about the secondary effects? The legislation would mean that a parent traveling with a small child would have to purchase an additional ticket, which will make it more expensive to fly. As a result, many families will choose to travel by auto rather than air. Because the likelihood of a serious accident per mile traveled in an automobile is several times higher than for air travel, more automobile travel will result in more injuries and fatalities. In fact, studies indicate that the increase in injuries and fatalities from additional auto travel will exceed the number of lives saved by airline safety seats.[6] Thus, even though the intentions of the proponents may well be lofty, there is reason to believe that the net impact of their proposal will be more fatalities and injuries than would be the case in the absence of the legislation.

The stated objective of the Endangered Species Act is to protect various species that are on the verge of extinction. Certainly, this is a admirable objective, but there is nonetheless reason to question the effectiveness of the Act itself. The Endangered Species Act allows the government to regulate the use of individual private property if an endangered species is found present on *or* near his or her land. To avoid losing control of their property, many landowners have taken steps to make their land less attractive as a natural habitat for these endangered species. For example, the endangered red-cockaded woodpecker nests primarily in old trees within southern pine ecosystems. Landowners have responded by cutting down trees the woodpeckers like to nest in to avoid having one nest on their land, which would result in the owner losing control of this part of their property. The end result is that the habitat for these birds has actually been disappearing more rapidly.

As you can see, good intentions are not enough. An unsound proposal will lead to undesirable outcomes even if it is supported by proponents with good intentions. But economic thinking can help us avoid this pitfall.

Association Is Not Causation

In economics, identifying cause-and-effect relationships is very important. But statistical association alone cannot establish this causation. Perhaps an extreme example will illus-

[6]For a detailed analysis of this subject, see Thomas B. Newman, Brian D. Johnston, and David C. Grossman, "Effects and Costs of Requiring Child-Restraint Systems for Young Children Traveling on Commercial Airplanes," *Archives of Pediatrics and Adolescent Medicine* 157 (October 2003): 969–74.

trate the point. Suppose that each November a witch doctor performs a voodoo dance designed to summon the gods of winter, and that soon after the dance is performed, the weather in fact begins to turn cold. The witch doctor's dance is associated with the arrival of winter, meaning that the two events appear to have happened in conjunction with one another. But is this really evidence that the witch doctor's dance actually caused the arrival of winter? Most of us would answer no, even though the two events seemed to happen in conjunction with one another.

Those who argue that a causal relationship exists simply because of the presence of statistical association are committing a logical fallacy known as the *post hoc propter ergo hoc* fallacy. Sound economics warns against this potential source of error.

The Fallacy of Composition: What's True for One Might Not Be True for All

What is true for the individual (or subcomponent) may not be true for the group (or the whole). If you stand up for an exciting play during a football game, you will be better able to see. But what happens if everyone stands up at the same time? Will everyone be better able to see? The answer is, of course, no. Thus, what is true for a single individual does not necessarily apply to the group as a whole. When everyone stands up, the view for individual spectators fails to improve; in fact, it may even become worse.

People who mistakenly argue that what is true for the part is also true for the whole are said to be committing the **fallacy of composition**. What is true for the individual can be misleading and is often fallacious when applied to the entire economy. The fallacy of composition highlights the importance of considering both a micro view and a macro view in the study of economics. **Microeconomics** focuses on the decision making of consumers, producers, and resource suppliers operating in a narrowly defined market, such as that for a specific good or resource. Because individual decision makers are the moving force behind all economic action, the foundations of economics are clearly rooted in a micro view.

As we have seen, however, what is true for a small unit may not be true in the aggregate. **Macroeconomics** focuses on how the aggregation of individual micro-units affects our analysis. Like microeconomics, it is concerned with incentives, prices, and output. Macroeconomics, however, aggregates markets, lumping together all 120 million households in this country. Macroeconomics involves topics like total consumption spending, saving, and employment, in the economy as a whole. Similarly, the nation's 7 million business firms are lumped together in "the business sector." What factors determine the level of aggregate output, the rate of inflation, the amount of unemployment, and interest rates? These are macroeconomic questions. In short, macroeconomics examines the forest rather than the individual trees. As we move from the micro-components to a macro view of the whole, it is important that we beware of the fallacy of composition.

Fallacy of composition
Erroneous view that what is true for the individual (or the part) will also be true for the group (or the whole).

Microeconomics
The branch of economics that focuses on how human behavior affects the conduct of affairs within narrowly defined units, such as individual households or business firms.

Macroeconomics
The branch of economics that focuses on how human behavior affects outcomes in highly aggregated markets, such as the markets for labor or consumer products.

LOOKING AHEAD

The primary purpose of this book is to encourage you to develop the economic way of thinking so that you can separate sound reasoning from economic nonsense. Once you have developed the economic way of thinking, economics will be relatively easy. Using the economic way of thinking can also be fun. Moreover, it will help you become a better citizen. It will give you a different and fascinating perspective on what motivates people, why they act the way they do, and why their actions sometimes go against the best interest of the community or nation. It will also give you valuable insight into how people's actions can be rechanneled for the benefit of the community at large.

ECONOMICS AS A CAREER

If you find yourself doing well in this course and discover that economics interests you, you may want to think about majoring in it. Graduating with a major in economics provides a variety of career choices. Many students go on to graduate school in economics, business, public administration, or law. Graduate M.B.A. and law programs find economics majors particularly attractive because of their strong analytical skills.

A graduate degree (a master's or doctorate) in economics is typically required to pursue a career as a professional economist. About one-half of all professional economists are employed by colleges and universities as teachers and researchers. Professional economists also work for the government or private businesses. Most major corporations have a staff of economists to advise them. Governments employ economists to analyze the impact of policy alternatives. The federal government's Council of Economic Advisers provides the president with analyses of how the activities of the government influence the economy.

Students who major in economics but who do not pursue graduate school still have many job opportunities. Because economics is a way of thinking, knowledge of it is a valuable decision-making tool that can be used in almost any job. Undergraduate majors in economics typically work in business, government service, banking, or insurance. Opportunities for people with undergraduate economics degrees to teach the subject at the high school level are also increasing. Arnold Schwarzenegger, Mick Jagger, and Ronald Reagan are among the long list of famous undergraduate economics majors!

The average salary of an economics graduate is comparable to that of finance and accounting graduates and is generally higher than those with management or marketing degrees. Professional economists with graduate degrees in economics who work for private business average approximately $90,000 per year, and those who choose to work as teachers and researchers at colleges and universities earn approximately $75,000 annually. Although salaries vary substantially, the point is that a career in economics can be rewarding both personally and financially.

Even if you choose not to major in economics, you will find that your economics courses will broaden your horizons and increase your ability to understand and analyze what is going on around you in the worlds of politics, business, and human relations. Economics is a social science that often overlaps with the fields of political science, sociology, and psychology. Because the economic way of thinking is so useful in making sense of the world around us, economics has sometimes been called the "queen of the social sciences." Reflecting this, economics is the only social science for which a Nobel Prize of the Swedish Academy of Science is awarded.

! KEY POINTS

- Scarcity and choice are the two essential ingredients of economic analysis. A good is scarce when the human desire for it exceeds the amount freely available from nature. Scarcity requires us to choose among available alternatives. Every choice entails a trade-off.
- Every society will have to devise some method of rationing the scarce resources among competing uses. Markets generally use price as the rationing device. Competition is a natural outgrowth of the need to ration scarce goods.
- Scarcity and poverty are not the same thing. Absence of poverty implies that some basic level of need has been met. An absence of scarcity implies that our desires for goods are fully satisfied. We may someday eliminate poverty, but scarcity will always be with us.
- Economics is a way of thinking that emphasizes eight points:
  1. The use of scarce resources to produce a good always has an opportunity cost.

2. Individuals make decisions purposefully, always seeking to choose the option they expect to be most consistent with their personal goals.
3. Incentives matter. The likelihood of people choosing an option increases as personal benefits rise and personal costs decline.
4. Economic reasoning focuses on the impact of marginal changes because it is the marginal benefits and marginal costs that influence choices.
5. Since information is scarce, uncertainty is a fact of life.
6. In addition to their direct impact, economic changes often generate secondary effects.
7. The value of a good or service is subjective and varies with individual preferences and circumstances.
8. The test of an economic theory is its ability to predict and to explain events in the real world.

- Economic science is positive; it attempts to explain the actual consequences of economic actions. Normative economics goes further, applying value judgments to make suggestions about what "ought to be."
- Microeconomics focuses on narrowly defined units, while macroeconomics is concerned with highly aggregated units. When shifting focus from micro to macro, one must beware of the fallacy of composition: What's good for the individual may not be good for the group as a whole.
- The origin of economics as a science dates to the publication of *The Wealth of Nations* by Adam Smith in 1776. Smith believed a market economy would bring individual self-interest and the public interest into harmony.

CRITICAL ANALYSIS QUESTIONS

1. Indicate how each of the following changes would influence the incentive of a decision maker to undertake the action described.

a. A reduction in the temperature from 80° to 50° on one's decision to go swimming

b. A change in the meeting time of the introductory economics course from 11:00 A.M. to 7:30 A.M. on one's decision to attend the lectures

c. **A** reduction in the number of exam questions that relate directly to the text on the student's decision to read the text

d. An increase in the price of beef on one's decision to buy steak

e. An increase in the rental rates of apartments on one's decision to build additional rental housing units

***2.** "The government should provide such goods as health care, education, and highways because it can provide them for free." Is this statement true or false? Explain your answer.

3. **a.** What method is used to ration goods in a market economy? How does this rationing method influence the incentive of individuals to supply goods, services, and resources to others?

b. How are grades rationed in your economics class? How does this rationing method influence student behavior? Suppose the highest grades were rationed to those whom the teacher liked best. How would this method of rationing influence student behavior?

***4.** In recent years, both the personal exemption and child tax credit have been increased in the United States. According to the basic principles of economics, how will the birthrate be affected by policies that reduce the taxes imposed on those with children?

***5.** "The economic way of thinking stresses that good intentions lead to sound policy." Is this statement true or false? Explain your answer.

6. Self-interest is a powerful motivator. Does this necessarily imply that people are selfish and greedy? Do self-interest and selfishness mean the same thing'?

***7.** Congress and government agencies often make laws to help protect the safety of consumers. New cars, for example. are required to have many safety features before they can be sold in the United States. These rules do indeed provide added safety for buyers, although they also add to the cost of making and price of buying the new vehicles. What secondary effects can you see happening as the result of mandating that automobiles have airbags? What incentives do you see changing for drivers as the result of making cars safer? Do you think the millions of dollars spent by consumers on air bags each year

could be better spent elsewhere to save even more lives?

*8. "Individuals who economize are missing the point of life. Money is not so important that it should rule the way we live." Evaluate this statement.

"9. "Positive economics cannot tell us which agricultural policy is better, so it is useless to policymakers." Evaluate this statement.

*10. "I examined the statistics for our basketball team's wins last year and found that, when the third team played more, the winning margin increased. If the coach played the third team more, we would win by a bigger margin." Evaluate this statement.

11. Which of the following are positive economic statements and which are normative?
   a. The speed limit should be lowered to 55 miles per hour on interstate highways.
   b. Higher gasoline prices cause the quantity of gasoline that consumers buy to increase.
   c. A comparison of costs and benefits should not be used to assess environmental regulations.
   d. Higher taxes on alcohol result in less drinking and driving.

1 "Economics is about trade-offs. If more scarce resources are used to produce one thing, fewer will be available to produce others." Evaluate this statement.

13. Do individuals "economize"? If so, what are they trying to do? Do you economize when you shop at the mall? Why or why not?

*14. Should the United States attempt to reduce air and water pollution to zero? Why or why not?

*Asterisk denotes questions for which answers are given in Appendix B

ADDENDUM

Understanding Graphs

Economists often use graphs to illustrate economic relations. Graphs are like pictures. They are visual aids that can communicate valuable information in a small amount of space. A picture may be worth a thousand words, but only to a person who understands the picture (and the graph).

This addendum illustrates the use of simple graphs as a way to communicate. Many students, particularly those with some mathematics background, are already familiar with this material, and can safely ignore it. This addendum is for those who need to be assured that they can understand graphic illustrations of economic concepts.

The Simple Bar Graph

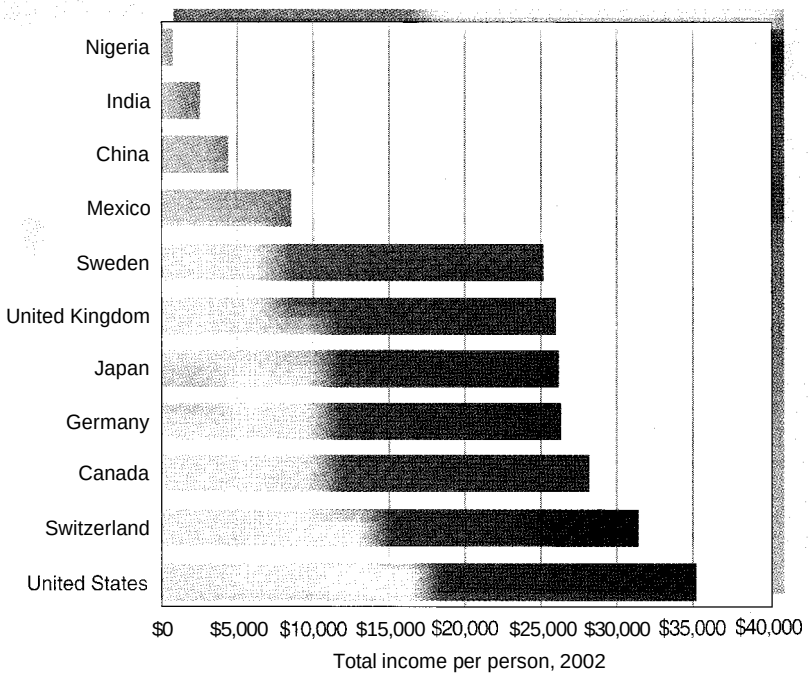
A simple bar graph helps us visualize comparative relationships and understand them better. It is particularly meful for illustrating how an economic indicator varies among countries, across time periods, or under alternative economic conditions.

Exhibit A-1, is a bar graph illustrating economic data. The table in part (a) presents data on the income per person in 2002 for several countries. Part (b) uses a bar graph to illustrate the same data. The horizontal scale of the graph indicates the total income per person. **A** bar is made for indicating the income level (see the dollar scale on the x-axis) of each country. The length of each bar is in pro-

EXHIBIT A-1
International Comparison of Income per Person

COUNTRY	TOTAL INCOME PER PERSON, 2002
United States	$35,060
Switzerland	31,250
Canada	28,070
Germany	26,220
Japan	26,070
United Kingdom	25,870
Sweden	25,080
Mexico	8,540
China	4,390
India	2,570
Nigeria	780

(a)



(b)

portion to the per-person income of the country. Thus, the length of the bars provides a visual illustration of how per capita income varies across the countries. For example, the extremely short bar for Nigeria shows immediately that income per person there is only a small fraction of the comparable income figure for the United States, Japan, Switzerland, and several other countries

Linear Graphic Presentation

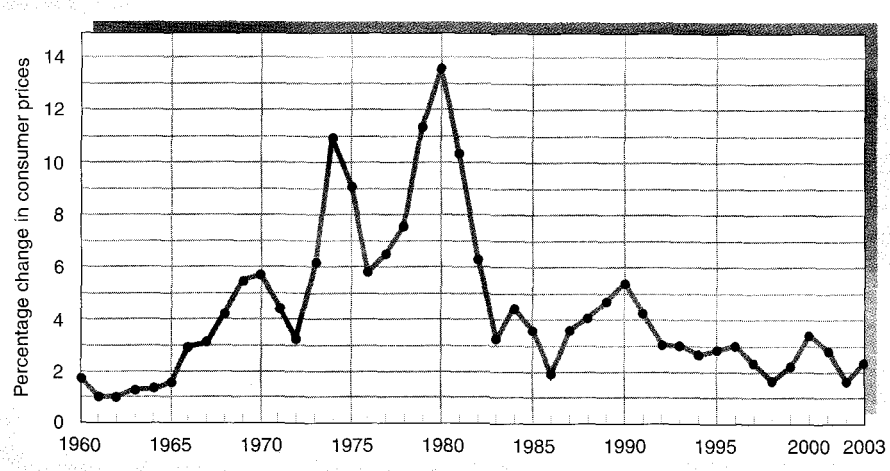
Economists often want to illustrate variations in economic variables with the passage of time. **A** linear graph with time on the horizontal axis and an economic variable on the vertical axis is a useful tool to indicate variations over time. **Exhibit A-2** illustrates a simple linear graph of changes in consumer prices (the inflation rate) in the United States between 1960 and 2003. The table of the exhibit presents data on the percentage change in consumer prices for each year. Beginning with 1960, the horizontal axis indicates the time period (year). The inflation rate is plotted vertically above each year. Of course, the height of the plot (line) indicates the inflation rate during that year. For example, in 1975 the inflation rate was 9.1 percent. This point is plotted at the 9.1 percent vertical distance directly above the year 1975. In 1976 the inflation rate fell to 5.8 percent. Thus, the vertical plot of the 1976 inflation rate is lower than that for 1975. The inflation rate for each year shown in part (a) is plotted at the corresponding height directly above the year in part (b). The linear graph

EXHIBIT A-2
Changes in Level of Prices in United States, 1960–2003

The tabular data (a) of the inflation rate are presented in graphic form in (b).

Year	Percent Change in Consumer Prices	Year	Percent Change in Consumer Prices
1960	1.7	1982	6.2
1961	1.0	1983	3.2
1962	1.0	1984	4.3
1963	1.3	1985	3.6
1964	1.3	1986	1.9
1965	1.6	1987	3.6
1966	2.9	1988	4.1
1967	3.1	1989	4.8
1968	4.2	1990	5.4
1969	5.5	1991	4.2
1970	5.7	1992	3.0
1971	4.4	1993	3.0
1972	3.2	1994	2.6
1973	6.2	1995	2.8
1974	11.0	1996	3.0
1975	9.1	1997	2.3
1976	5.8	1998	1.6
1977	6.5	1999	2.2
1978	7.6	2000	3.4
1979	11.3	2001	2.8
1980	13.5	2002	1.6
1981	10.3	2003	2.3

(a)



(b)

Source: Bureau of Labor Statistics (http://stats.bls.gov/cpihome.htm).

is simply a line connecting the points plotted for each of the years.

The linear graph is a visual aid to understanding what happens to the inflation rate during the period. As the graph shows, the inflation rate rose sharply between 1967 and 1969, 1972 and 1974, and 1978 and 1980. It was substantially higher during the 1970s than in the early 1960s or the mid-1980s and 1990s. Most importantly, the inflation rate has been lower and more stable since 1983 than in the period before. Although the linear graph does not communicate any information not in the table, it does make it easier to see the pattern of the data. Thus, economists often use simple graphics rather than tables to communicate information.

Direct and Inverse Relationships

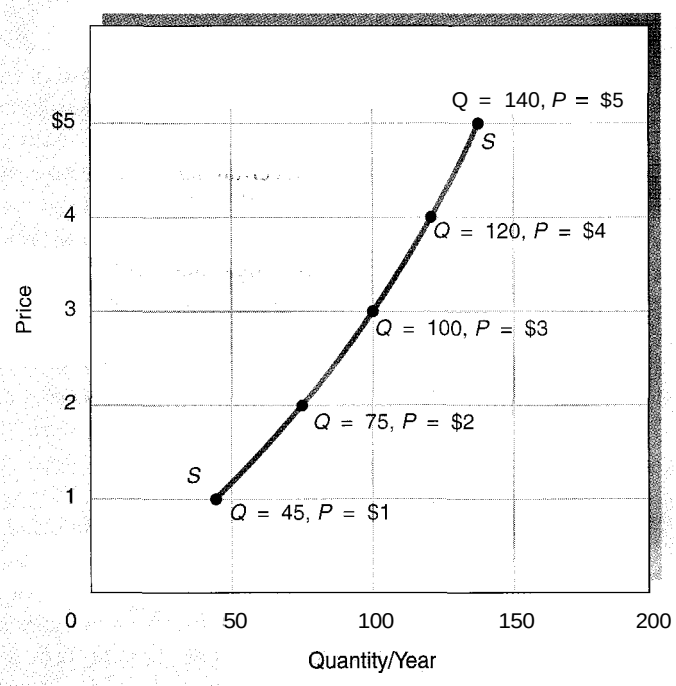
Economic logic often suggests that two variables are linked in a specific way. Suppose an investigation reveals that, other things being constant, farmers supply more wheat as the price of wheat increases. **Exhibit A-3** presents hypothetical data on the relationship between the price of wheat and the quantity supplied by farmers, first in tabular form in part (a) and then as a simple two-dimensional graph in part (b). Suppose we measure the quantity of wheat supplied by farmers on the x-axis (the horizontal axis) and the price of wheat on the y-axis (the vertical axis). Points indicating the value of x (quantity supplied) at alternative values of y (price of wheat) can then be plotted. The line (or curve) linking the points illustrates the relationship between the price of wheat and the amount supplied by farmers.

In the case of price and quantity supplied of wheat, the two variables are directly related. When the y-variable increases, so does the x-variable. When two variables are directly related, the graph illustrating the linkage between the two will slope upward to the right, as in the case of SS in part (b).

Sometimes the x-variable and the y-variable are inversely related. A decline in the y-variable is associated with an increase in the x-variable. Therefore, a curve picturing the inverse relationship between x and y slopes

Price	Amount of Wheat Supplied by Farmers per Year (Millions of Bushels)
$1	45
2	75
3	100
4	120
5	140

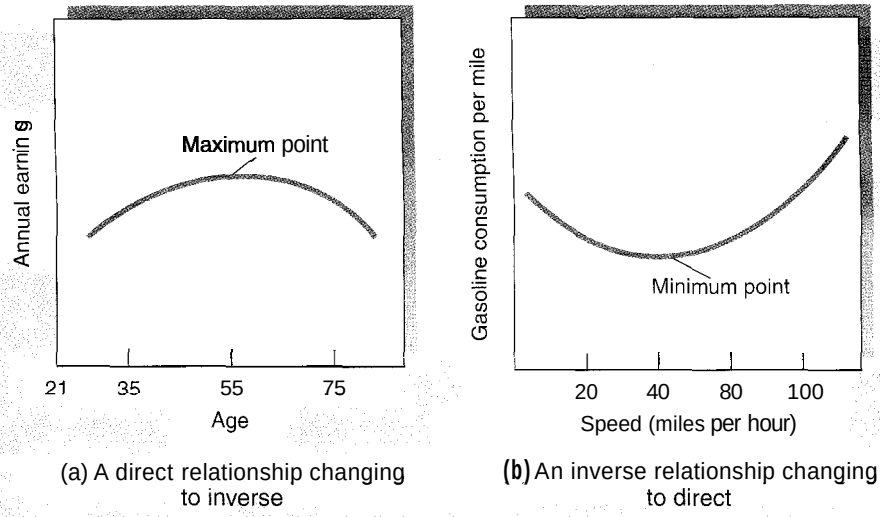
(a)



(b)

EXHIBIT A-3
Direct Relationship Between Variables

As the table (a) indicates, farmers are willing to supply more wheat at a higher price. Thus, there is a direct relation between the price of wheat and the quantity supplied. When the *x*- and y-variables are directly related, a curve mapping the relationship between the two will slope upward to the right like SS.



(a) A direct relationship changing to inverse

(b) An inverse relationship changing to direct

EXHIBIT A-5
Complex Relationships Between Variables

At first, an increase in age (and work experience) leads to a higher income, but later earnings decline as the worker approaches retirement (a). Thus, age and annual income are initially directly related, but at approximately age 55 an inverse relationship emerges. Part (b) illustrates the relationship between travel speed and gasoline consumption per mile. Initially, gasoline consumption per mile declines as speed increases (an inverse relationship), but as speed increases above 40 mph, gasoline consumption per mile increases with the speed of travel (direct relationship).

and more gasoline per mile is required to maintain the additional speed. At very high speeds, gasoline consumption per mile increases substantially with speed of travel. Thus, gasoline consumption per mile reaches a minimum, and a direct relationship between the *x*- and y-variables describes the relationship beyond that point (40 mph).

Slope of a Straight Line

In economics, we are often interested in how much the *y*-variable changes in response to a change in the x-variable. The slope of the line or curve reveals this information. Mathematically, the slope of a line or curve is equal to the change in the y-variable divided by the change in the *x*-variable.

Exhibit A-6 illustrates the calculation of the slope for a straight line. The exhibit shows how the daily earnings (*y*-variable) of a worker change with hours worked (the *x*-variable). The wage rate of the worker is $10 per hour, so when 1 hour is worked, earnings are equal to $10. For 2 hours of work, earnings jump to $20, and so on. A 1-hour change in hours worked leads to a $10 change in earnings. Thus, the

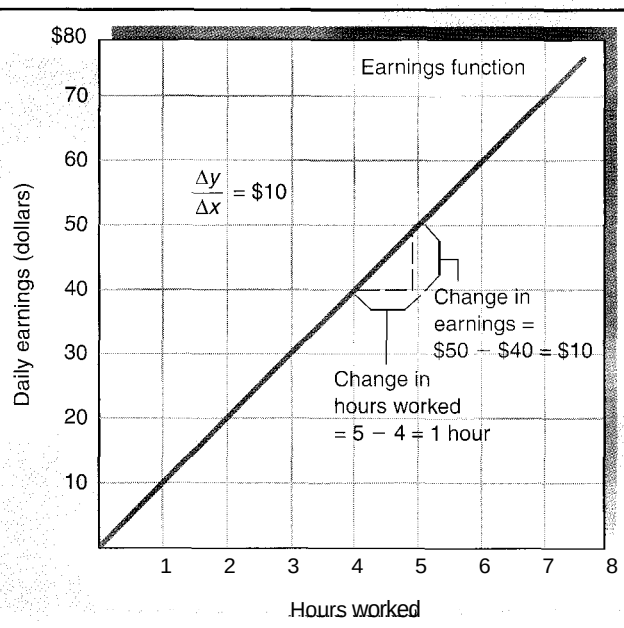
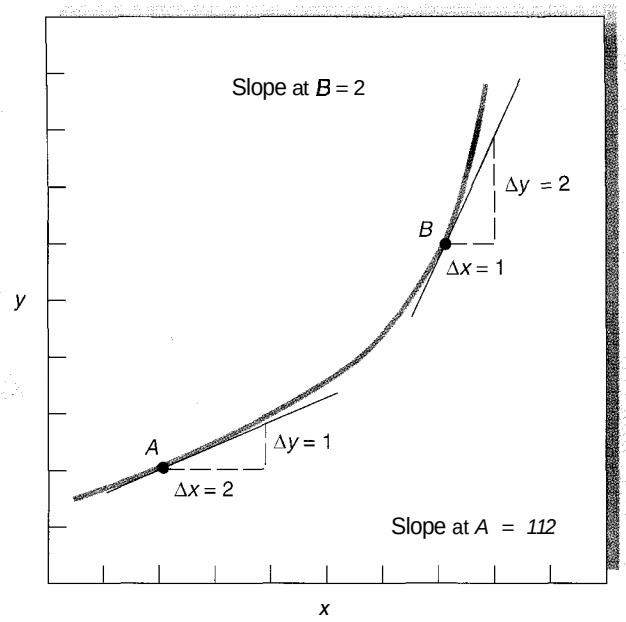


EXHIBIT A-6
Slope of a Straight Line

The slope of a line is eaual to the change in *y* divided by the change in *x*. The line opposite illustrates the case in which daily earnings increase by $10 per hour worked. Thus, the slope of the earnings function is 10 ($10 ÷ 1 hr). For a straight line, the slope is constant at each point on the line.

EXHIBIT A-7
Slope of a Nonlinear Curve

The slope of a curve at any point is equal to the slope of the straight line tangent to the curve at the point. As the lines tangent to the curve at points A and 6 illustrate, the slope of a curve will change from point to point along the curve.



slope of the line *(AylAx)* is equal to 10. (The symbol A means "change in.") In the case of a straight line, the change in *y,* per unit change in *x,* is equal for all points on the line. Thus, the slope of a straight line is constant for all points along the line. Exhibit A-6 illustrates a case in which a direct relation exists between the *X*- and y-variables. For an inverse relation, the y-variable decreases as the x-variable increases. So, when *X* and y are inversely related, the slope of the line will be negative.

Slope of a Curve

In contrast with a straight line, the slope of a curve is different at each point along the curve. The slope of a curve at a specific point is equal to the slope of a line tangent to the curve at the point, meaning a line that just touches the curve.

Exhibit A-7 illustrates how the slope of a curve at a specific point is determined. First, consider the slope of the curve at point A. A line tangent to the curve at point **A** indicates that *y* changes by one unit when *X* changes by two units at point A. Thus, the slope ***(AylAx)*** of the curve at **A** is equal to 0.5.

Now consider the slope of the curve at point *B*. The line tangent to the curve at *B* indicates that y changes by two units for each one-unit change in *X* at point *B*. Thus, at *B* the slope ($\Delta y/\Delta x$) is equal to 2. At point *B*, a change in the x-variable leads to a much larger change in y than it does at point **A.** The greater slope of the curve at *B* reflects this greater change in y per unit change in *x* at *B* relative to **A.**

Graphs Are Not a Substitute for Economic Thinking

By now you should have a fairly good understanding of how to read a graph. If you still feel uncomfortable with graphs, try drawing (graphing) the relationship between several things with which you are familiar. If you work, try graphing the relationship between your hours worked *(X*-axis) and your weekly earnings (y-axis). Exhibit A-3 could guide you with this exercise. Can you graph the relationship between the price of gasoline and your expenditures on gasoline? Graphing these simple relationships will give you greater confidence in your ability to grasp more complex economic relationships presented in graphs.

This text uses only simple graphs. Thus, there is no reason for you to be intimidated. Graphs look much more complex than they really are. In fact, they are nothing more than a simple device to communicate information quickly and concisely. Nothing can be communicated with a graph that cannot be communicated verbally.

Most important, graphs are not a substitute for economic thinking. Although a graph may illustrate that two variables are related, it tells us nothing about the cause-and-effect relationship between the variables. To determine probable cause and effect, we must rely on economic theory. Thus, the economic way of thinking, not graphs, is the power station of economic analysis.

Some Tools of the Economist

CHAPTER 2

Chapter Focus

- **What is opportunity cost? Why do economists place so much emphasis on it?**
- Why do people engage in exchange?

  How does private ownership affect the use of resources? Will private owners pay any attention to the desires of others?
- What does a production possibilities curve demonstrate?
- What are the sources of gains from trade? How does trade influence our **modern living standards?**
- What are the two major methods of economic organization? How do **they differ?**

The key insight of Adam Smith's* Wealth of Nations *is misleadingly simple: if an exchange between two parties is voluntary, it will not take place unless both believe they will benefit from it. Most economic fallacies derive from the neglect of this simple insight, from the tendency to assume that there is a fixed pie, that one party can gain only at the expense of another.

—Milton and Rose Friedman[1]

Milton and Rose Friedman, *Free to Choose* (Harcourt Brace, 1990), 13.

In the preceding chapter, you were introduced to the economic way of thinking. We will now begin to apply that approach. This chapter focuses on five topics: opportunity cost, trade, property rights, the potential output level of an economy, and the creation of wealth. These seemingly diverse topics are in fact highly interrelated. For example, the opportunity cost of goods determines which ones an individual or a nation should produce and which should be acquired through trade. In turn, the ways in which trade and property rights are structured influence the amount of output and wealth an economy can create. These tools of economics are important for answering the basic economic questions: what to produce, how to produce it, and for whom it will be produced. We will begin by first explaining in more detail what opportunity cost is.

WHAT SHALL WE GIVE UP?

Because of scarcity we can't have everything we want. As a result, we constantly face choices that involve trade-offs between our competing desires. Most of us would like to have more time for leisure, recreation, vacations, hobbies, education, and skill development. We would also like to have more wealth, a larger savings account, and more consumable goods. However, all these things are scarce, in the sense that they are limited. Our efforts to get more of one will conflict with our efforts to get more of others.

Opportunity Cost

An unpleasant fact of economics is that the choice to do one thing is, at the same time, a choice *not* to do something else. Your choice to spend time reading this book is a choice not to spend the time playing video games, listening to a math lecture, or going to a party. These things must be given up because you decided to read this book instead. As we indicated in Chapter 1, the highest valued alternative sacrificed in order to choose an option is called the *opportunity cost* of that choice. In economics when we refer to the "cost" of an action, we are referring to its opportunity cost.

Opportunity costs are subjective because they depend upon how the decision maker values his or her options. They are also based on the expectations of the decision maker—what he or she expects the value of the forgone alternatives will be. Because of this, opportunity cost can never be directly measured by someone other than the decision maker. Only the person choosing can know the value of what is given up.[2] This makes it difficult for someone other than the decision maker—including experts and elected officials—to make choices on that person's behalf. Moreover, not only do people differ in the trade-offs they prefer to make, but their preferences also change with time and circumstances. Thus, the decision maker is the only person who can properly evaluate the options and decide which is the best, given his or her preferences and current circumstances.

Monetary costs can be measured objectively in terms of dollars and cents (or yen, lira, and so forth). They also represent an opportunity cost. If you spend $20 on a new CD, you must now forgo the other items you could have purchased with the $20—a new shirt, for example. However, it is important to recognize that monetary costs do not represent the total opportunity cost of an option. The total cost of attending a football game, for example, is the highest valued opportunity lost as a result of both the time you spend at the game and the amount of money you pay for your ticket. In cases like the purchase of a CD, where there is minimal outlay of time, effort, and other resources to make the purchase, the monetary cost will approximate the total cost. Contrast this with a decision to sit on your sofa and listen to your new CD, which involves little or no monetary cost, but has a clear opportunity cost of your time. In this second case, the monetary cost is a poor measure of the total cost.

[2]See James M. Buchanan, *Cost and Choice* (Chicago: Markham, 1969), for a classic work on the relationship between cost and choice.

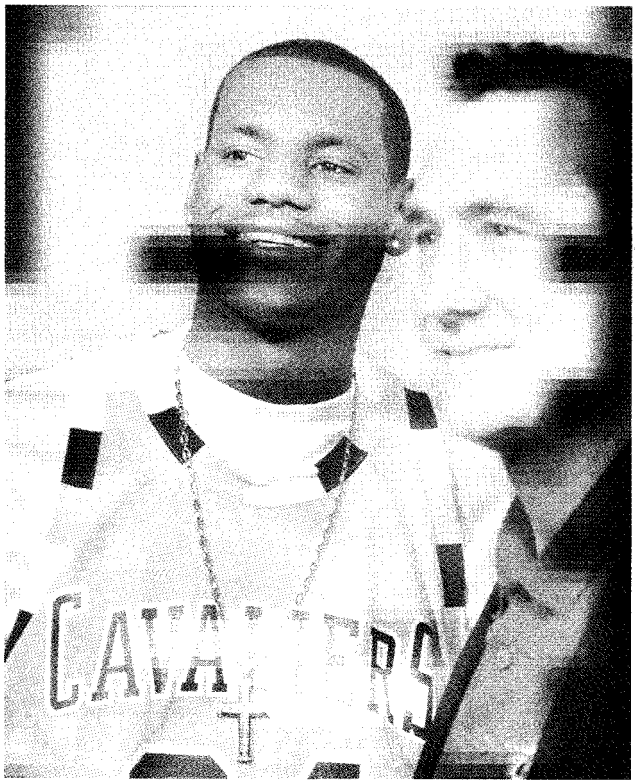
Opportunity Cost and the Real World

Is real-world decision making influenced by opportunity costs? Consider your own decision to attend college. Your opportunity cost of going to college is the value of the next best alternative, which could be measured as the salary you would earn if you had chosen to go directly into full-time work instead. Every year you stay in college, you give up what you could have earned by working that year. Typically, students incur opportunity costs of $80,000 or more in forgone income during their stay in college.

But what if the opportunity cost of attending college changes? How will it affect your decision? Suppose, for example, that you received a job offer today for $250,000 per year as an athlete or an entertainer, but the job would require so much travel that school would be impossible. Would this change in the opportunity cost of going to college affect your choice as to whether to continue in school? It likely would. Going to college would mean you would have to say goodbye to the huge salary you've been offered. (See the accompanying illustration on LeBron James for a good example.) You can clearly tell from this example that the monetary cost of college (tuition, books, and so forth) isn't the only factor influencing your decision. Your opportunity cost plays a part, too.

Even when their parents pay all the monetary expenses of their college education, some students are surprised to learn that they are actually incurring more of the total cost of going to college than their parents. For example, the average monetary cost (tuition, room and board, books, and so forth) for a student attending college is about $10,000 per year ($40,000 over four years). Even if the student's next best alternative were working at a job that paid only $15,000 per year, over four years, that would amount to $60,000 in forgone earnings, So, the total cost of the student's education would be $100,000 ($40,000 in monetary costs paid by the parents and $60,000 in opportunity costs incurred by the student).[3]

Now consider another decision made by college students — whether to attend a particular class meeting. The monetary cost of attending class (bus fare, parking, gasoline costs,



AP PHOTO/MARK DUNCAN

LeBron James (shown here with Cleveland Cavaliers General Manager Jim Paxson) understands opportunity cost. As a high school player, James was already one of the best basketball players in the nation. He had received numerous scholarship offers and was considering attending college at Ohio State, the University of North Carolina, Michigan State, or the University of California. However, after high school graduation, LeBron decided to go directly into the NBA because the opportunity cost of college was simply too high. He was selected as the first pick in the 2003 NBA draft, signing a three-year contract worth almost $13 million, with an option for a fourth year at $5.8 million. Had he decided to go to college instead, James would have incurred an opportunity cost of at least $19 million in forgone income to earn a four-year college degree! Would you have skipped college if your opportunity cost had been that high?

[3] From the standpoint of the family's total economic cost of sending a child to college, some of the monetary costs, such as room and board are not costs of choosing to go to college The cost of living does have to be covered, hut it would be incurred whether or not the student went to college

and so on) remains fairly constant from day to day. Why then do students choose to attend class on some days and not on others? Even though the monetary cost of attending class is fairly constant, a student's opportunity cost can change dramatically from day to day. Some days the next best alternative to attending class may be sleeping in or watching TV. Other days, the opportunity cost may be substantially larger, perhaps the value of attending a big football game, getting an early start on spring break, or having additional study time for a crucial exam in another class. As options like these increase the cost of attending class, more students will decide not to attend.

Failure to consider opportunity cost often leads to unwise decision making. Suppose that your community builds a beautiful new civic center. The mayor, speaking at the dedication ceremony, tells the world that the center will improve the quality of life in your community. People who understand the concept of opportunity cost may question this view. If the center had not been built, the resources might have funded construction of a new hospital, improvements to the educational system, or housing for low-income families. Will the civic center contribute more to the well-being of the people in your community than these other facilities? If so, it was a wise investment. If not, your community will be worse off than it would have been if decision makers had chosen a higher valued project.

TRADE CREATES VALUE

Why do individuals trade with each other, and what is the significance of this exchange? We have learned that value is subjective. It is wrong to assume that a particular good or service has a fixed objective value just because it exists.[4] The value of goods and services generally depends on who uses them, and on circumstances, such as when and where they are used, as well as on the physical characteristics. Some people love onions, whereas others dislike them exceedingly. Thus, when we speak of the "value of an onion," this makes sense only within the context of its value to a specific person. Similarly, to most people an umbrella is more valuable on a rainy day than on a sunny one.

Consider the case of Janet, who loves tomatoes but hates onions, and Brad, who loves onions but hates tomatoes. They go out to dinner together and the waiter brings their salads. Brad turns to Janet and says, "I'll trade you the tomatoes on my salad for the onions on yours." Janet gladly agrees to the exchange. This simple example will help us illustrate two important aspects of voluntary exchange.

OUTSTANDING ECONOMIST

Thomas Sowell (1930–)

Thomas Sowell, a senior fellow at the Hoover Institution, recognizes the critical importance of the institutions—the "rules of the game"—that shape human interactions. His book *Knowledge and Decisions* stresses the role of knowledge in the economy and how different institutional arrangements compare at using scarce information. Sowell is the author of many books and journal articles and writes a nationally syndicated column that appears in more than 150 newspapers. His writings address subjects ranging from race preferences and cultural differences to the origins and ideology of political conflict.

[4]An illuminating discussion of this subject, termed the "physical fallacy," is found in Thomas Sowell, *Knowledge and Decisions* (New York: Basic Books, 1980), 67–72.

1. When individuals engage in a voluntary exchange, both parties are made better *off.* In the above example, Janet has the option of accepting or declining Brad's offer of a trade. If she accepts his offer, she does so *voluntarily.* Janet would agree to this exchange only if she expects to be better off as a result. Because she likes tomatoes better than onions, Janet's enjoyment of her salad will be greater with this trade than without it. On the other side, Brad has voluntarily made this offer of an exchange to Janet because Brad believes he will also be better off as a result of the exchange.

People tend to think of making, building, and creating things as productive activities. Agriculture and manufacturing are like this. They create something genuinely new, something that was not there before. On the other hand, trade—the mere exchange of one thing for another—does not create new material items. You might be tempted to think that if goods are merely being traded, one party will be better off and the other worse off. A closer look at the motivation for trade helps us see through this popular fallacy. Exchange takes place because both parties expect it will make them better off. If they didn't, they wouldn't agree to do it. For example, if Janet liked onions better than tomatoes, she wouldn't have traded with Brad. The fact that she agreed to the trade means she thinks she has something to gain by doing so. Brad thinks the same thing when it comes to his tomatoes. In other words, because their exchange is voluntary, *both* Janet and Brad are made better off. As the chapter-opening quotation illustrates, most errors in economic reasoning happen when we forget that voluntary trades, like the one between Janet and Brad, make both parties better off.

2. By channeling goods and resources to those who value them most, trade creates value and increases the wealth created by a society's resources. Because preferences differ among individuals, the value of an item can vary greatly from one person to another. Therefore, trade can create value by moving goods from those who value them less to those who value them more. The simple exchange between Janet and Brad also illustrates this point. Imagine for a moment that Brad and Janet had never met and instead were both eating their salads alone. Without the ability to engage in this exchange, both would have eaten their salads but not had as much enjoyment from them. When goods are moved to individuals who value them more, the total value created by a society's limited resources is increased. The same two salads create more value when the trade occurs than when it doesn't.

It is easy to think of material things as wealth, but material things are not wealth until they are in the hands of someone who values them. A highly technical book on electronics that is of no value to an art collector may be worth several hundred dollars to an engineer. Similarly, a painting that is unappreciated by an engineer may be of great value to an art

ECONOMICS AT THE MOVIES

20TH CENTURY FOX/THE KOBAL COLLECTION

Wall Street (1987)

Michael Douglas won an Oscar for his performance in *Wall Street,* but he gets a failing grade for his understanding of economics. In response to a question Charlie Sheen poses to him about how much money is "enough," Michael Douglas replies: "It's not a question of enough, pal. It's a zero-sum game. Somebody wins; somebody loses. Money itself isn't lost or gained, it's simply transferred from one person to another." This is false. In the real world, voluntary trade is a positive-sum game, meaning that wealth is created, and both parties gain. It is not a zero-sum game, where the gains to one person result in losses to another.

collector. Therefore, a voluntary exchange that moves the electronics book to the engineer and the painting to the art collector will increase the value of both goods. By channeling goods and resources toward those who value them most, trade creates wealth for both the trading partners and for the nation.

Transaction Costs—A Barrier to Trade

How many times have you been sitting home late at night, hungry, wishing you had a meal from your favorite fast-food restaurant? You would gladly pay the $4 price for the value meal you have in mind, but you feel it is just not worth the time and effort to get dressed and make that drive. The costs of the time, effort, and other resources necessary to search out, negotiate, and conclude an exchange are called **transaction costs.** High transaction costs can be a barrier to potentially productive exchange.

Transaction costs
The time, effort, and other resources needed to search out, negotiate, and complete an exchange.

Because of transaction costs, we should not expect all potentially valuable trades to take place, any more than we expect all useful knowledge to be learned, all safety measures to be taken, or all potential "A" grades to be earned. Frequent fliers know that if they never miss a flight, they are probably spending too much time waiting in airports. Similarly, the seller of a car, a house, or a ballet ticket knows that finding the one person in the world willing to pay the most money for the good is not worth the enormous effort it would take to find him or her. Information is costly. That is one reason that perfection in exchange, as in most things we do, is seldom worth achieving.

The Internet has significantly lowered transaction costs. The auction Web site eBay enables sellers to reach millions of potential buyers with little effort and few costs. Buyers can easily search eBay for items they want to buy, even if the items are located halfway around the world. Other Web sites, such as MySimon and Pricescan, scour online shopping sites for the lowest prices so buyers don't have to. Consumers can also readily find detailed information about products on any number of sites. Amazon.com posts prices, product information provided by manufacturers, and reviews from other buyers. By reducing transaction costs, the Internet creates value and wealth. It expands the number of trades that are made, and makes it faster and easier to make them.

The Middleman as a Cost Reducer

Because it is costly for buyers and sellers to find each other and to negotiate the exchange, an entrepreneurial opportunity exists for people to become **middlemen**. Middlemen provide buyers and sellers information at a lower cost and arrange trades between them. Many people think middlemen just add to the buyer's expense without performing a useful function. However, because of transaction costs, without middlemen, many trades would never happen (nor would the gains from them be realized). An auto dealer, for example, is a middleman. An auto dealer helps both the manufacturer and the buyer. The dealer helps buyers by maintaining an inventory of vehicles for them to choose from. Knowledgeable salespeople hired by the dealer help car shoppers quickly learn about the vehicles they're interested in and the pros and cons of each. Car buyers also like to know that a local dealer will honor the manufacturer's warranty and provide parts and service for the car. The dealer helps manufacturers by handling tasks like these so they can concentrate on designing and making better cars.

Middleman
A person who buys and sells goods or services or arranges trades. A middleman reduces transaction costs.

Grocers are also middlemen. Each of us could deal with farmers directly to buy our food—probably at a lower monetary cost. But that would have a high opportunity cost. Finding and dealing with different farmers for every product we wanted to buy would take a lot of time. Alternatively, we could form consumer cooperatives, banding together to eliminate the middleman, using our own warehouses and our own volunteer labor to order, receive, display, distribute, and collect payment for the food. In fact, some cooperatives like this do exist. But most people prefer instead to pay a grocer to provide all of the goods they want rather than trying to trade with different farmers.

Stockbrokers, realtors, publishers, and merchants of all sorts are other kinds of middlemen. For a fee, they reduce transaction costs for both buyers and sellers. By making exchanges cheaper and more convenient, middlemen cause more efficient trades to happen. In so doing, they themselves create value.

THE IMPORTANCE OF PROPERTY RIGHTS

Private ownership provides people with a strong incentive to take care of things and develop resources in ways that are highly valued by others.

Private Ownership

The buyer of an apple, a CD, a television set, or an automobile generally takes the item home. The buyer of a steamship or an office building, though, may never touch it. When exchange occurs, it's really the **property rights** of the item that change hands.

Private-property rights involve three things:

(1) the right to exclusive use of the property (that is, the owner has sole possession, control, and use of the property—including the right to exclude others);

(2) legal protection against invasion from other individuals who would seek to use or abuse the property without the owner's permission; and

(3) the right to transfer, sell, exchange, or mortgage the property.

Property rights
The rights to use, control, and obtain the benefits from a good or service.

Private-property rights
Property rights that are exclusively held by an owner and protected against invasion by others. Private property can be transferred, sold, or mortgaged at the owner's discretion.

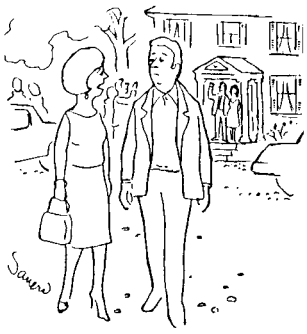
Private owners can do anything they want with their property as long as they do not use it in a manner that invades or infringes on the rights of another. For example, I cannot throw the hammer that I own through the television set that you own. If I did, I would be violating your property right to your television. The same is true if I operate a factory spewing out pollution harming you or your land.[5] Because an owner has the right to control the use of property, the owner also must accept responsibility for the outcomes of that control.

In contrast to private ownership, common-property ownership occurs when multiple people simultaneously have or claim ownership rights to a good or resource. None of the common owners can prevent the others from using or damaging the property. Most beaches, lakes, and parks are examples of commonly owned property. The distinction between private- and common-property ownership is important because common ownership does not create the same powerful incentives as private ownership. Economists are fond of saying that when everybody owns something, nobody owns it.

Clearly defined and enforced private-property rights are a key to economic progress because of the powerful incentive effects that follow from private ownership of goods and resources. The following four incentives are particularly important:

1. Private owners can gain by employing their resources in ways that are beneficial to others, and they bear the opportunity cost of ignoring the wishes of others. Realtors often advise home owners to use neutral colors for countertops and walls in their house because they will improve the resale value of the home. As a private owner you could install bright green fixtures and paint your walls deep purple, but you will bear the cost (in terms of a lower selling price) of ignoring the wishes of others who might want to buy your house later. On the other hand, by fixing up a house and doing things to it that others find beneficial, you can reap the benefit of a higher selling price. Similarly, you could spray paint orange designs all over the outside of your brand-new car. but private ownership gives you an incentive not to do so because the resale value of the car depends on the value that *others* place on it.

Consider a parcel of undeveloped privately owned land near a university. The private owner of the land can do many things with it. For example, she could leave it undeveloped. turn it into a metered parking lot, erect a restaurant, or build rental housing. Will the wishes and desires of the nearby students be reflected in her choice, even though they are not the owners of the property? Yes. Whichever use is more highly valued by potential customers will earn her the highest investment return. If housing is relatively hard to find but there are plenty of other restaurants, the profitability of using her land for housing will be higher than



"Their house looks so nice They must be getting ready to sell it."

FROM THE *WALL STREET JOURNAL*—PERMISSION, CARTOON FEATURES SYNDICATE

A private owner has a strong incentive to do things with his or her property that increase its value to others.

[5]For a detailed explanation of how property rights protect the environment, with several real-world examples, see Roger E. Meiners and Bruce Yandle, *The Common Law: How It Protects the Environment* (Bozeman. Mont.: PERC, 1998), available online at http://www.perc.org.

the profitability of using it for a restaurant. Private ownership gives her a strong incentive to use her property in a way that will also fulfill the wishes of others. If she decides to leave the property undeveloped instead of erecting housing that would benefit the students, she will bear the opportunity cost of forgone rental income from the property.

As a second example, consider the owner of an apartment complex near your campus. The owner may not care much for swimming pools, workout facilities, study desks, washers and dryers, or green areas. Nonetheless, private ownership provides the owner with a strong incentive to provide these items if students and other potential customers value them more than it costs to provide them. Why? Because tenants will be willing to pay higher rents to live in a complex with amenities that they value. The owners of rental property can profit by providing an additional amenity that tenants value as long as the tenants are willing to pay enough additional rent to cover the cost of providing it. Because renters differ in their preferences and willingness to pay for amenities, some will prefer to live in less expensive apartments with fewer amenities, while others will prefer to live in more expensive apartments with a greater range of amenities. By choosing among potential apartment complexes, renters are able to buy as few or as many of these amenities as they wish.

2. Private owners have a strong incentive to care for and properly manage what they own. Will Ed regularly change the oil in his car? Will he see to it that the seats don't get torn? Probably so, since being careless about these things would reduce the car's value, both to him and to any future owner. The car and its value—the sale price if he sells it—belong just to Ed, so he would bear the burden of a decline in the car's value if the oil ran low and ruined the engine, or if the seats were torn. Similarly, he would capture the value of an expenditure that improved the car, like a new paint job. As the owner, Ed has both the authority and the incentive to protect the car against harm or neglect and even to enhance its value. Private-property rights give owners a strong incentive for good stewardship.

Do you take equally good care not to damage an apartment you rent as you would your own house? If you share an apartment with several roommates, are the common areas of the apartment (such as the kitchen and living room) as neatly kept as the bedrooms? Based on economic theory, we guess that the answer to both of these questions is probably "No."

In 1998, the student government association at Berry College in Georgia purchased 20 bicycles to be placed around campus for everyone's use.[6] These $200 Schwinn Cruiser bicycles were painted red and were marked with a plate reading "Berry Bike." The bikes were available on a first-come, first-served basis, and students were encouraged to take them whenever they needed them and leave them anywhere on campus for others to use when they were finished. What do you think happened to these bikes? Within two months,

When apartments and other investment properties are owned privately, the owner has a strong incentive to provide amenities that others value highly relative to their cost.

© INC. MEDIOIMAGES/ NDEX STOCK IMAGERY

© JEFF GREENBERG/PHOTOEDIT

[6]Daniel L Alban and E **Frank** Stephenson, "The 'Berry Bikes' A Lesson in Private Pioperty," *Ideas on Liberty* 49, no 10 (October 1999) 8–9

most of these top-quality bikes were severely damaged or lost. The campus newspaper reported on the "mangled corpses of twisted red metal that lie about campus." Over the summer break the student government replaced or fixed the bikes, but despite its pleas to "treat the bikes as if they were your own property," the same thing happened the following fall precisely because the bikes weren't the students' own property. It wasn't that the students at Berry College were inherently destructive; after all, there were no problems on campus with privately owned bikes being lost or abused during this time. It was a matter of the different incentives they faced. The student government association eventually abandoned the program and began leasing the remaining bikes to individual students instead. As you can see, there is no denying the strong incentive that private ownership creates for owners to care for their property (or the lack of incentive when private ownership is not clearly defined and enforced).

The incentive for owners to care for and properly manage their property is strong. The owner of a hotel doesn't want to neglect fixing electrical or plumbing problems if it means fewer repair costs due to electrical fires or water leaks in the future. The owner knows travelers aren't going to want to stay in a charred or water-damaged hotel. Poor management will reduce the hotel's value and the owner's personal wealth. This gives the owner an incentive to manage the asset properly.

3. Private owners have an incentive to conserve for the future—particularly if the property is expected to increase in value. People have a much stronger incentive to conserve privately owned property than they do commonly owned property. For example, when Steven was in college, the general rule among his roommates was that any food or drink in the house was common property—open game for the hungry or thirsty mouth of anyone who stumbled across it. There was never a reason for Steven to conserve food or drinks in the house because it would be quickly consumed by a roommate coming in later that night. When Steven first started living alone, he noticed a dramatic change in his behavior. When he ordered a pizza, he would save some for the next day's lunch rather than eating it all that night. Steven began counting his drinks before he had one to make sure there were enough left for the next day. When Steven was the sole owner, he began delaying his current consumption to conserve for the future because he was the one, not his roommates, who reaped the benefit from his conservation.

Similarly, when more than one individual has the right to drill oil from an underground pool of oil, each has an incentive to extract as much as possible, as quickly as possible. Any oil conserved for the future will probably be taken by someone else. In contrast, when only one owner has the right to drill, the oil will be extracted more slowly. The same applies to the common-property problems involved in overfishing of the sea compared with fisheries that use privately owned ponds.

COURTESY OF BERRY COLLEGE, MOUNT BERRY GA

Without clearly defined private-property rights, there is less of an incentive to take proper care of things—as the student government administration at Berry College found out when it provided common-property bikes to be used around campus.

Someone who owns land, a house, or a factory, has a strong incentive to bear costs now, if necessary, to preserve the asset's value for the future. The owner's wealth is tied up in the value of the property, which reflects nothing more than the net benefits that will be available to a future owner. Thus, the wealth of private owners is dependent on their willingness and ability to look ahead, maintain, and conserve those things that will be more highly valued in the future. This is why private ownership is particularly important for the optimal conservation of natural resources.

4. Private owners have an incentive to lower the chance that their property will cause damage to the property of others. Private ownership links responsibility with the right of control. Private owners can be held accountable for damage done to others through the misuse of their property. A car owner has a right to drive his car, but will be held accountable if the brakes aren't maintained and the car damages someone else's property. Similarly, a chemical company has control over its products, but, exactly for that reason, it is legally liable for damages if it mishandles the chemicals. Courts of law recognize and enforce the authority granted by ownership, but they also enforce the responsibility that goes with that authority. Because private-property owners can be held accountable for damages they cause, they have an incentive to use their property responsibly and take steps to reduce the likelihood of harm to others. A property owner, for example, has an incentive to cut down a dying tree before it falls into a neighbor's house and to leash or restrain his or her dog if it's likely to bite others.

Private Ownership and Markets

Private ownership and competitive markets provide the foundation for cooperative behavior among individuals. When private-property rights are protected and enforced, the permission of the owner must be sought before anyone else can use the property. Put another way, if you want to use a good or resource, you must either buy or lease it from the owner. This means that each of us must face the cost of using scarce resources. Furthermore, market prices give private owners a strong incentive to consider the desires of others and use their resources in ways others value.

F. A. Hayek, the winner of the 1974 Nobel Prize in economics, used the expression "the extended order" to refer to the tendency for markets to lead perfect strangers from different backgrounds around the world to cooperate with one another. Let's go back to the example of the property owner who has the choice of leaving her land idle or building housing to benefit students. The landowner might not know any students in her town nor particularly care about providing them housing. However, because she is motivated by market prices, she might build an apartment complex and eventually do business with a lot of students she never intended to get to know. In the process, she will purchase materials, goods, and services produced by other strangers.

Things are different in countries that don't recognize private-ownership rights or enforce them. In these countries, whoever has the political power or authority can simply seize property from whomever might have it without compensating them. In his book *The Mystery of Capital,* economist Hernando DeSoto argues that the lack of well-defined and enforced property rights explains why some underdeveloped countries (despite being market based) have made little economic progress. DeSoto points out that in many of these nations, generations of people have squatted on the land without any legal deed giving them formal ownership. The problem is these squatters cannot borrow against the land to generate capital because they don't have a deed to it, nor can they prevent someone else from arbitrarily taking the land away from them.

Private ownership and markets can also play an important role in environmental protection and natural-resource conservation. Ocean fishing rights, tradable rights to pollute, and private ownership of endangered species are just some examples. The accompanying Applications in Economics feature, "Protecting Endangered Species and the Environment with Private-Property Rights," explores some of these issues.

APPLICATIONS IN ECONOMICS

Protecting Endangered Species and the Environment with Private-Property Rights

Column 1	Column 2
cows	African Rhinos
Pigs	Bald Eagles
Chickens	Spotted Owls
Dogs	American Bison
Cats	African Elephants

Compare the two columns of animals above. The animals listed in column 2 are endangered species, whereas those in column 1 are not. Why the difference? The answer may surprise you—all of the animals listed in column 1 can be privately owned, whereas those in column 2 generally cannot. In this chapter you have learned about the powerful incentives for careful management and conservation created by private-property rights. This application considers how the power of these incentives is being harnessed to protect endangered species and the environment.

What do you think would happen to the total population of cows if people wanted less beef? Beef prices would fall, and the incentive for individuals to dedicate land and other resources to raising cattle would fall. It is precisely the market demand for beef that *creates* the incentive for suppliers to maintain herds of cattle and to protect them from harm.

In some ways, the rhinoceros is similar to a cow. It is large and rather unpredictable—a rhino, like a large bull in a cattle herd, may charge if disturbed. And at *3,000* pounds, a charging rhino can be very dangerous. Also like cattle, rhinos can be valuable, and they are significantly rarer. A single horn from a black rhino can sell for as much as $30,000. That makes it a favorite target of poachers—people who hunt illegally.

Rhinos are very different from cattle in one important respect: In most of Africa where they naturally range, the rhinoceros cannot be privately owned or sold by anyone who might protect them. In those areas, poachers may be assisted by local people eager to see fewer rhino present because they make life risky for humans and compete for food and water. Under these circumstances, the rhino is in danger of becoming extinct.

One reaction to this problem is to outlaw rhinoceros hunting and to forbid the sale of any rhino parts. That happened in 1977, when many nations signed an international treaty outlawing sales of black rhino products. Nearly 20 years later, however, in 1994, the black rhino was closer to extinction than ever before. According to South African economist Michael 't Sas Rolfes, the trade ban "has not had a discernible effect on rhino numbers and does not seem to have stopped the trade in rhino horn. If anything, the . . . listings led to a sharp increase in the black market price of rhino horn, which simply fueled further poaching and encouraged speculative stockpiling of horn." Incentives for poachers and local people had not changed, and between 1970 and 1994, black rhinos suffered a 95 percent decline in Africa.[1]

Then a very different strategy for the black rhino emerged in the southern African nation of Zimbabwe. Although rhinos cannot be privately owned there, landowners can fence and manage the game animals on their property. Many of the remaining black rhinos were relocated to private land in the early 1990s. Because they could profit from protecting the big animals, some ranchers shifted their operations from cattle to wildlife protection, ecotourism, and hunting. Often, they combined several ranches into one conservancy, since some wild animals, including rhinos, range over a large area and are difficult to fence in. Revenues from the conservancies come both from hunting many big game animals—not rhinos, however—and from nonconsumptive uses of wildlife, such as photo safaris.

In 1997, a stay at the Barberton Lodge in the Bubiana Conservancy in Zimbabwe cost about $160 per night for a photo safari. Other Bubiana properties charged between $500 and $1,000 per day for a hunting safari, on top of any trophy fees (such as *$3,000* or more for a leopard). No elephants were hunted in the conservancies, but elsewhere in Zimbabwe, hunters paid up to $36,000 for a three-week chance at tracking and killing an elephant. Although no black rhino could be hunted in Zimbabwe, estimates of what the fee would be, if hunting were allowed, reached $250,000. Because of the low overhead and high return, hunting is the reason that, even without the hunting of elephant and rhino. all of the ranchers in the Bubiana partnership have been able to turn a profit from their wildlife operations.

Cattle were introduced to the area in the 1950s and 1960s. When they were removed from some of the conservancies to make way for rhinos, native grasses and shrubs came back in strength, and so did other big game and other forms of wildlife. By the turn of the century, not a single animal had been poached on these private conservancies, and rhino populations climbed. A similar success story is found in South Africa for the African white rhino.[2]

In Africa, elephant numbers also show the value of property rights and market tools for conservation. Zimbabwe and Botswana have for years allowed landowners and local

[1]See Michael De Alessi, *Private Conservation and Black Rhinos in Zimbabwe The Savé Valley and Bubiana ConTenancies*, available online at http //www pnvatecon servation org/pubs/studies/Rhino PDF The rhino story *is* one of many case studies available from the Center for Prirate Conservation at www pnvateconservation org

[2]See "The Rhinos Are Baaack!" *Smithsonian Magazine* (March 2001): 76–86..

(continued)

APPLICATIONS IN ECONOMICS

tribes to benefit financially from the presence of elephants. They have allowed domestic trade in ivory. Other countries, such as Kenya, have banned the ivory trade and have forbidden such gains to landowners from the elephants, instead making their government responsible for protecting them. From 1979 to 1989, property rights and market conservation helped push elephant numbers from 50,000 up to 94,000 in Zimbabwe and Botswana, while Kenya's elephant population fell from 65,000 to 19,000. The trend did not stop there. From 1989 to 1995, elephant populations in Zimbabwe and Botswana rose by about 15 percent, while the rest of Africa lost about 20 percent of their elephants.

The story is not entirely a happy one, however. Success is crucially dependent on protection of property rights to wildlife, and property rights are not always secure in southern Africa. Zimbabwe is currently in the midst of a virtual civil war. Terry Anderson, a resource economist and hunter, writes that the Bubiana Conservancy has experienced severe devastation: "20,000 trees have been felled, 22 buildings razed, staff assaulted, and 50 percent of the wildlife killed." The value of the wild animals lost is estimated at $1.5 million.[3]

In South Africa, however, property rights to wildlife are much stronger. Once a landowner "game-fences" his or her property (builds a tall fence using 12 strands of high-tensile wire), wild animals become the owner's property. These owners "have an incentive to manage the wild animals as they might their cattle, paying close attention to carrying capacity, habitat, and water," writes Anderson. He observes that this is an improvement over the United States, where landowners "cannot capture the benefits of wildlife and consider it more of a nuisance than an asset." In sum, says Anderson, "My experiences in Africa show that private ownership and a focus on rewarding good stewardship are the key to protecting wildlife habitat and wildlife populations."

[3]Terry L. Anderson, "My Love Affair with Africa," *PERC Reports*, June 2004, available at http://perc.org/publications/percreports/june2004/africa.php?s=2. Much more is available on this Web site (http://www.perc.org) on ways that property rights can preserve and enhance environmental quality.

PRODUCTION POSSIBILITIES CURVE

Production possibilities curve A curve that outlines all possible combinations of total output that could be produced, assuming (1) a fixed amount of productive resources, (2) a given amount of technical knowledge, and (3) full and efficient use of those resources. The slope of the curve indicates the amount of one product that must be given up to produce more of the other.

People try to get the most from their limited resources by making purposeful choices and engaging in economizing behavior. This can be illustrated using a conceptual tool called the **production possibilities curve.** The production possibilities curve shows the maximum amount of any two products that can be produced from a fixed set of resources, and the possible trade-offs in production between them. Admittedly, this is an oversimplified model because economies obviously produce more than just two products. Nonetheless, the production possibilities curve can help us understand a number of important economic ideas.

Exhibit 1 illustrates the production possibilities curve for Susan, an intelligent economics major. This curve indicates the combinations of English and economics grades that she thinks she can earn if she spends a total of 10 hours per week studying for the two subjects. Currently, she is choosing to study the material in each course that she expects will help her grade the most, for the time spent, and she is allocating 5 hours of study time to each course. She expects that this amount of time, carefully spent on each course, will allow her to earn a B grade in both, indicated at point *T*. But if she were to take some time away from studying one of the two subjects and spend it studying the other, she could raise her grade in the course receiving more study time. However, it would come at the cost of a lower grade in the course she spends less time studying for. If she were to move to point S by spending more hours on economics and fewer on English, for example, her expected economics grade would rise, while her expected English grade would fall. This illustrates the first important concept shown in the production possibilities framework—the idea of trade-offs in the use of scarce resources. Whenever more of one thing is produced, there is an opportunity cost in terms of something else that now must be forgone.

You might notice that Susan's production possibilities curve indicates that the additional study time required to raise her economics grade by one letter, from a B to an A (moving from point *T* to point S), would require giving up two letter grades in her English class, not just one, reducing her English grade from a B to a D. If, alternatively, Susan

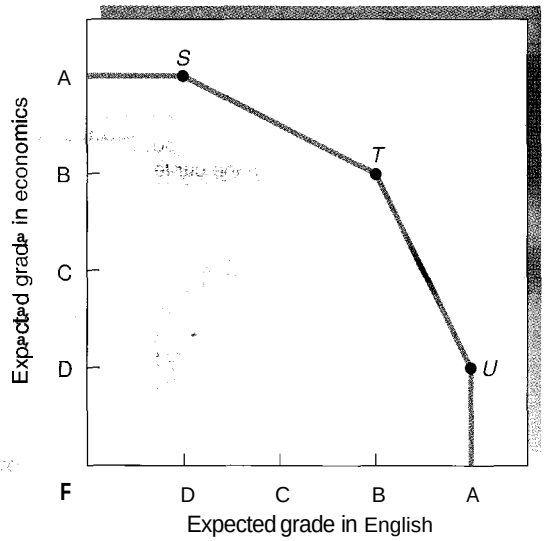


EXHIBIT 1
Production Possibilities Curve for Susan's Grades in English and Economics

The production possibilities for Susan, in terms of grades, are illustrated for 10 hours of total study time. If Susan studied 10 hours per week in these two classes, she could attain a D in English and an A in economics (point S), a B in English and a B in economics (point *T*), or a D in economics and an A in English (point *U*).

were to move from point *T* to point *U*, the opposite would be true—she would improve her English grade by one letter at the expense of two letter grades in economics. You can understand this by thinking about your own studying behavior. When you have only a limited amount of time to study a subject, you begin by studying the most important (grade-increasing) material first. As you spend additional time on that subject, you begin studying topics that are of decreasing importance for your grade. Thus, adding an hour of study time to the subject Susan studies least will have a larger impact on her grade than will taking away an hour from the subject on which she currently spends more time.

This idea of increasing opportunity cost is reflected in the slope of the production possibilities curve. The curve is flatter to the left of point *T*, and steeper to the right, showing that, as Susan takes more and more of her resources (time, in this case) from one course and puts it into the other, she must give up greater and greater amounts of productivity in the course getting fewer resources.

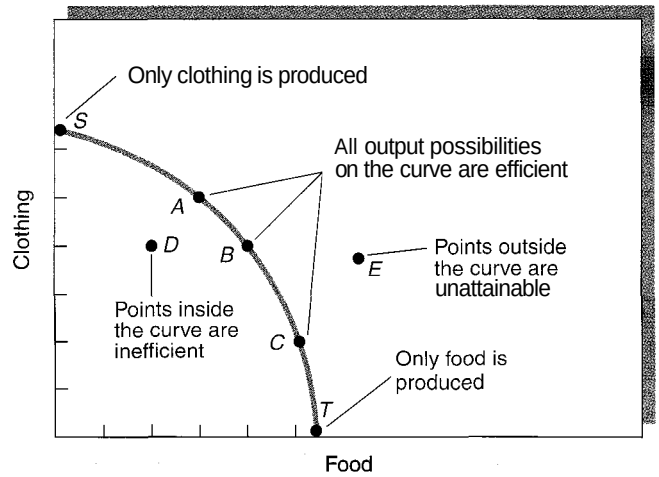
Of course, Susan could study more economics *without* giving up her English study time, if she gave up some leisure, or study time for other courses, or her part-time job in the campus bookstore. If she gave up leisure or her job and added those hours to the 10 hours of study time for economics and English, the entire curve in Exhibit 1 would shift outward. She could get better grades in both classes by having more time to study.

Can the production possibilities concept be applied to the entire economy? Yes. We can grow more soybeans if we grow less corn, since both can be grown on the same land. Beefing up the nation's military would mean we would have to produce fewer nonmilitary goods than we could otherwise. When scarce resources are being used efficiently, getting more of one requires that we sacrifice others.

Exhibit 2 shows a hypothetical production possibilities curve for an economy with a limited amount of resources that produces only two goods: food and clothing. The points along the curve represent all possible combinations of food and clothing that could be produced with the current level of resources and technology of the economy (assuming the resources are being used efficiently). A point outside the production possibilities curve (such as point *E*) would be considered unattainable, at the present time. **A** point inside the production possibilities curve (such as point *D*) is attainable. but producing that amount would mean that the economy is not making maximum use of its resources (some resources are being underutilized). Thus, point *D* is considered inefficient.

More specifically, the production possibilities curve shows all of the maximum combinations of two goods that an economy will be able to produce: (1) given a fixed quantity of resources, (2) holding the level of technology constant, and (3) assuming that all re-

EXHIBIT 2
Concept of Production Possibilities Curve for an Economy



When an economy is using its limited resources efficiently, production of more clothing requires that the economy give up some other goods—such as food in this example. In time, improved technology, more resources, or improvement in its economic organization could make it possible to produce more of both goods by shifting the production possibilities curve outward.

When these three conditions are met, the economy will be at the edge of its production possibilities frontier (where points **A,** *B,* and C lie), and producing more of one good will necessitate producing less of others. If condition 3 above is not met, and resources are being used inefficiently, an economy would be operating inside its production possibilities curve. If the level of resources and technology change (conditions I and 2), it will result in an outward shift in the production possibilities curve. We will return to these factors that can shift the production possibilities curve in a moment.

Notice that the production possibilities curve is convex, or bowed out from the origin, just as Susan's was in Exhibit 1 because of the concept of increasing opportunity cost. Here, the convexity reflects the fact that an economy's resources are not equally well suited to produce food and clothing. If an economy were using all its resources to produce clothing (point S), transferring those resources least suited for producing clothing toward food production will reduce clothing output a little but increase food output a lot. Since the resources transferred would be those better-suited for producing food and less-suited for producing clothing, the opportunity cost of producing additional food (in terms of clothing forgone) is low—near point *S*. However, as more and more resources are devoted to food production, and successively larger amounts of food are produced (moving the economy from *S* to **A** to *B* and so on), the opportunity cost of food will rise. This is because, as more and more food is produced, additional food output can be achieved only by using resources that are less and less suitable for the production of food relative to clothing. Thus, as food output is expanded, successively larger amounts of clothing must be forgone per unit of additional food. This is similar to what happened to Susan when she diverted study hours from one course to another. Only this time, we are talking about an entire economy.

Shifting the Production Possibilities Curve Outward

What restricts an economy—once its resources are fully utilized—from producing more of everything? Why can't we get more of something produced without having to give up the production of something else? The same constraint that kept Susan from simultaneously making a higher grade in both English and economics—a lack of resources. As long as all current resources are being used efficiently, the only way to get more of one good is to sacrifice some of the other. Over time, however, it is possible for a country's production possibilities curve to shift outward, making it possible for more of all goods to be

produced. Below we address four factors that could potentially shift the production possibilities curve outward.

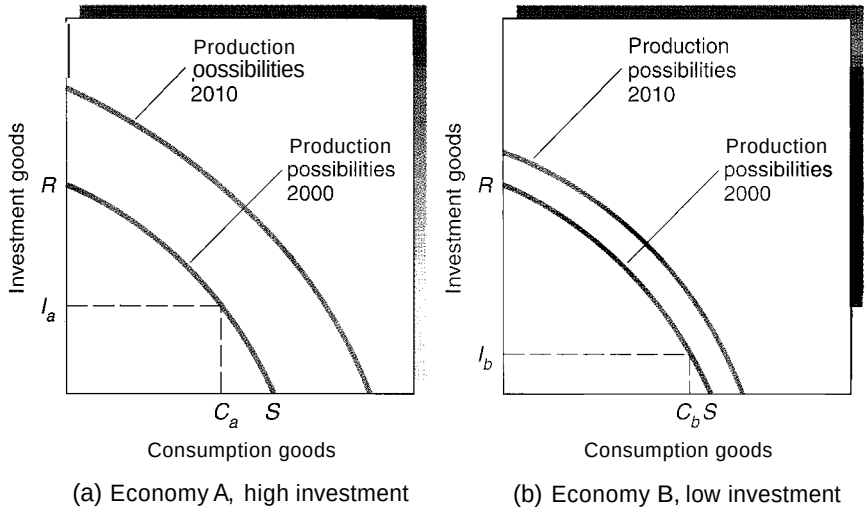
1. An increase in the economy's resource base would expand our ability to produce goods and services. If we had more or better resources. we could produce a greater amount of all goods. Resources such as machinery. buildings. tools. and education are human-made, and thus we can expand our resource base by devoting some of our efforts to producing them. This **investment** would provide us with better tools and skills and increase our ability to produce goods and services in the future. However. like with the production of other goods, devoting effort and resources toward producing these long-lasting physical assets means fewer resources are available to produce other things, in this case goods for current consumption. Thus. the choice between using resources to produce goods for current consumption and using them to produce investment goods for the future can also be illustrated within the production possibilities framework. The two economies illustrated in **Exhibit 3** begin with identical production possibilities curves (*RS*). Notice that Economy A dedicates more of its output to investment (shown by I_a) than Economy B (shown by I_b). Economy B, on the other hand, consumes more than Economy A. Because Economy A allocates more of its resources to investment and less to consumption, A's production possibilities curve shifts outward over time by a greater amount than B's. In other words, the growth rate of Economy A—the expansion of its ability to produce goods—is enhanced by this investment. But more investment in machines and human skills requires a reduction in current consumption.

Investment
The purchase, construction, or development of resources, including physical assets, such as plants and machinery, and human assets. such as better education. Investment expands an economy's resources. The process of investment is sometimes called capital formation.

2. Advancements in technology can expand the economy's production possibilities. **Technology** determines the maximum amount of output an economy can produce given the resources it has. New and better technology makes it possible for us to get more output from our resources. An important form of technological change is **invention**—the use of science and engineering to create new products or processes. In recent years, for example, inventions have allowed us to develop photographs faster and more cheaply, process data more rapidly, get more oil from existing fields, and send information instantly and cheaply by satellite. Such technological advances increase our production possibilities, shifting our economy's entire production possibilities curve outward.

Technology
The technological knowledge available in an economy at any given time. The level of technology determines the amount of output we can generate with our limited resources.

Invention
The creation of a new product or process, often facilitated by the knowledge of engineering and science.



(a) Economy A, high investment (b) Economy B, low investment

EXHIBIT 3
Investment and Production Possibilities in the Future

Here we illustrate two economies (A and B) that initially confront identical production possibilities curves *(RS)*. Economy A allocates a larger share of its output to investment (I_a, compared to I_b for Economy B). As a result, the production possibilities curve of the high-investment economy (Economy A) will tend to shift outward by a larger amount over time than the low-investment economy's will.

Innovation
The successful introduction and adoption of a new product or process; the economic application of inventions and marketing techniques.

Entrepreneur
A person who introduces new products or improved technologies and decides which projects to undertake. A successful entrepreneur's actions will increase the value of resources and expand the size of the economic pie.

An economy can also benefit from technological change through **innovation**—the practical and effective adoption of new techniques. Such innovation is commonly carried out by an **entrepreneur**—a person who introduces new products or improved techniques to satisfy consumers at a lower cost. To make a profit, an entrepreneur must convert or rearrange resources in a way that increases their value. This also pushes the production possibilities curve outward.

Take, for example, Henry Ford, an entrepreneur who changed how cars were made by pioneering the assembly line. With the same amount of labor and materials, Ford made more cars, more cheaply. Another entrepreneur, the late Ray Kroc, purchased a hamburger restaurant from Richard McDonald and built it into the world's largest fast-food chain. Kroc revolutionized fast food by offering attractive food at economical prices. He also developed a franchising system that resulted in uniform quality across the many different McDonald's restaurants worldwide. More recently, entrepreneurs like Steven Jobs (Apple Computer) and Bill Gates (Microsoft) helped develop the personal computer and software programs that dramatically increased their usefulness to businesses and households. It is interesting to think about how a few famous entrepreneurs have improved our productivity and changed our lives so much.

3. An improvement in the rules under which the economy functions can also increase output. The legal system of a country influences the ability of people to cooperate with one another and produce goods. Changes in legal institutions that promote social cooperation and motivate people to produce will also push the production possibilities curve outward. On the other hand, poor institutions can reduce both the level of resources used (shifting the curve inward) and how efficiently they are used (causing the economy to operate inside its production possibilities curve).

Historically, legal innovations have been an important source of economic progress. During the eighteenth century, a system of patents was established in Europe and North America, giving inventors private-property rights to their ideas. At about the same time, laws were passed allowing businesses to establish themselves legally as corporations, reducing the cost of forming large firms that were often required for the mass production of manufactured goods. Both of these legal changes improved economic organization and accelerated the growth of output by shifting the production possibilities curve outward more rapidly.

Sometimes governments, perhaps because of ignorance or prejudice, adopt legal institutions that reduce production possibilities. Laws that restrict or prohibit trade is one example. For almost a hunderd years following the American Civil War, the laws of several southern states prohibited hiring African-Americans for certain jobs and restricted other economic exchanges between blacks and whites. The legislation not only was harmful to African-Americans, it also retarded progress and reduced the production possibilities of these states.

The collapse of communism in the 1980s also illustrates the importance of economic institutions. After the collapse, Russia was unable to develop legal institutions protecting property rights and enforcing contracts. The absence of these institutions hampered investment and the gains from trade. Investors moved their money to countries with more secure property rights, and resources within the country were used inefficiently because trade was hindered. As a result, even though Russia has a well-educated labor force and abundant natural resources, its economic performance has been poor.

4. By working harder and giving up current leisure, we could increase our production of goods and services. Hypothetically, the production possibilities curve would shift outward if everyone worked more hours and took less leisure time. Strictly speaking, however, leisure is also a good, so we would simply be giving up leisure to have more of other things. If we were to construct a production possibilities curve for leisure versus other goods, this would be shown as simply a movement along the curve. However, if we restrict our model to only material goods and services, a change in the amount we work would be shown as a shift in the curve.

How much people work depends not only on their personal preferences but also on public policy. For example, high tax rates on personal income may cause people to work less. This is because high tax rates reduce the payoff from working. When this happens, people spend more time doing other, untaxed activities—like leisure activities. This will move the production possibilities curve for material goods inward because the economy can't produce as much when people work less.

Production Possibilities and Economic Growth

Economic growth is one of the most important topics in modern economics for good reason. An economic growth rate of 3 percent per year will result in living standards doubling approximately every 24 years. On the other hand. in a country experiencing **an** economic growth rate of only 1 percent, it will take 75 years for living standards to double. Within the production possibilities framework. economic growth is simply an outward shift in the curve through time. The more rapidly the curve shifts outward. the more rapid is economic growth. There are other economic models that are used to analyze economic growth; however. they all share the production possibilities curve as a foundation.

TRADE, OUTPUT, AND LIVING STANDARDS

Trade makes it possible for people to generate more output through specialization and division of labor, large-scale production processes, and the dissemination of improved products and production methods.

Gains from Trade

As we previously discussed, trade creates value by moving goods from people who value them less to people who value them more. However, this is only part of the story. Trade also makes it possible for people to expand their output through specialization and **division of labor,** large-scale production, and the dissemination of better products and production methods.

Division of labor
A method that breaks down the production of a product into a series of specific tasks, each performed by a different worker.

Gains from Specialization and Division of Labor

Businesses can achieve higher output levels and greater productivity from their workers through specialization and division of labor. More than 200 years ago, Adam Smith noted the importance of this factor. Observing the operation of a pin manufacturer, Smith noted that when each worker specialized in a separate function needed to make pins, 10 workers together were able to produce 48,000 pins per day, or 4,800 pins per worker. Smith doubted an individual worker could produce even 20 pins per day working alone from start to finish on each pin.[7]

The division of labor separates production tasks into a series of related operations. Each worker performs one or a few of perhaps hundreds of tasks necessary to produce something. This process makes it possible to assign different tasks to those individuals who are able to accomplish them most efficiently (that is, at the lowest cost). Furthermore, a worker who specializes in just one narrow area becomes more experienced and more skilled in that task over time.

Trading partners can also benefit from specialization and the division of labor. The **law of comparative advantage**, developed in the early 1800s by the great English economist David Ricardo, explains why this is true. ***The law of comparative advantage states***

Law of comparative advantage
A principle that states that individuals, firms, regions, or nations can gain by specializing in the production of goods that they produce cheaply (at a low opportunity cost) and exchanging them for goods they cannot produce cheaply (at a high opportunity cost).

[7] *See* **Adam** Smith, *An Inquiry* **into** *the Nature and Causes of the Wealth of Nations* (1776, Cannan's ed , Chicago University of Chicago **Press,** 1976), 7–16, for additional detail on the Importance of the division of labor

that the total output of a group of individuals, an entire economy, or a group of nations will be greatest when the output of each good is produced by the person (or firm) with the lowest opportunity cost.

Comparative advantage applies to trade among individuals, business firms, regions, and even nations. When trading partners are able to use more of their time and resources to produce the things each is best at, they will be able to produce more together than would otherwise have been possible. In turn, the mutual gains they get from trading will result in higher levels of income for each. It's a win-win situation for both.

If a good or service can be obtained more economically through trade, it makes sense to get it that way rather than producing it for yourself. For example, even though most doctors might be good at record keeping and arranging appointments, it's generally better for them to hire someone to perform these services for them. That's because the time doctors spend keeping records is time they could have spent seeing patients. The revenue forgone as a result of seeing fewer patients would be greater than the cost of hiring the worker. The issue is not whether doctors are better record keepers than the assistants they could hire, but rather how they should use their time most efficiently.

If you think about it, the law of comparative advantage is common sense. If someone else is willing to supply you with a good at a lower cost than you can produce it yourself, doesn't it make sense to trade for it and use your time and resources to produce more of the things you can produce most efficiently? Consider the situation of Andrea, an attorney who earns $100 per hour providing legal services. She has several documents that need to be typed, and she is thinking about hiring a typist earning $15 per hour to do it. Andrea is an excellent typist, much faster than the prospective employee. She could do the job in 20 hours, whereas the typist would take 40 hours.

Because of her greater typing speed, some might think Andrea should handle the job herself. This is not the case. If she types the documents, the job will cost her $2,000—the opportunity cost of 20 hours of practicing law at $100 per hour. Alternatively, the cost of having the documents typed by the typist is only $600 (40 hours at $15 per hour). Andrea's comparative advantage lies in practicing law. By hiring the typist, she will increase her own productivity and make more money.

The implications of the law of comparative advantage are universal. Any group will be able to produce more output from its available resources when each good or service is produced by the person with the lowest opportunity cost. This insight is particularly important in understanding the way a market economy works. Purposeful decision making

© JEFF GREENBERG/PHOTOEDIT

© JAMES NUBILE/THE IMAGE WORKS

© WESLEY BOCXE/THE IMAGE WORKS

Trade channels goods to those who value them most. Trade also helps disseminate ideas for improved products and makes production methods such as specialization, the division of labor, and mass production more feasible. Over the years, trade has enabled us to produce more with our limited resources, dramatically improving our living standards.

indicates that buyers will try to get the most for their money. They will not knowingly choose a high-cost option when a lower-cost alternative is available. This places low-cost suppliers at a competitive advantage. Thus, they will generally survive and prosper in a market economy. As a result, the production of goods and resources will naturally tend to be allocated according to comparative advantage.

Most people recognize that Americans benefit from trade among the nation's 50 states. For example, the residents of Nebraska and Florida are able to produce a larger joint output and achieve higher income levels when Nebraskans specialize in producing wheat and other grain products, and Floridians specialize in producing oranges and other citrus products. The same is true for trade among nations. Like Nebraskans and Floridians, people in different nations will be better off if they specialize in the goods and services they can produce at a low cost and trade them for goods they produce at a high cost. See the addendum to this chapter for additional evidence on this point.

Gains from Mass Production Methods

Trade also promotes economic progress by making it possible for firms to lower their per-unit costs with mass production. Say a nation isolated itself and refused to trade with other countries. In an economy like this, self-sufficiency and small-scale production would be the norm. If trade were allowed, however, the nation's firms could sell their products to customers around the world. This would make it feasible for the firms to adopt more efficient, large-scale production processes. Mass production often leads to labor and machinery efficiencies that increase enormously the output per worker. But without trade, these gains could not be achieved.

Gains from Innovation

Trade also makes it possible to realize gains from the discovery and dissemination of innovative products and production processes. Economic growth involves brain power, innovation, and the application of technology. Without trade, however, the gains derived from the discovery of better ways of doing things would be stifled. Furthermore, observing and interacting with other people using different and better technologies often encourages others to copy successful approaches. People also modify the technology they observe, adapting it for their own purposes. This sometimes results in new, and even better, technologies. Again, gains from these sources would be far more limited in a world without trade.

Can you imagine the difficulty involved in producing your own housing, clothing, and food, to say nothing of radios, television sets, dishwashers, automobiles, and telephone services? Yet, most families in North America, Western Europe, Japan, and Australia enjoy all these conveniences. They are able to do so largely because their economies are organized in such a way that individuals can cooperate. specialize. and trade. thereby reaping the benefits of the enormous increases in output—in both quantity and diversity—that can be generated. On the other hand, countries that impose obstacles that retard exchange—either domestic or international—hinder their citizens from achieving these gains and more prosperous lives.

HUMAN INGENUITY AND THE CREATION OF WEALTH

Economic goods are the result of human ingenuity and action; thus, the size of the "economic pie" is variable, not fixed.

Human Ingenuity

The size of a country's "economic pie" is most easily thought of as the total dollar value of all goods and services produced during some period of time. This economic pie is the total amount of wealth (or value) created in the economy. It is not some fixed total waiting

to be divided up among people. It is simply a statistic—a grand total, calculated by adding up the wealth created by each of the individuals in the economy. **As** the quotation at the chapter opening suggests, many errors in economic reasoning stem from the incorrect notion that the size of the economic pie is fixed.

On the contrary, the size of the economic pie reflects the physical effort and ingenuity of human beings. It is not an endowment from nature. Economic output expands as we discover better ways of doing things. So over time, it is human knowledge and ingenuity—perhaps more than anything else—that limits our economic progress. If Jim, a local farmer who normally produces $30,000 worth of corn each year, finds a better growing method enabling him to produce $40,000 of corn each year, he has created additional wealth. But Jim has actually created more than the $10,000 in extra wealth. The $10,000 is only his share of the gains from the additional trades made possible by the extra corn he grew. Exchange makes both buyer and seller better off, so the total wealth created by Jim includes not only his $10,000 but also the gains of all of the buyers who purchased corn from him as well.

This highlights an important point: in a market economy, a larger income for one person does not mean a smaller income for another. In fact, it is just the opposite. When a person earns income, he or she expands the economic pie by more than the amount of the slice that he or she gets, making it possible for the rest of us to have a bigger slice, too. When a wealthy entrepreneur, such as Bill Gates or Henry Ford, has an income of, say, $1 billion per year earned through voluntary exchanges in the marketplace, he has enlarged the economic pie for everyone by an even larger amount. Here's how:

Suppose that Linda, a freelance graphic artist, pays $175 for a new software program developed by Bill Gates. As a result, she can do twice as much work in the same amount of time. Because she's more productive, Linda can earn more than enough additional income with the software to justify her purchase. In addition, the businesses she serves are also likely to be better off because the software makes it possible for her to give them more and better service and a lower price for her services. Thus, while Bill Gates gained, so, too, did Linda and her customers.

Similarly, although Henry Ford certainly became rich, he also greatly increased our ability to transport goods, services, and people. In the process, he made it possible for many others to achieve higher living standards than would have been possible in his absence. Had Stephen King never written a novel, not only would he not be as rich, but we would all be poorer for never having had the opportunity to read his novels. When income is acquired through voluntary exchange, people who earn income also help others earn more income and live better, too.

ECONOMIC ORGANIZATION

Market organization
A method of organization in which private parties make their own plans and decisions with the guidance of unregulated market prices. The basic economic questions of consumption, production, and distribution are answered through these decentralized decisions.

Capitalism
An economic system in which productive resources are owned privately and goods and resources are allocated through market prices.

Every economy faces three basic questions: (1) What will be produced? (2) How will it be produced? and (3) For whom will it be produced? These problems are highly interrelated. Throughout the book, we will consider how different types of economies solve these issues. There are two broad ways that an economy can be organized: markets and government (political) planning. Let us briefly consider each.

Market Organization

Private ownership of productive assets, voluntary contracts (often verbal), and market prices are the distinguishing features of **market organization**. Market organization is also known as **capitalism**.[8] Under market organization, private parties are permitted to buy and sell ownership rights of their assets at mutually acceptable prices. The government plays the limited role of rule maker and referee. It develops the rules, or the legal structure, that recognize, define, and protect private ownership rights. It enforces contracts and protects

[8] *Capitalism* is a term coined by Karl Marx

people from violence and fraud. But the government is not an active player in the economy. Ideally, it avoids modifying market outcomes in an attempt to favor some people at the expense of others. For example, it doesn't prevent sellers from slashing prices or improving the quality of their products to attract customers from other competitors. Nor does it prevent buyers from outbidding others for products and productive resources. No legal restraints limit potential buyers or sellers from producing. selling. or buying in the marketplace.

Under market organization, no single individual or group of individuals guides the economy. There is no central planning authority, only individual planning. The three basic questions are solved independently in the marketplace by individual buyers and sellers making their own decentralized decisions. Buyers and sellers decide on their own what to produce, how to produce it, and whom to trade it to. based on the prices they themselves decide to charge.

In markets, individual buyers and sellers communicate their desires and preferences both directly and indirectly. They directly voice their desires when they buy or sell by advertising, whether in print or broadcast. or informally by word of mouth, on bulletin boards, and by letters of request and complaint and other means. They communicate indirectly by exiting or entering exchange relationships, as when they stop purchasing Coke and switch to Pepsi. The indirect, or "exit," option gives special power to their voiced, or direct, statements. Indeed, sellers, when markets are competitive, often hire experts to seek out the statements and desires of potential buyers. Buyers, too, are eager to know what sellers want—special terms of payment or delivery, for example—hoping that sellers might be willing to reward cooperation with a better deal.

Political Planning

The major alternative to market organization is **collective decision making,** whereby the government, through the political process, makes decisions for buyers and sellers in an attempt to solve the basic economic questions facing the economy. The government may maintain private ownership, but uses taxes, subsidies, and regulations to resolve the basic economic questions. Alternatively, an economic system in which the government also owns the income-producing assets (machines, buildings, and land) and directly determines what goods will be produced is called **socialism**. Either way, individual planning and decisions are replaced by central planning and decisions made through the political process. These decisions can be made by a single dictator or a group of experts, or through democratic voting. Political rather than market forces direct the economy, and government officials and planning boards hand down decisions to expand or contract the output of education, medical services, automobiles, electricity, steel, consumer durables, and thousands of other commodities.

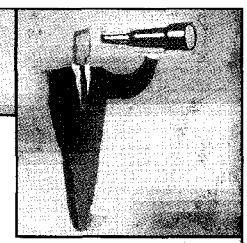
Collective decision making
The method of organization that relies on public-sector decision making (voting, political bargaining, lobbying, and so on) to resolve basic economic questions.

Socialism
A system of economic organization in which (1) the ownership and control of the basic means of production rest with the state, and (2) resource allocation is determined by centralized planning rather than market forces.

This is not to say that the preferences of individuals carry no weight. If the government officials and central planners are influenced by the democratic process, they must consider how their actions will influence their reelection prospects. That means they will listen to the voices of the voters to win over a majority of them. Otherwise, like the firm in a market economy that produces a product that consumers do not want, their tenure of service is likely to be short. However, under central planning the indirect exit method of communicating is much more difficult. Although people can use the direct or voice method to communicate their preferences by lobbying government officials or casting votes in an election, they generally cannot use the indirect exit option because they cannot refuse to pay taxes or to quit purchasing a good or service that is provided by government. For example, families who send their children to private school must continue to pay the same amount in taxes to support the public school system as they would if they kept their child in public school. Oftentimes, people "vote with their feet" and leave one political jurisdiction to move to another. This is frequently seen when people move to better school districts. It is much easier, however, to move between school districts than between states or nations.

In summary, both market organization and central planning face the same basic economic questions. **A** basic difference between them is that the market system, with its exit option, allows for a wider variety of products and creates constant competition among suppliers, whereas the central planning system, in a democracy, responds primarily to the votes of the majority. In varying degrees, all economies use a combination of both of these methods of economic organization. Even predominantly market economies will still use taxes, subsidies, and some government ownership to direct and control resources. Similarly, predominantly socialist economies will, to some degree, use markets to allocate certain goods and services.

LOOKING AHEAD



The next two chapters present an overview of the market sector and explain how supply and demand for goods and services work. Chapters 5 and 6focus on potential shortcomings of the market and how the collective decision-making process works in a democracy. **As** *we proceed, the tools of economics will be used to analyze both the market and political sectors. We think this approach is important and that you will find it both interesting and enlightening.*

! KEY POINTS

- The highest valued activity sacrificed when a choice is made is the opportunity cost of the choice; differences (or changes) in opportunity costs help explain human behavior.
- Mutual gain is the foundation of trade. When two parties engage in voluntary exchange, they are both made better off. Trade creates value because it channels goods and resources to those who value them the most.
- Transaction costs—the time, effort, and other resources necessary to search out, negotiate, and conclude an exchange—hinderthe gains from trade in an economy. Middlemen perform a productive function by reducing transaction costs.
- Private-property rights motivate owners to use their resources in ways that benefit others and avoid doing harm to them. These rights also motivate owners to take proper care of their resources and conserve them.
- The production possibilities curve shows the maximum combination of any two products that can be produced with a fixed quantity of resources.
- Over time, the production possibilities curve of an economy can be shifted outward by (1) investment, (2) technological advances, (3) improved institutions, and (4) greater work effort (forgoing leisure). The size of the economic pie is variable, not fixed. It can grow (or shrink) over time.
- The law of comparative advantage indicates that the joint output of individuals, regions, and nations will be maximized when each productive activity is undertaken by the low-opportunity-cost supplier. When a good can be acquired through trade more economically than it can be produced directly, it makes sense to trade for it.
- In addition to the gains that occur when goods are moved toward those who value them most, trade also makes it possible to expand output through specialization, division of labor, mass production processes, and innovation. These improved production techniques have contributed greatly to our modern living standards.
- Economies can either be organized by decentralized markets (capitalism) or they can be centrally planned by government through political decision making. Under central planning, buyers and sellers are more limited in their ability to communicate their desires.

CRITICAL ANALYSIS QUESTIONS

1. "If Jones trades a used car to Smith for $5,000, nothing new is created. Thus, there is no way the transaction can improve the welfare of people." Is this statement true? Why or why not?

"2. Economists often argue that wage rates reflect productivity. Yet, the wages of house painters have increased nearly as rapidly as the national average, even though these workers use approximately the same production methods as they did 50 years ago. Can you explain why the wages of painters have risen substantially even though their productivity has changed so little?

3. It takes one hour to travel from New York City to Washington, D.C., by air, but it takes five hours by bus. If the airfare is $110 and the bus fare is $70, which would be cheaper for someone whose opportunity cost of travel time is $6 per hour? For someone whose opportunity cost is $10 per hour? $14 per hour?

"4. "People in business get ahead by exploiting the needs of their consumers. The gains of business are at the expense of suffering imposed on their customers." Evaluate this statement.

5. What is the objective of the entrepreneur when it comes to the use of his or her resources? What is the major function of the middleman? Is the middleman an entrepreneur?

6. If you have a private-ownership right to something, what does this mean? Does private ownership give you the right to do anything you want with the things that you own? Explain. How does private ownership influence the incentive of individuals to (a) take care of things, (b) conserve resources for the future, and (c) develop and modify things in ways that are beneficial to others? Explain.

7. What is the law of comparative advantage? According to the law of comparative advantage, what should be the distinguishing characteristics of the goods a nation produces? What should be the distinguishing characteristics of the goods a nation imports? How will international trade influence people's production levels and living standards? Explain.

"8. Does a 60-year-old tree farmer have an incentive to plant and care for Douglas fir trees that will not reach optimal cutting size for another 50 years?

'"9. What forms of competition does a private-property, market-directed economy authorize? What forms does it prohibit?

10. What are the major sources of gains from trade? Why is exchange important to a nation's prosperity? How does trade influence the quantity of output that trading partners are able to produce? In a market economy, will there be a tendency for both resources and products to be supplied by low-cost producers? Why or why not? Does this matter? Explain.

***11.** Chick-fil-A's "Eat Mor Chikin" advertising campaign features three cows holding signs that say things like: "Save the cows, eat more chicken." If consumers began eating more chicken and less beef, would the cattle population increase or decrease? Explain.

***12.** In many states. ticket scalping, or reselling tickets to entertainment events at prices above the original purchase price. is prohibited. Who is helped and who is hurt by such prohibitions? How can owners who want to sell their tickets get around the prohibition? Do you think it would be a good idea to prohibit the resale of other things — automobiles, books, works of art, or stock shares at prices higher than the original purchase price? Why or why not?

13. Consider the choices of two groups of women ages 30 to 50. All the women in one group have a college education. All the women in the other group have less than a high school education. Which of the two groups will participate more in the workforce? Which of the two groups will bear a larger number of children on average? Explain your answers based on the concept of opportunity cost.

14. Consider the questions below:

a. Do you think that your work effort is influenced by whether there is a close link between personal output and personal compensation (reward)? Explain.

b. Suppose the grades in your class were going to be determined by a random drawing at the end of the course. How would this influence your study habits?

c. How would your study habits be influenced if everyone in the class were going to be given an **A** grade? How about if grades were based entirely on examinations composed of the multiple-choice questions in the course book for this textbook?

d. Do you think the total output of a nation will be influenced by whether or not there is a close link between the productive contribution of individuals and their personal reward? Why or why not?

15. In the chapter it was stated that a private-property right also involves having the right to transfer or exchange what you own with others. However, selling your organs is a violation of federal law, a felony punishable by up to five years in prison or a $50,000 fine. In 1999 eBay intervened when a person put one of his kidneys up for sale on the auction site (the bidding

reached $5.7 million before the auction was halted). Does this lack of legal ability to exchange mean that individuals do not own their own organs? Explain.

16. During the last three decades entrepreneurs like Michael Dell, Sam Walton, and Ted Turner earned billions of dollars. Do you think the average American is better or worse off as the result of the economic activities of these individuals? Explain your response.

***17.** As the skill level (and therefore earnings rate) of, say, an architect, computer specialist, or chemist increases, what happens to his or her opportunity cost of doing other things? How is the time spent on leisure likely to change?

18. Two centuries ago there were more buffalo than cattle in the United States. Even though millions of cattle are killed for beef consumption each year, the cattle population continues to grow while the buffalo are virtually extinct. Why?

19. The tables below show the production possibilities for two hypothetical countries, Italia and Nire. Which country has the comparative advantage in producing butter? Which country has the comparative advantage in producing guns? What would be a mutually agreeable rate of exchange between the countries?

Italia		Nire	
Guns	Butter	Guns	Butter
12	0	16	0
8	2	12	1
4	4	8	2
0	6	4	3
		0	4

***Asterisk** denotes questions for which answers are given in **Appendix** B

ADDENDUM

Comparative Advantage, Specialization, and Gains from Trade

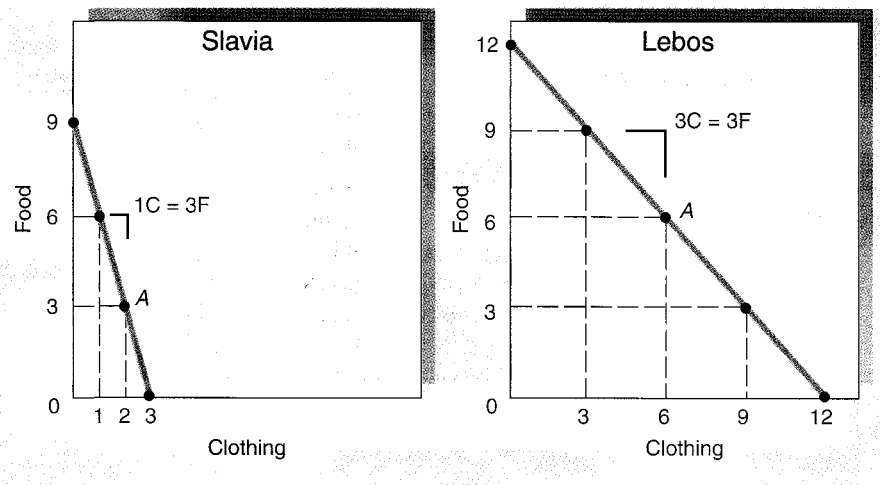
This addendum is for instructors who want to assign a more detailed numerical example demonstrating comparative advantage, specialization, and mutual gains from trade. Students who are uncertain about their understanding of these topics may also find this material enlightening. The international-trade chapter later in the text provides still more information on trade and how it affects our lives.

We begin with hypothetical production possibilities curves for two countries, Slavia and Lebos, shown in **Exhibit A-1.** The numerical tables represent selected points from each country's production possibilities curve. To make calculations easier, we have assumed away increasing opportunity costs in production so that the production possibilities curves are linear.

Without trade, each country would be able to consume only what it can produce for itself. Let's arbitrarily assume

EXHIBIT A-1
Production Possibilities for Slavia and Lebos

For Slavia, the opportunity cost of producing 1 unit of clothing is equal to 3 units of food (1C = 3F). For Lebos, the opportunity cost of producing 3 units of clothing is equal to 3 units of food (3C = 3F or 1C = 1F). The difference in the opportunity costs of production will make possible mutually beneficial trade between the countries, with each specializing in its area of comparative advantage.



SLAVIA

Food		Clothing
9		0
6	3F = 1C	1
3		2
0		3

LEBOS

Food		Clothing
12		0
9	3F = 3C	3
6		6
3		9
0		12

that for survival Slavia requires 3 units of food and Lebos requires 6 units of food. As can be seen by point *A* in the exhibit, if Slavia were to produce the 3 units of food it requires, it would have enough resources remaining to produce 2 units of clothing. Similarly, if Lebos were to produce the 6 units of food it requires, it would have enough resources left to produce 6 units of clothing, again shown by point **A** in the exhibit. As we proceed we will use this outcome as our benchmark outcome that occurs in the absence of specialization and trade between the countries.

Economic analysis suggests that both countries could gain if each were to specialize in the production of the good for which they have the comparative advantage and then trade for the other. First, let's figure out which country has a comparative advantage in the production of clothing. Doing so requires calculating the opportunity cost of producing clothing for each country. Because, in this example, the opportunity costs are constant at all points along the production possibilities curve, rather than increasing, this can be found by first selecting any two points on the production possibilities curve (or equivalently by comparing any two rows of numbers in the numerical tables given in the exhibit). For Slavia, moving from the point of producing 6 food units and 1 clothing unit to the alternative point of producing 3 food units and 2 clothing units, we see that Slavia gains 1 clothing unit but must give up 3 units of food. For simplicity, the opportunity cost for Slavia can be written as 1C = 3F, where C stands for clothing and F for food. You might note that this same numerical trade-off is true for Slavia anywhere along its production possibilities curve (for example, beginning from 9 food units and 0 clothing units, it would also have to give up 3 food units to gain 1 unit of clothing).

Using a similar approach (taking any 2 points or 2 rows in the table) for Lebos shows that for every 3 units of clothing the country wishes to produce, it must give up 3 units of food (3C = 3F). This can be treated as any other mathematical equation, and can be simplified by dividing both sides by 3, resulting in an opportunity cost of 1 clothing unit equals 1 food unit (1C = 1F). Now, compare this to the opportunity cost for Slavia (1C = 3F). Slavia must give up the production of 3 units of food for every 1 unit of clothing it produces, whereas Lebos must give up only 1 unit of food for every 1 unit of clothing it produces. Thus, Lebos gives up the production of ***less*** food for every unit of clothing. Lebos is the low-opportunity-cost producer of clothing, and thus it has a comparative advantage in the production of clothing.

Because comparative advantage is a relative comparison, if one country has the comparative advantage in the production of one of the products, the other country must have the comparative advantage for the other good. Thus, because Lebos has the comparative advantage in clothing, it will be true that Slavia has the comparative advantage in food. However, it is worthwhile to show this here as well. To produce 1 unit of food, Lebos must give up 1 unit of clothing (recall the 1C = 1F opportunity cost). To produce 1 unit of food, Slavia must give up the production of only one-third of a unit of clothing (recall the 1C = 3F opportunity cost and rewrite the equation as 1/3 C = IF by dividing both sides of the equation by 3). Thus, Slavia gives up the production of *less* clothing for every unit of food produced. Slavia is the low-opportunity-cost producer of food, and thus has a comparative advantage in the production of food.

Suppose that, according to their comparative advantages, Lebos specializes in producing clothing and Slavia in food. From the last row of the table for Lebos. you can see that it can produce 12 units of clothing (and 0 food) if it specializes in producing only clothing. From the top row of the table for Slavia, you can see that it can produce 9 units of food (and 0 clothing) if it specializes in producing only food. Note that this joint output (9 food and 12 clothing) is greater than the benchmark joint output (9 food and 8 clothing) produced and consumed without trade.

If they are to trade, the countries now must find a mutually agreeable rate of exchange. Any rate of exchange ***between*** the two opportunity costs of 1C = 3F and 3C = 3F would be mutually agreeable. Here we will use 2C = 3F.

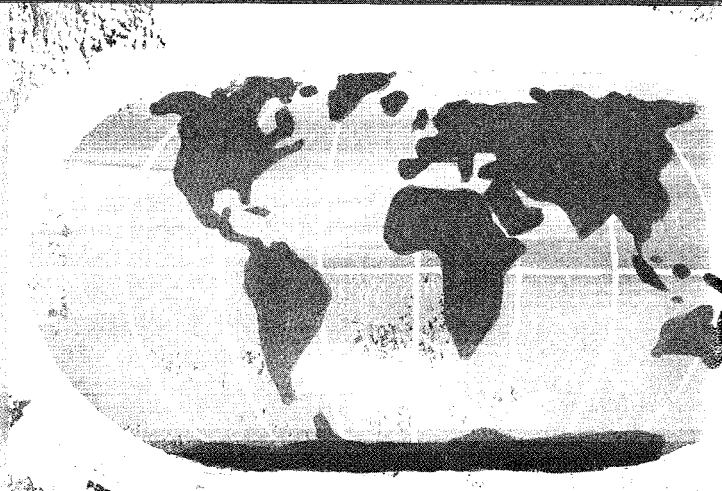
Recall that Slavia requires 3 units of food for survival. Now, however, they are specializing and producing 9 units of food. Using the rate of exchange above, Slavia would send its extra 6 units of food to Lebos in exchange for 4 units of clothing. After trade, Slavia would then have 3 units of food and 4 units of clothing. Compare this to the situation that existed before specialization and trade, in which Slavia had only 3 units of food and 2 units of clothing to consume. Specialization and trade have created 2 additional units of clothing for Slavia that it would not have had without trade.

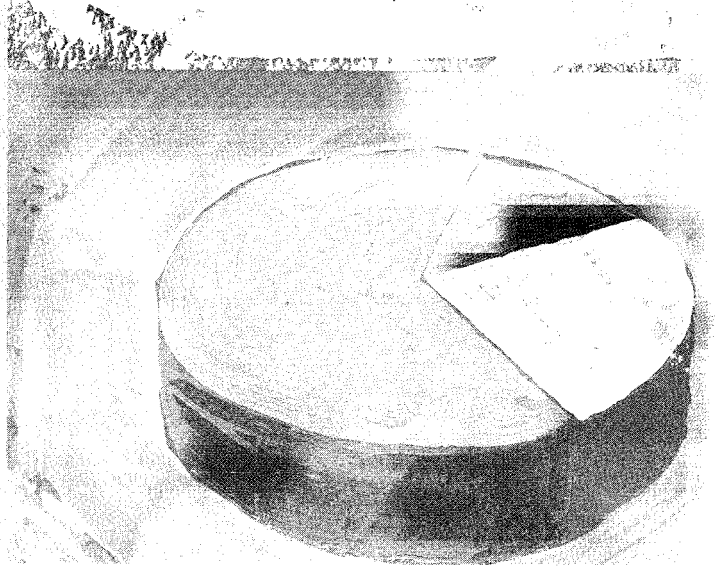
With specialization, Lebos is producing 12 units of clothing. In the trade with Slavia, Lebos gave up 4 units of clothing to obtain 6 units of food. After trade. Lebos has 8 units of clothing remaining and 6 units of food imported from Slavia. Compare this to the situation that existed before specialization and trade. in which Lebos had only 6 units of food and 6 units of clothing to consume. For Lebos. specialization and trade have also created 2 additional units of clothing that it would not have had without trade.

As this simple example shows, total output is greater and *both* countries are better off when they specialize in the area in which they have a comparative advantage. By doing so. each is able to consume a bundle of goods and services that exceeds what it could have achieved in the absence of trade. This concept applies equally to individuals, states. or nations. The typical worker could not begin to produce alone all of the things he or she can afford to buy with the money earned in a year by specializing and working in a single occupation. As our world has become more integrated over the past several hundred years, the gains that have occurred from specialization and trade are at the root of the significant improvements in well-being that we have experienced.

PART 2

"There are two primary methods of allocating scarce resources: markets and governments"





Markets and Government

PART 2

Economics has a great deal to say about how both markets and governments allocate scarce resources. It gives us insight about the conditions under which each will likely work well (and each will likely work poorly). The next four chapters will focus on this topic.

MARKET ALLOCATION OF RESOURCES

Business firms purchase resources like materials, labor services, tools, and machines from households in exchange for income, bidding the resources away from their alternate uses. The firms then transform the resources into products like shoes, automobiles, food products, and medical services and sell them to households. In a market economy, businesses will continue to supply a good or service only if the revenues from the sale of the product are sufficient to cover the cost of the resources required for its production.

GOVERNMENT ALLOCATION OF RESOURCES

Resource allocation by the government involves a more complex, three-sided exchange. In a democratic political setting, a legislative body levies taxes on voter-citizens, and these revenues are subdivided into budgets, which are allocated to government bureaus and agencies. In turn, the bureaus and agencies use the funds from their budgets to supply goods, services, and income transfers to voter-citizens. The legislative body is like a board of directors elected by the citizens. The competitive pressure to get elected gives legislators a strong incentive to cater to the wishes of voters. In turn, voters will be more likely to support a legislator if the value of the goods, services, and transfers received by them is high relative to the taxes they have to pay. In other words, goods, services, and income transfers will be supplied by the government if, and only if, a majority of legislators believe it will improve their election prospects. As you can tell, this is quite different from the way markets allocate goods and services! Now let's see just how really different it is and what impact it has on you and the rest of the economy.

CHAPTER 3

Supply, Demand, and the Market Process

I am convinced that if [the market system] were the result of deliberate human design, and if the people guided by the price changes understood that their decisions have significance far beyond their immediate aim, this mechanism would have been acclaimed as one of the greatest triumphs of the human mind.

—*Friedrich Hayek, Nobel laureate*[1]

From the point of view of physics, it is a miracle that [7 million New Yorkers are fed each day] without any control mechanism other than sheer capitalism.

—*John H. Holland, scientist, Santa Fe Institute*

Chapter Focus

- **What are the laws of demand and supply?**

  How do consumers decide whether to purchase a good? How do producers decide whether to supply it?
- **How do buyers and sellers respond to changes in the price of a good?**
- **What role do profits and losses play in an economy? What must a firm do to make a profit?**
- **How is the market price of a good determined?**
- **How do markets adjust to changes in demand? How do they adjust to changes in supply?**
- **What is the "invisible hand" principle?**

[1]Friedrich Hayek, "The Use of Knowledge in Society," *American Economic Review* 35 (September 1945) 519–30

To those who study art, the *Mona Lisa* is much more than a famous painting of a woman. Looking beyond the overall picture, they see and appreciate the brush strokes, colors, and techniques embodied in the painting. Similarly, studying economics can help you to gain an appreciation for the details behind many things in your everyday life. During your last visit to the grocery store, you probably noticed the fruit and vegetable section. Next time, take a moment to ponder how potatoes from Idaho, oranges from Florida, apples from Washington, bananas from Honduras, kiwi fruit from New Zealand, and other items from around the world got there. Literally thousands of different individuals, *working independently,* were involved in the process. Their actions were so well coordinated, in fact, that the amount of each good was just about right to fill exactly the desires of your local community. Furthermore, even the goods shipped from halfway around the world were fresh and reasonably priced.

How does all this happen? The short answer is that it is the result of market prices and the incentives and coordination that flow from them. To the economist, the operation of markets—including your local grocery market—is like the brush strokes underlying a beautiful painting. Reflecting on this point, Professor Hayek speculates that if the market system had been deliberately designed, it would be "acclaimed as one of the greatest triumphs of the human mind." Similarly, computer scientist John H. Holland argues that, from the viewpoint of physics, the feeding of millions of New Yorkers day after day with very few shortages or surpluses is a miraculous feat (see the quotations at the chapter opening).

Amazingly, markets coordinate the actions of millions of individuals *without* central planning. There is no individual, political authority, or central planning committee in charge. Considering that there are nearly 300 million Americans with widely varying skills and desires, and roughly 25 million businesses producing a vast array of products ranging from diamond rings to toilet paper, the coordination derived from markets is indeed an awesome achievement.

This chapter focuses on supply, demand, and the determination of market prices. For now, we will analyze the operation of competitive markets—that is, markets with unrestricted numbers of buyers and sellers. We will also assume that the property rights are well defined. Later, we will consider what happens when these conditions are absent.

On eBay, sellers enter their reserve prices—the minimum prices they will accept for goods; buyers enter their maximum bids—the maximum prices they are willing to pay for goods. The process works the same way when a person runs a

© KAYTE M. DEIOMAI/PHOTOEDIT

The produce section of your local grocery store is a great place to see economics in action. Literally millions of individuals from around the world have been involved in the process of getting these goods to the shelves in just the right quantities. Market prices underlie this feat.

newspaper ad to sell a car. The seller has in mind a minimum price he or she will accept for the car. A potential buyer, on the other hand, has in mind a maximum price he or she will pay for the car. If the buyer's maximum price is greater than the seller's minimum price, the exchange will occur at a price somewhere in between. As these examples show, the buyers' and sellers' desires and incentives determine prices and make markets work. We will begin with the demand (buyer's) side, and then turn to the supply (seller's) side of the market.

CONSUMER CHOICE AND THE LAW OF DEMAND

Clearly, prices influence our decisions. As the price of a good increases, we have to give up more of *other* goods if we want to buy it. Thus, as the price of a good rises, its opportunity cost increases (in terms of other goods that must be forgone to purchase it).

Law of demand
A principle that states there is an inverse relationship between the price of a good and the quantity of it buyers are willing to purchase As the price of a good increases, consumers will wish to purchase less of it. As the price decreases, consumers will wish to purchase more of it.

Substitutes
Products that serve similar purposes. An increase in the price of one will cause an increase in demand for the other (examples are hamburgers and tacos, butter and margarine, Microsoft Xbox and Sony Play-Station, Chevrolets and Fords).

A basic principle of economics is that if something becomes more costly, people will be less likely to buy it. This principle is called the **law of demand**. ***The law of demand states that there is an inverse (or negative) relationship between the price of a good or service and the quantity of it that consumers are willing to purchase.*** This inverse relationship means that price and the quantity consumers wish to purchase move in opposite directions. As the price increases, buyers purchase less—and as the price decreases, buyers purchase more.

The availability of **substitutes**—goods that perform similar functions—helps explain this inverse relationship. No single good is absolutely essential; everything can be replaced with something else. A chicken sandwich can be substituted for a cheeseburger. Wood, aluminum, bricks, and glass can take the place of steel. Going to the movies, playing tennis, watching television, and going to a football game are substitute forms of entertainment. When the price of a good increases, people cut back on it and buy substitute products.

The Market Demand Schedule

The lower portion of **Exhibit 1** shows a hypothetical *demand schedule* for cellular telephone service.[2] **A** demand schedule is simply a table listing the various quantities of something consumers are willing to purchase at different prices. In Exhibit 1, notice that the price is the average monthly cost of purchasing cellular phone service. The quantity demanded is the number of people willing to subscribe to cellular service at each price. When the price of cell phone service is $143 per month, just over 2 million people subscribe. As the price falls to $85, the quantity of subscribers rises to 11 million; when the price falls to $41 per month, the quantity of subscribers increases to just over 69 million.

The upper portion of Exhibit 1 shows what the demand schedule would look like if the various prices and corresponding quantity of subscribers were plotted on a graph and connected by a line. This is called the *demand curve.* When representing the demand schedule graphically, economists measure price on the vertical or y-axis, and the amount demanded on the horizontal or x-axis. Because of the inverse relationship between price and amount purchased, the demand curve will have a negative slope—that is, it will slope downward to the right. More of a good will be purchased as its price decreases. This is the law of demand.

Read horizontally, the demand curve shows how much of a particular good consumers will buy at a given price. Read vertically, the demand curve shows how much consumers value the good. ***The height of the demand curve at any quantity shows the maximum price consumers are willing to pay for an additional unit.*** If consumers value highly an additional unit of a product, they will be willing to pay a large amount for it. Conversely, if they place a low value on the additional unit, they will be willing to pay only a small amount for it.

Because the amount a consumer is willing to pay for a good is directly related to the good's value to them, the demand curve indicates the marginal benefit (or value)

[2]These data are actual pnces (adjusted to 2000 dollars) and quantities annually for 1988 to 1998 taken from *Statistical Abstract of the United States* (Washington, D C U S Bureau of the Census, various years) *If we could assume that other demand determinants (income,* **prices** *of related goods, etc) had remained constant,* then this hypothetical demand schedule would be accurate for that time period Because it is possible that some of these other factors changed, we treat the numbers as hypothetical, depicting alternative pnces and quantities *at a given time*

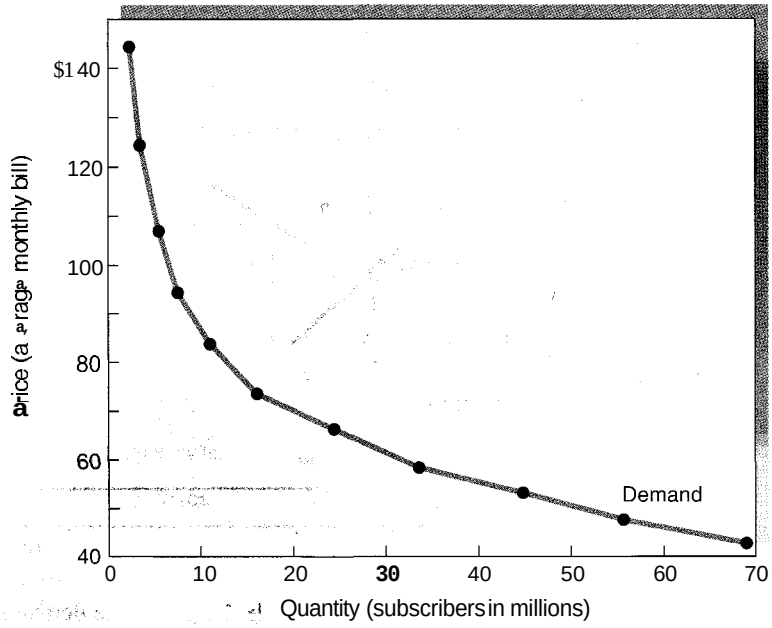


EXHIBIT 1
Law of Demand

As the demand schedule shown in the table indicates, the number of people subscribing to cellular phone service (just like the consumption of other products) is inversely related to price. The data from the table are plotted as a demand curve in the graph. The inverse relation between price and amount demanded reflects the fact that consumers will substitute away from a good as it becomes more expensive.

CELLULAR PHONE PRICE (AVERAGE MONTHLY BILL)	QUANTITY OF CELLULAR PHONE SUBSCRIBERS (IN MILLIONS)
$143	2.1
124	3.5
107	5.3
92	7.6
85	11.0
73	16.0
65	24.1
58	33.7
53	44.0
46	55.3
41	69.2

consumers receive from additional units. (Recall that we briefly discussed marginal benefit in Chapter 1.)When viewed in this manner, the demand curve reveals that as consumers have more and more of a good or service, they value additional units less and less.

Consumer Surplus

Previously, we indicated that voluntary exchanges make both buyers and sellers better off. The demand curve can be used to illustrate the gains to consumers. Suppose you value a particular good at $50, but you are able to purchase it for only $30. Your net gain from buying the good is the $20 difference. Economists call this net gain of buyers **consumer surplus**. Consumer surplus is simply the difference between the maximum amount consumers would be willing to pay and the amount they actually pay for a good.

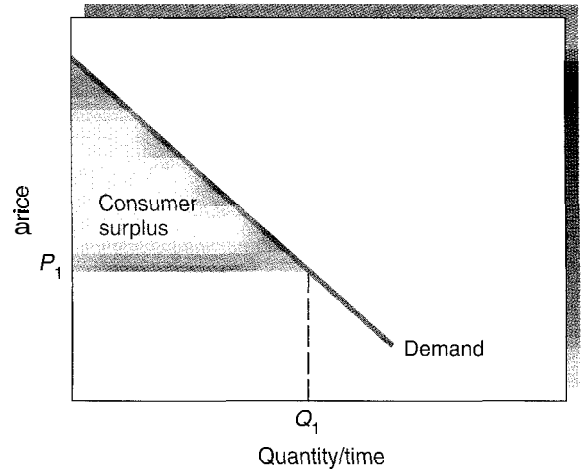
Consumer surplus
The difference between the maximum price consumers are willing to pay and the price they actually pay. It is the net gain derived by the buyers of the good.

Exhibit 2 shows the consumer surplus for an entire market. The height of the demand curve measures how much buyers in the market value each unit of the good. The price indicates the amount they actually pay. The difference between these two—the triangular area below the demand curve but above the price paid—is a measure of the total consumer surplus generated by all exchanges of the good. The size of the consumer surplus, or triangular area, is affected by the market price. If the market price for the goods falls, more of it will be purchased, resulting in a larger surplus for consumers. Conversely, if the market price rises, less of it will be purchased, resulting in a smaller surplus (net gain) for consumers.

Because the value a consumer places on a particular unit of a good is shown by the corresponding height of the demand curve, we can use the demand curve to clarify the

EXHIBIT 2
Consumer Surplus

Consumer surplus is the area below the demand curve but above the actual price paid. This area represents the net gains to buyers from market exchange.



difference between the *marginal value* and *total value* of a good—a distinction we introduced briefly in Chapter 1. Returning to Exhibit 2, if consumers are currently purchasing Q, units, the marginal value of the good is indicated by the height of the demand curve at Q_1—the last unit consumed (or purchased). So at each quantity, the height of the demand curve shows the marginal value of that unit, which as you can see, declines along a demand curve. The *total* value of the good, however, is equal to the combined value of all units purchased. This is the sum of the value of each unit (the heights along the demand curve) on the x-axis, out to, and including, unit Q,. This total value is indicated graphically as the entire area under the demand curve out to Q, (the triangular area representing consumer surplus *plus* the unshaded rectangular area directly below it).

You can see that the total value to consumers of a good can be far greater than the marginal value of the last unit consumed. When additional units are available at a low price, the marginal value of a good may be quite low, even though its total value to consumers is exceedingly high. This is usually the case with water, for example, because it is essential for life. The value of the first few units of water consumed per day will be exceedingly high. The consumer surplus derived from these units will also be large when water is plentiful at a low price. As more and more units are consumed, however, the *marginal value* of even something as important as water will fall to a low level. Thus, when water is cheap, people will use it not only for drinking, cleaning, and cooking, but also for washing cars, watering lawns, flushing toilets, and maintaining fish aquariums. Thus, although the total value of water is rather large, its marginal value is quite low.

Consumers will tend to expand their consumption of a good until its price and *mayginal value* are equal (which occurs at Q, in Exhibit 2 at a price of P_1). Thus, the price of a good (which equals marginal value) reveals little about the *total value* derived from the consumption of it. This is the reason that the market price of diamonds (which reflects their high marginal value) is greater than the market price of water (which has a low marginal value), even though the total value of diamonds is far less than the total value of water. Think of it this way, beginning from your current levels of consumption, if you were offered a choice between one diamond or one gallon of water right now, which would you take? You would probably take the diamond, because at the margin it has more value to you than additional water. However, if given a choice between giving up *all* of the water you use or *all* of the diamonds you have, you would probably keep the water over diamonds, because in total water has more value to you.

Responsiveness of Quantity Demanded to Price Changes

Elastic and Inelastic Demand Curves

As we previously noted, the availability of substitutes is the main reason why the demand curve for a good slopes downward. Some goods, however, are much easier than others to substitute away from. As the price of tacos rises, most consumers find hamburgers a reasonable substitute. Because of the ease of substitutability, the quantity of tacos

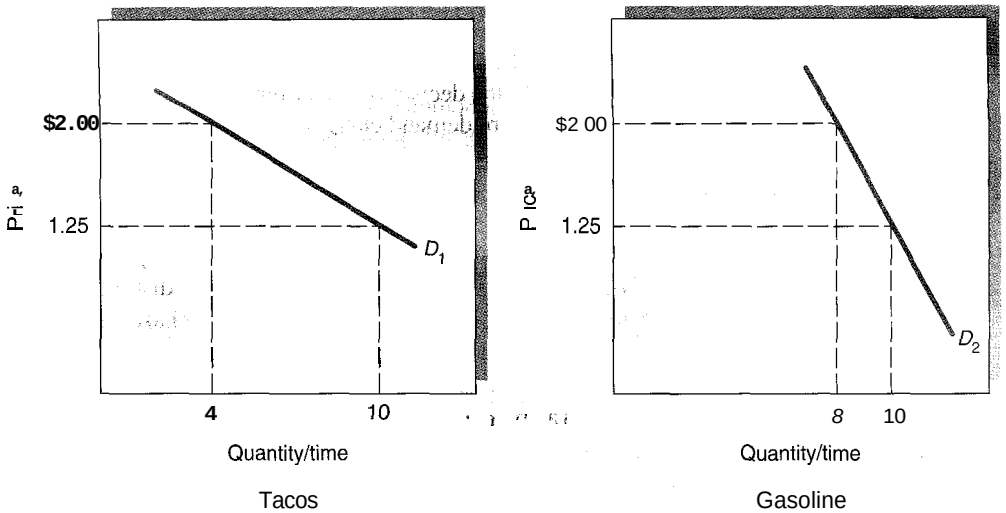


EXHIBIT 3
Elastic and Inelastic Demand Curves

The responsiveness of consumer purchases to a change in price is reflected in the steepness of the demand curve. The flatter demand curve (D_1) for tacos shows a higher degree of responsiveness and is called relatively elastic, while the steeper demand curve (D_2) for gasoline shows a lower degree of responsiveness and is called relatively inelastic.

demanded is quite sensitive to a change in their price. Economists would say that the demand for tacos is relatively *elastic* because a small price change will cause a rather large change in the amount purchased. Alternatively, goods like gasoline and electricity have fewer close substitutes. When their prices rise, it is harder for consumers to find substitutes for these products. When close substitutes are unavailable, even a large price change may not cause much of a change in the quantity demanded. In this case, an economist would say that the demand for such goods is relatively *inelastic.*

Graphically, this different degree of responsiveness is reflected in the steepness of the demand curve, as shown in **Exhibit 3.** The flatter demand curve (D_1, left frame) is for a product like tacos, for which the quantity purchased is highly responsive to a change in price. As the price increases from $1.25 to $2.00, the quantity demanded falls sharply from 10 to 4 units. The steeper demand curve (D_2, right frame) is for a product like gasoline, where the quantity purchased is much less responsive to a change in price. For gasoline, an increase in price from $1.25 to $2.00 results in only a small reduction in the quantity purchased (from 10 to 8 units). An economist would say that the flatter demand curve D_1 is "relatively elastic," whereas the steeper demand curve D_2 is "relatively inelastic." The availability of substitutes is the main determinant of a product's elasticity or inelasticity and thus how flat or steep its demand curve is.

What would a demand curve that was perfectly vertical represent? Economists refer to this as a "perfectly" inelastic demand curve. It would mean that the quantity demanded of the product never changes—regardless of its price. Although it is tempting to think that the demand curves are vertical for goods essential to human life (or goods that are addictive), this is inaccurate for two reasons. First, in varying degrees, there are substitutes for everything. As the price of a good rises, the incentive increases for suppliers to invent even more new substitutes. Thus, even for goods that currently have few substitutes, if the price were to rise high enough, alternatives would be invented and marketed, reducing the quantity demanded of the original good. Second, our limited incomes restrict our ability to afford goods when they become very expensive. As the price of a good rises to higher and higher levels, if we do not cut back on the quantity purchased, we will have less and less income to spend on other things. Eventually, this will cause us to cut back on our purchases of it. Because of these two reasons, the demand curve for every good will slope downward to the right.

CHANGES IN DEMAND VERSUS CHANGES IN QUANTITY DEMANDED

The purpose of the demand curve is *to* show what effect a price change will nave on the quantity demanded (or purchased) of a good. Economists refer to a change in the quantity of a good purchased in response solely to a price change as a "change in *quantity*

demanded." A change in quantity demanded is simply a movement along a demand curve from one point to another.

Changes in factors other than a good's price—such as consumers' income and the prices of closely related goods—will also influence the decisions of consumers to purchase a good. If one of these other factors changes, the entire demand curve will *shift* inward or outward. Economists refer to a shift in the demand curve as a "change in demand."

Failure to distinguish between a change in demand and a change in quantity demanded is one of the most common mistakes made by beginning economics students? A change in demand is a shift in the entire demand curve. A change in quantity demanded is a movement along the same demand curve. The easiest way to distinguish between these two concepts is the following: If the change in consumer purchases is caused by a change in the price of the good, it is a change in quantity demanded—a movement along the demand curve. If the change in consumer purchases is due to a change in anything other than the price of the good (a change in consumer income, for example), it is a change in demand—a shift in the demand curve.

Let us now take a closer look at some of the factors that cause a "change in demand"—an inward or outward shift in the entire demand curve.

1. Changes in consumer income. An increase in consumer income makes it possible for consumers to purchase more goods. If you were to win the lottery, or if your boss were to give you a raise, you would respond by increasing your spending on many products. Alternatively, when the economy goes into a recession, falling incomes and rising unemployment cause consumers to reduce their purchases of many items. A change in consumer income will result in consumers buying more or less of a product at all possible prices. When consumer income increases, in the case of most goods, individuals will purchase more of the good even if the price is unchanged. This is shown by a shift to the right—an outward shift—in the demand curve. Such a shift is called an increase in demand. A reduction in consumer income generally causes a shift to the left—an inward shift—in the demand curve, which is called a decrease in demand. Note that the appropriate terminology here is an increase or decrease in demand, not an increase or decrease in quantity demanded.

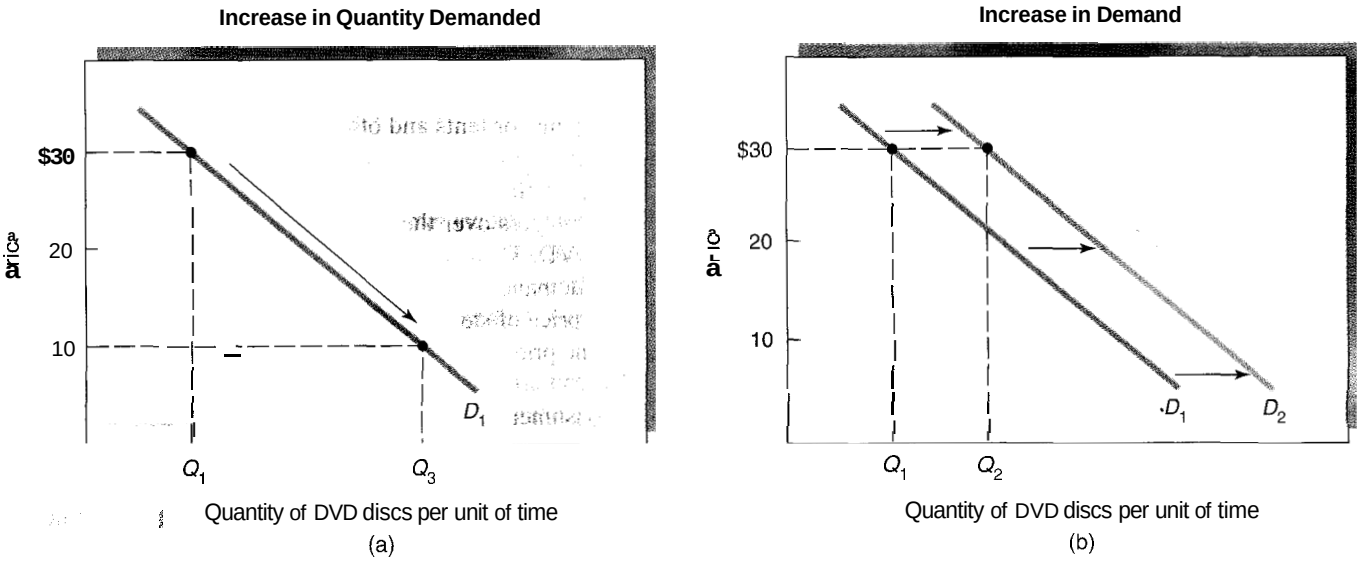
Exhibit 4 highlights the difference between a change in demand and a change in quantity demanded. The demand curve D, indicates the initial demand curve for DVDs. At a price of $30, consumers will purchase Q_1 units. If the price were to decline to $10, the quantity demanded would increase from Q, to Q_3. The arrow in panel (a) indicates the change in quantity demanded—a movement along the original demand curve D_1 in response to the change in price. Now, alternatively suppose there were an increase in income that caused the demand for DVDs to shift from D_1 to D_2. As indicated by the arrows in panel (b), the entire demand curve would shift outward. At the higher income level, consumers would be willing to purchase more DVDs than before. This is true at a price of $30, $20, $10, and every other price. The increase in income leads to an increase in *demand*—a shift in the entire curve.

2. Changes in the number of consumers in the market. Businesses that sell products in college towns are greatly saddened when summer arrives. As you might expect in these towns, the demand for many items—from pizza delivery to beer—falls during the summer. **Exhibit 5** shows how the falling number of consumers in the market caused by students going home for the summer affects the demand for pizza delivery. With fewer customers, the demand curve shifts inward from D_1 to D_2. There is a decrease in demand; pizza stores sell fewer pizzas than before regardless of what price they originally charged. Had their original price been $20, then demand would fall from 200 pizzas per week to only 100. Alternatively, had their original price been $10, then demand would fall from 300 pizzas to 200. When autumn arrives and the students come back to town, there will be

[3]Questions designed to test the ability of students to inake this distinction are favorite? of many economics instructors **A** word to the wise should be sufficient

EXHIBIT 4
Change in Demand Versus Change in Quantity Demanded

Panel (a) shows a change in *quantity demanded,* a movement along the demand curve D_1, in response to a change in the price of DVD discs. Panel (b) shows a change in *demand,* a shift of the entire curve. in this case due to an increase in consumer income.



an increase in demand that will restore the curve to about its original position. As cities grow and shrink, and as international markets open up to domestic firms, changes in the number of consumers affect the demand for many products.

3. Changes in the price of a related good. Changes in prices of closely related products also influence the choices of consumers. Related goods may be either substitutes or complements. When two products perform similar functions or fulfill similar needs, they are substitutes. Economists define goods as substitutes when there is a direct relationship between the price of one and the demand for the other—meaning an increase in the price of one leads to an increase in demand for the other (they move in the same direction). For example, margarine is a substitute for butter. If the price of butter rises, it will increase the demand for margarine as consumers substitute margarine for the more expensive butter. Similarly, lower butter prices will reduce the demand for margarine, shifting the entire

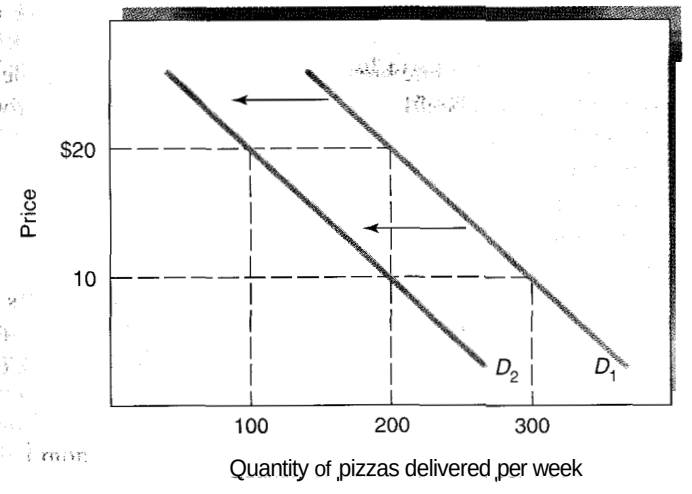


EXHIBIT 5
A Decrease in Demand

In college towns the demand for pizza delivery decreases substantially when students go home for the summer. A decrease in demand is a leftward shift in the entire demand curve. Fewer pizzas are demanded at every price.

demand curve for margarine to the left. A substitute relationship exists between beef and chicken, pencils and pens, apples and oranges, coffee and tea, and so forth.

Note that although a change in the price of butter shifts the demand curve for margarine (a change in demand), it will only result in a movement along the demand curve for butter (a change in the quantity demanded). The reason is that the demand curve for butter already shows the relationship between the price of butter and the quantity of butter desired. An increase in the price of butter makes consumers willing to purchase more margarine, holding constant the price of margarine.

Other products are consumed jointly, so the demands for them are linked together as well. Examples of goods that "go together" include peanut butter and jelly, DVDs and DVD players, hot dogs and hot dog buns, or tents and other camping equipment. These goods are called **complements**. For complements, a decrease in the price of one will not only increase its quantity demanded, it will also increase the demand for the other good. For example, lower prices for DVD players over the past decade have substantially increased the demand for movies on DVD. The reverse is also true. As a complement becomes more expensive, the quantity demanded of it will fall, and so will the demand for its complements. For example, if the price of steak rises, grocery stores can expect to sell fewer bottles of steak sauce, even if the price of steak sauce remains unchanged.

Complements
Products that are usually consumed jointly (for example, bread and butter, hot dogs and hot dog buns). A decrease in the price of one will cause an increase in demand for the other.

4. Changes in expectations. Consumers' expectations about the future also can affect the current demand for a product. If consumers begin to expect that a major hurricane will strike their area, the current demand for batteries and canned food will rise. Expectations about the future direction of the economy can also affect current demand. If consumers become pessimistic about the economy, they might start spending less, causing the current demand for goods to fall. Perhaps most important is how a change in the expected future price of a good affects current demand. When consumers expect the price of a product to rise in the future, their current demand for it will increase. Gasoline is a good example. If you expect the price to increase soon, you'll want to fill up your tank now before the price goes up. On the other hand, consumers will delay a purchase if they expect the item to decrease in price. No doubt you have heard someone say, "I'll wait until it goes on sale." When consumers expect the price of a product to fall, current demand for it will decline.

5. Demographic changes. The demand for many products is strongly influenced by the demographic composition of the market. An increase in the elderly population in the United States in recent years has increased the demand for medical care, retirement housing, and vacation travel. The demand curves for these goods have shifted to the right. During the 1980s, the number of people ages 15–24 fell by more than 5 million. Because young people are a major part of the U.S. market for jeans, the demand for jeans fell by more than 100 million pairs over the course of the decade.[4]

6. Changes in consumer tastes and preferences. Why do preferences change? Preferences change because people change and because they acquire new information. Consider how consumers are responding to the popularity of the Atkins diet. The demand for high-carbohydrate foods like white bread has fallen substantially, while the demand for low-carbohydrate foods like beef has risen. This is a major change from the 1990s when the demand for beef fell because of the "heart-healthy'' eating habits consumers preferred then. Trends in the markets for clothing, toys, collectibles, and entertainment are constantly causing changes in the demand for these products as well. Firms may even try to change consumer preferences for their own products through advertising and information brochures.

The accompanying **Thumbnail Sketch** summarizes the major factors that cause a change in demand—a shift of the entire demand curve—and points out that quantity demanded (but not demand) will change in response to a change in the price of a good.

[4]These figures are from Suzanne Tregarthen, "Market for Jeans Shrink?," *The Margin* 4, no 3 (January–February 1991) 28

THUMBNAIL SKETCH

Factors That Cause Changes in Demand and Quantity Demanded

This factor changes the quantity demanded of a good:

1. The price of the good: A higher price decreases the quantity demanded; a lower price increases the quantity demanded.

These factors change the demand for a good:

1. Consumer income: Lower consumer income decreases demand; higher consumer income increases demand.
2. Number of consumers in the market: Fewer consumers decreases demand; more consumers increases demand.

3a. Price of a substitute good: A decrease in the price of a substitute decreases the demand for the original good; an increase in the price of a substitute increases the demand for the original good.

3b. Price of a complementary good: An increase in the price of a complement decreases the demand for the original *good;* a decrease in the price of a complement increases the demand for the original good.

4. Expected future price of the good: If the price of a good is expected to fall in the future, the current demand for it will decrease; if the price of a good is expected to rise in the future, the current demand for it will increase.
5. Demographic changes: Population trends in age, gender, race, and other factors can increase or decrease demand for specific goods.
6. Consumer preferences: Changes in consumer tastes and preferences can increase or decrease demand for specific goods.

PRODUCER CHOICE AND THE LAW OF SUPPLY

Now let's shift our focus to producers and the supply side of the market. How does the market process determine the amount of each good that will be produced? To figure this out, we first have to understand what influences the choices of producers. Producers convert resources into goods and services by doing the following:

1. organizing productive inputs and resources, like land, labor, capital, natural resources, and intermediate goods;
2. transforming and combining these inputs into goods and services; and
3. selling the final products to consumers.

Producers have to purchase the resources at prices determined by market forces. Predictably, the owners of these resources will supply the resources only at prices at least equal to what they could earn elsewhere. Put another way, each resource the producers buy to make their product has to be bid away from all other potential uses. Its owner has to be paid its opportunity cost. ***The sum of the producer's cost of each resource used to produce a good will equal the* opportunity cost of production.**

There is an important difference between the opportunity cost of production and standard accounting measures of cost. Accountants generally do not count the cost of the firm's assets, such as its buildings, equipment, and financial resources, when they calculate a product's cost. But economists do. Economists consider the fact that these assets could be used some other way—in other words, that they have an opportunity cost. Unless these opportunity costs are covered, the resources will eventually be used in other ways.

The opportunity cost of these assets to the firm is the amount of money the firm could earn from the assets if they were used another way. Consider a manufacturer that invests $10 million in buildings and equipment to produce shirts. Instead of buying buildings and equipment, the manufacturer could simply put the $10 million in the bank and let it draw interest. If the $10 million were earning, say, 10 percent interest, the firm would make $1 million on that money in a year's time. This $1 million in forgone interest is part of the firm's opportunity cost of producing shirts. Unlike an accountant, an economist will take

Opportunity cost of production
The total economic cost of producing a good or service. The cost component includes the opportunity cost of all resources, including those owned by the firm. The opportunity cost is equal to the value of the production of other goods sacrificed as the result of producing the good.

that $1 million opportunity cost into account. If the firm plans to invest the money in shirt-making equipment, it had better earn more from making the shirts than the $1 million it could earn by simply putting the money in the bank. If the firm can't generate enough to cover all of its costs, including the opportunity cost of assets owned by the firm, it will not continue in business.

The Role of Profits and Losses

Profits direct producers toward activities that increase the value of resources; losses impose a penalty on those who reduce the value of resources.

Profits and Losses

Profit
An excess of sales revenue relative to the opportunity cost of production. The cost component includes the opportunity cost of all resources, including those owned by the firm. Therefore, profit accrues only when the value of the good produced is greater than the value of the resources used for its production.

Firms earn a **profit** when the revenues from the goods and services that they supply exceed the opportunity cost of the resources used to make them. Consumers will not buy goods and services unless they value them at least as much as their purchase price. For example, Susan would not be willing to pay $40 for a pair of jeans unless she valued them by at least that amount. At the same time, the seller's opportunity cost of supplying a good will reflect the value consumers place on *other* goods that could have been produced with those same resources. This is true precisely because the seller has to bid those resources away from other producers wanting to use them.

Think about what it means when, for example, a firm is able to produce jeans at a cost of $30 per pair and sell them for $40, thereby reaping a profit of $10 per pair. The $30 opportunity cost of the jeans indicates that the resources used to produce the jeans could have been used to produce other items worth $30 to consumers (perhaps a denim backpack). In turn, the profit indicates that consumers value the jeans more than other goods that might have been produced with the resources used to supply the jeans.

The willingness of consumers to pay a price greater than a good's opportunity cost indicates that they value the good more than other things that could have been produced with the same resources. Viewed from this perspective, profit is a reward earned by entrepreneurs who use resources to produce goods consumers value more highly than the other goods those resources could have produced. In essence, this profit is a signal that an entrepreneur has increased the value of the resources under his or her control.

Loss
A deficit of sales revenue relative to the opportunity cost of production. Losses are a penalty imposed on those who produce goods even though they are valued less than the resources required for their production.

Business decision makers will seek to undertake production of goods and services that will generate profit. However, things do not always turn out as expected. Sometimes business firms are unable to cover their costs. **Losses** occur when the revenue derived from sales is insufficient to cover the opportunity cost of the resources used to produce a good or service. Losses indicate that the firm has reduced the value of the resources it has used. In other words, consumers would have been better off if those resources had been used to produce something else. In a market economy, losses will eventually cause firms to go out of business, and the resources they previously utilized will be directed toward other things valued more highly.

Profits and losses play a very important role in a market economy. They determine which products (and firms) will expand and survive, and which will contract and be driven from the market. Clearly, there is a positive side to business failures. As our preceding discussion highlights, losses and business failures free up resources being used unwisely so they can be put to a use producing other things that people value more highly.

Supply and the Entrepreneur

Entrepreneurs organize the production of new products. In doing so, they take on significant risk in deciding what to produce and how to produce it. Their success or failure depends upon how much consumers eventually value the products they develop relative to other products that could have been produced with the resources. Entrepreneurs figure out which projects are likely to be profitable and then try to persuade a corporation, a banker, or individual investors to invest the resources needed to give their new idea a

chance. Studies indicate, however, that only about 55 to 65 percent of the new products introduced are still on the market five years later. Being an entrepreneur means you have to **risk** failing.

To prosper, entrepreneurs must convert and rearrange resources in a manner that will increase their value. A person who purchases 100 acres of raw land, puts in a street and a sewage-disposal system, divides the plot into 1-acre lots, and sells them for 50 percent more than the opportunity cost of all resources used is clearly an entrepreneur. This entrepreneur profits because the value of the resources has increased. Sometimes entrepreneurial activity is less complex, though. For example, a 15-year-old who purchases a power mower and sells lawn services *to* his neighbors is also an entrepreneur seeking to profit by increasing the value of his resources—time and equipment.

Market Supply Schedule

How will producer-entrepreneurs respond to a change in product price? Other things constant, a higher price will increase the producer's incentive to supply the good. Established producers will expand the scale of their operations, and over time new entrepreneurs, seeking personal gain, will enter the market and begin supplying the product, too. ***The* law of supply *states that there is a direct (or positive) relationship between the price of a good or service and the amount of it that suppliers are willing to produce. This direct relationship means that price and the quantity producers wish to supply move in the same direction. As the price increases, producers will supply more—and as the price decreases, they will supply less.***

Law of supply
A principle that states there is a direct relationship between the price of a good and the quantity of it producers are willing to supply. As the price of a good increases, producers will wish to supply more of it. As the price decreases, producers will wish to supply less.

Like the law of demand, the law of supply reflects the basic economic principle that incentives matter. Higher prices increase the reward entrepreneurs get from selling their products. The more profitable producing a product becomes, the more of it they will be willing to supply. Conversely, as the price of a product falls, so does its profitability and the incentive to supply it. Just think about how many hours of tutoring services you would be willing to supply for different prices. Would you be willing to spend more time tutoring students if instead of $5 per hour, tutoring paid $50 per hour? The law of supply suggests you would, and producers of other goods and services are no different.

Exhibit 6 illustrates the law of supply. The curve shown in the exhibit is called a *supply curve*. Because there is a direct relationship between a good's price and the amount offered for sale by suppliers, the supply curve has a positive slope. It slopes upward to the right. Read horizontally, the supply curve shows how much of a particular good producers are willing to produce and sell at a given price. Read vertically, the supply curve reveals important information about the cost of production. ***The height of the supply curve indicates both (I) the minimum price necessary to induce producers to supply that***

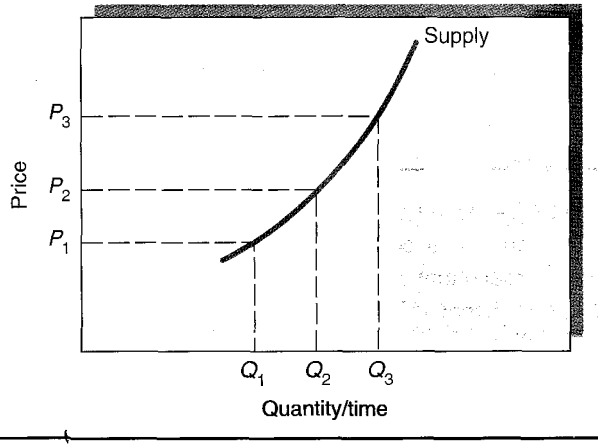
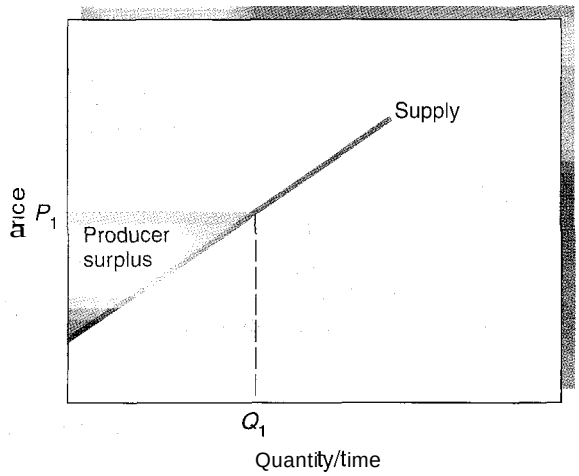


EXHIBIT 6
Supply Curve

As the price of a product increases, other things constant, producers *will* increase the amount of the product supplied to the market.

EXHIBIT 7
Producer Surplus

Producer surplus is the area above the supply curve but below the actual sales price. This area represents the net gains to producers and resource suppliers from production and exchange.



additional unit and (2) the opportunity cost of producing that additional unit. These are both measured by the height of the supply curve because the minimum price required to induce a supplier to sell a unit is precisely the marginal cost of producing it.

Producer Surplus

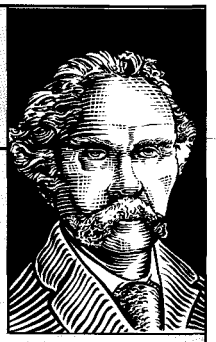
We previously used the demand curve to illustrate consumer surplus, the net gains of buyers from market exchanges. The supply curve can be used in a similar manner to illustrate the net gains of producers and resource suppliers. Suppose that you are an aspiring musician and are willing to perform a two-hour concert for $500. If a promoter offers to pay you $750 to perform the concert, you will accept, and receive $250 more than your minimum price. This $250 net gain represents your **producer surplus**. In effect, producer surplus is the difference between the amount a supplier actually receives (based on the market price) and the minimum price required to induce the supplier to produce the given units (their marginal cost). The measurement of producer surplus for an entire market is illustrated by the shaded area of **Exhibit 7.**

Producer surplus
The difference between the minimum price suppliers are willing to accept and the price they actually receive. It measures the net gains to producers and resource suppliers from market exchange. It is not the same as profit.

It's important to note that producer surplus represents the gains received by all parties contributing resources to the production of a good. In this respect, producer surplus is fundamentally different from profit. Profit accrues to the owners of the business firm producing the good, whereas producer surplus encompasses the net gains derived by all people who help produce the good, including those employed by or selling resources to the firm.

OUTSTANDING ECONOMIST

Alfred Marshall (1842–1924)



British economist Alfred Marshall was one of the most influential economists of his era. Many concepts and tools that form the core of modern microeconomics originated with Marshall in his famous *Principles of Economics,* first published in 1890. Marshall introduced the concepts of supply and demand, equilibrium, elasticity, consumers' and producers' surplus, and the idea of distinguishing between short-run and long-run changes.

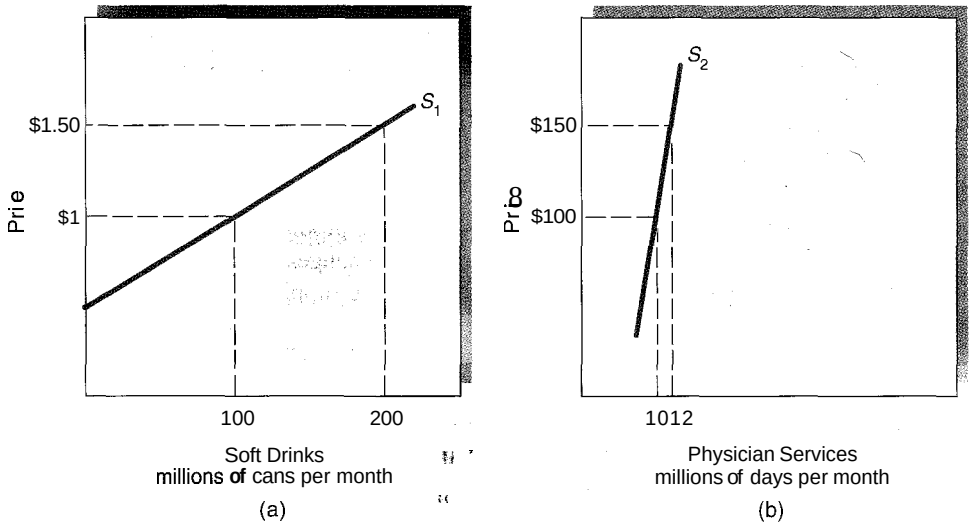


EXHIBIT 8
Elastic and Inelastic Supply Curves

Frame (a) illustrates a supply curve that is relatively elastic and therefore the quantity supplied is highly responsive to a change in price. Soft drinks provide an example. Frame (b) illustrates a relatively inelastic supply curve, one where the quantity supplied increase is by only a small amount in response to a change in price. This is the case for physician services.

Responsiveness of Quantity Supplied to Price Changes: Elastic and Inelastic Supply Curves

Like the quantity demanded, the responsiveness of the quantity supplied to a change in price is different for different goods. The supply curve is said to be elastic when a modest change in price leads to a large change in quantity supplied. This is generally true when the additional resources needed to expand output can be obtained with only a small increase in their price. Consider the supply of soft drinks. The contents of soft drinks—primarily carbonated water, sugar, and flavoring—are abundantly available. **A** sharp increase in the use of these ingredients by soft drink producers is unlikely to push up their price much. Therefore, as **Exhibit 8** illustrates, if the price of soft drinks were to rise from $1 to $1.50, producers would be willing to expand output sharply from 100 million to 200 million cans per month. **A** 50 percent increase in price leads to a 100 percent expansion in quantity supplied. The larger the increase in quantity in response to a higher price, the more elastic the supply curve. The flatness of the supply curve for soft drinks reflects the fact that it is highly elastic.

In contrast, when the quantity supplied is not very responsive to a change in price, supply is said to be inelastic. Physicians' services are an example. If the earnings of doctors increase from $100 to $150 per hour, there will be some increase in the quantity of the services they provide. Some physicians will work longer hours; others may delay retirement. Yet, these adjustments are likely to result in only a small increase in the quantity supplied because it takes a long time to train a physician and the number of qualified doctors who are working in other occupations or who are outside of the labor force is small. Therefore, as Exhibit 8 (right frame) shows, a 50 percent increase in the price of physician services leads to only a 20 percent expansion in the quantity supplied. Unlike soft drinks, higher prices for physician services do not generate much increase in quantity supplied. Economists would say that the supply of physician services is relatively inelastic.

CHANGES IN SUPPLY VERSUS CHANGES IN QUANTITY SUPPLIED

Like demand, it is important to distinguish between a change in the *quantity supplied* and a change in *supply*. When producers change the number of units they are willing to supply in response to a change in price, this movement along the supply curve is called a "change in *quantity supplied.*" **A** change in any factor *other than the price* shifts the supply curve and is called a "change in *supply.*"

As we previously discussed, profit-seeking entrepreneurs will produce a good only if its sales price is expected to exceed its opportunity cost of production. Therefore, changes that affect the opportunity cost of supplying a good will also influence the amount of it producers are willing to supply. These other factors, such as the prices of resources used to make the good and the level of technology available, are held constant when we draw the supply curve. The supply curve itself reflects quantity changes only in response to price changes. Changes in these other factors shift the supply curve. Factors that increase the opportunity cost of providing a good will discourage production and decrease supply, shifting the entire curve inward to the left. Conversely, changes that lower the opportunity cost of producers will encourage production and increase supply, shifting the entire curve outward to the right.

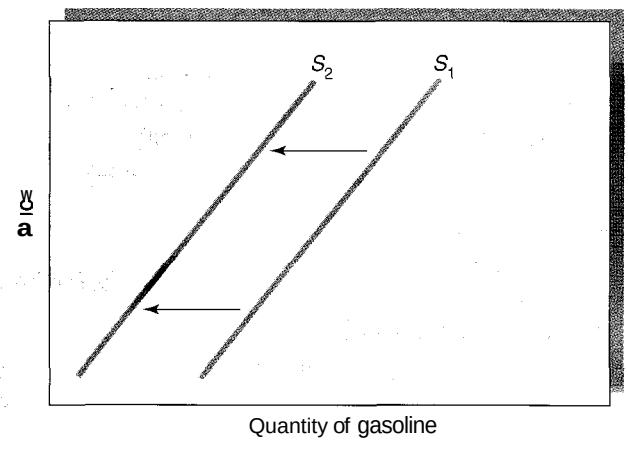
Let us now take a closer look at the primary factors that will cause a change in supply and shift the entire curve right or left.

1. Changes in resource prices. How will an increase in the price of a resource, such as wages of workers or the materials used to produce a product, affect the supply of a good? Higher resource prices will increase the cost of production, reducing the profitability of firms supplying the good. The higher cost will induce firms to reduce their output. With time, some may even be driven out of business. As **Exhibit 9** illustrates, higher resource prices will reduce the supply of the good, causing a shift to the left in the supply curve from S_1 to S,. Alternatively, a reduction in the price of a resource used to produce a good will cause an increase in supply—a rightward shift in the supply curve—as firms expand output in response to the lower costs and increased profitability of supplying the good.

2. Changes in technology. Like lower resource prices, technological improvements—the discovery of new, lower-cost production techniques—reduce production costs, and thereby increase supply. Technological advances have affected the cost of almost everything. Before the invention of the printing press, books had to be handwritten. Just imagine the massive reduction in cost and increase in the supply of books caused by this single invention. Similarly, improved farm machinery has vastly expanded the supply of agricultural products through the years. Robotics have reduced the cost of producing airplanes, automobiles, and other types of machinery. Better computer chips have drastically reduced the cost of producing electronics. As recently as 35 years ago, a simple calculator cost more than $100; a microwave oven almost $500; and a VCR approximately $1,000. When introduced in the mid-1980s, a cellular telephone cost more than $4,000. You probably noticed that the prices of flat-screen computer monitors and plasma-screen televisions have fallen due to technological advances in recent years.

EXHIBIT 9
A Decrease in Supply

Crude oil is a resource used to produce gasoline. When the price of crude oil rises, it increases the cost of producing gasoline and results in a decrease in the supply of gasoline.



THUMBNAIL SKETCH

Factors That Cause Changes in Supply and Quantity Supplied

This factor changes the quantity supplied of a good:

1. The price of the good: A lower price decreases the quantity supplied; a higher price increases the quantity supplied.

These factors change the supply of a good:

1. Resource prices (the prices of things used to make the good): Lower resource prices increase supply; higher resource prices decrease supply.
2. Technological change: A technological improvement increases supply; a technological setback decreases supply.
3. Weather or political conditions: Favorable weather or good political conditions increase supply; adverse weather conditions or poor political conditions decrease supply.
4. Taxes imposed on the producers of a good: Lower taxes increase supply; higher taxes decrease supply.

3. Elements of nature and political disruptions. Natural disasters and changing political conditions can also alter supply, sometimes dramatically. In some years, good weather leads to "bumper crops," increasing the supply of agricultural products. At other times, droughts lead to poor harvests, reducing supply. War and political unrest in the Middle East region have had a major impact on the supply of oil several times during the past few decades. Similarly, during the summer of 2003, massive U.S. electricity blackouts in the Midwest and along the East Coast temporarily shut down oil refineries, reducing the supply of oil. Factors such as these will reduce supply.

4. Changes in taxes. If the government increases the taxes on the sellers of a product, the result will be the same as any other increase in the cost of doing business. The added tax that sellers have to pay will reduce their willingness to sell the product at any given price. Each unit must now be sold for a price that covers not only the opportunity cost of production, but also the tax. For example, the Superfund law, passed by Congress in 1980, placed a special tax on petroleum producers based on their output. That raised the cost of producing petroleum products, decreasing the amount producers were willing to supply.

The accompanying **Thumbnail Sketch** summarizes the major factors that change *supply—a* shift of the entire supply curve; and quantity supplied—a movement along the supply curve.

HOW MARKET PRICES ARE DETERMINED: SUPPLY AND DEMAND INTERACT

Consumer-buyers and producer-sellers make decisions independent of each other, but markets coordinate their choices and influence their actions. To the economist, a **market** is not a physical location but an abstract concept that encompasses the forces generated by the decisions of buyers and sellers. A market may be quite narrow (for example, the market for grade A jumbo eggs), or it may be quite broad like when we lump diverse goods into a single market, such as the market for all "consumer goods." There is also a wide range of sophistication among markets. The New York Stock Exchange is a highly formal, computerized market. Each weekday, buyers and sellers, who seldom meet, electronically exchange corporate shares they own worth billions of dollars. In contrast, a neighborhood market for baby-sitting services or tutoring in economics, may be highly informal, bringing together buyers and sellers primarily by word of mouth.

Equilibrium ***is a state in which the conflicting forces of supply and demand are in balance. When a market is in equilibrium, the decisions of consumers and producers***

Market
An abstract concept encompassing the forces of supply and demand, and the interaction of buyers and sellers with the potential for exchange to occur.

Equilibrium
A state in which the conflicting forces of supply and demand are in balance. When a market is in equilibrium, the decisions of consumers and producers are brought into harmony with one another, and the quantity supplied will equal the quantity demanded.

are brought into harmony with one another, and the quantity supplied will equal the quantity demanded. In equilibrium, it is possible for both buyers and sellers to realize their choices simultaneously. What could bring these diverse interests into harmony? We will see the answer is market prices.

Market Equilibrium

As we have learned, a higher price will reduce the quantity of a good demanded by consumers. On the other hand, a higher price will increase the quantity of a good supplied by producers. The market price of a good will tend to change in a direction that will bring the quantity of a good consumers want to buy into balance with the quantity producers want to sell. If the price is too high, the quantity supplied by producers will exceed the quantity demanded. They will be unable to sell as much as they would like unless they reduce their price. Alternatively, if the price is too low, the quantity demanded by consumers will exceed the quantity supplied. Some consumers will be unable to get as much as they would like, unless they are willing to pay a higher price to bid some of the good away from other potential customers. Thus, there will be a tendency for the price in a market to move toward the price that brings the two into balance.

People have a tendency to think of consumers wanting lower prices and producers wanting higher prices. Although this is true, price changes frequently trend toward the middle of the two extremes. When a local store has an excess supply of a particular item, how does it get rid of it? By having a sale or somehow otherwise lowering its price (a "blue-light special"). Firms often lower their prices in order to get rid of excess supply.

On the other hand, excess demand is solved by consumers bidding up prices. Children's toys around Christmas provide a perfect example. When first introduced, items such as the Sony PlayStation 2, Furby, and the Tickle-Me-Elmo doll were immediate successes. The firms producing these products had not anticipated the overwhelming demand and every child wanted one for Christmas. Some stores raised their prices, but the demand was so strong that lines of parents were forming outside stores before they even opened. Often, only the first few in line were able to get the toys (a sure sign that the store had set the price below equilibrium). Out in the parking lots, in the classified ads, and on eBay, parents were offering to pay even higher prices for these items. If stores were not going to set the prices right, parents in these informal markets would! These examples show that rising prices are often the result of consumers bidding up prices when excess demand is present. A similar phenomenon can be seen in the market for Ty Beanie Babies (or concert tickets) as their immediate value on the resale market can be much higher than the original retail price if, at that price, the original quantity supplied is not adequate to meet the quantity demanded.

As these examples illustrate, whenever quantity supplied and quantity demanded are not in balance, there is a tendency for price to change in a manner that will correct the imbalance. It is possible to show this process graphically with the supply and demand curves we have developed in this chapter. **Exhibit 10** shows the supply and demand curves in the market for a basic calculator. At a high price—$12, for example—producers will plan to supply 600 calculators per day, whereas consumers will choose to purchase only 450. An excess supply of 150 calculators (shown by distance *ab* in the graph) will result. Unsold calculators will push the inventories of producers upward. To get rid of some of their calculators in inventory, some producers will cut their price to increase their sales. Other firms will have to lower their price, too, as a result, or sell even fewer calculators. This lower price will make supplying calculators less attractive to producers. Some of them will go out of business. Others will reduce their output or perhaps produce other products. How low will the price of calculators go? As the figure shows, when the price has declined to $10, the quantity supplied by producers and the quantity demanded by consumers will be in balance at 550 calculators per day. At this price ($1 0), the quantity demanded by consumers just equals the quantity supplied by producers, and the choices of the two groups are brought into harmony.

What will happen if the price per calculator is lower—$8, for example? In this case, the amount demanded by consumers (650 units) will exceed the amount supplied by producers (500 units). An excess demand of 150 units (shown by the distance *cd* in the graph)

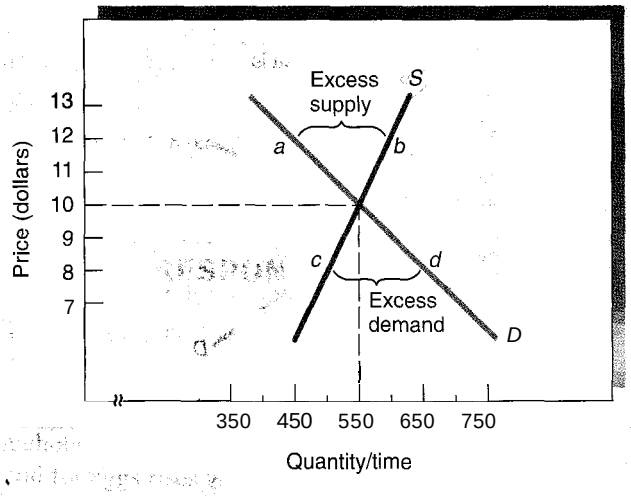


EXHIBIT 10
Supply and Demand

The table indicates the supply and demand conditions for calculators. These conditions are also illustrated by the graph. When the price exceeds $10, an excess supply is present, which places downward pressure on price. In contrast, when the price is less than $10, an excess demand results, which causes the price to rise. Thus, the market price will tend toward $10, at which point the quantity demanded will be equal to the quantity supplied.

Price of Calculators (Dollars)	Quantity Supplied (Per Day)	Quantity Demanded (Per Day)	Condition in the Market	Direction of Pressure on Price
$13	625	400	Excess supply	Downward
12	600	450	Excess supply	Downward
11	575	500	Excess supply	Downward
10	550	550	Balance	Equilibrium
9	525	600	Excess demand	Upward
8	500	650	Excess demand	Upward
7	475	700	Excess demand	Upward

will be the result. Some consumers who are unable to purchase the calculators at $8 per unit because of the inadequate supply would be willing to pay a higher price. Recognizing this fact, producers will raise their price. As the price increases to $10, producers will expand their output and consumers will cut down on their consumption. At the $10 price, equilibrium will be restored.

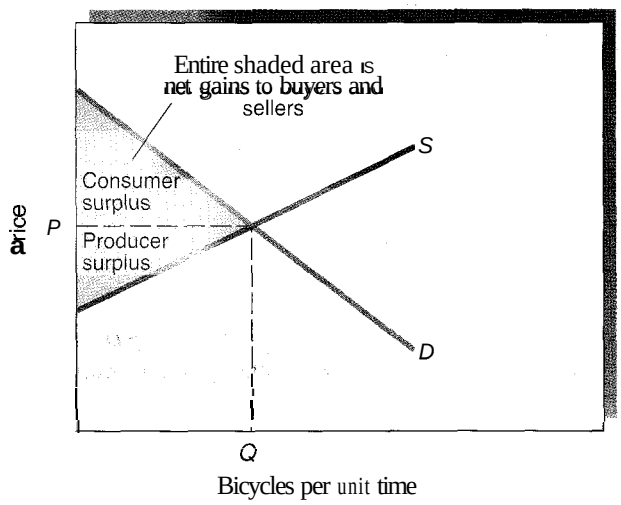
Efficiency and Market Equilibrium

When a market reaches equilibrium, all the gains from trade have been fully realized and **economic efficiency** is present. Economists often use economic efficiency as a standard to measure outcomes under alternative circumstances. The central idea of efficiency is a cost-versus-benefit comparison. Undertaking an economic action will be efficient only if it generates more benefit than cost. On the other hand, undertaking an action that generates more cost than benefit is inefficient. For a market to be efficient, all trades that generate more benefit than cost need to be undertaken. In addition, economic efficiency requires that no trades creating more cost than benefit be undertaken.

Economic efficiency
A situation in which all of the potential gains from trade have been realized. An action is efficient only if it creates more benefit than cost. With well-defined property rights and competition, market equilibrium is efficient.

A closer look at the way in which markets work can help us understand the concept of efficiency. The supply curve reflects producers' opportunity cost. Each point along the supply curve indicates the minimum price for which the units of a good could be produced without a loss to the seller. Assuming no other third parties are affected by the production of this good, then the height of the supply curve represents the opportunity cost to society of producing and selling the good. On the other side of the market, each point along the demand curve indicates how consumers value an extra unit of the good—that is, the maximum amount the consumer is willing to pay for the extra unit. Again assuming that no other third parties are affected, the height of the demand curve represents the benefit to society of producing and selling the good. Any time the consumer's valuation of a unit (the benefit) exceeds the producer's minimum supply price (the cost), producing and selling the unit is consistent with economic efficiency. The

EXHIBIT 11
Economic Efficiency



When markets are competitive and property rights are well defined, the equilibrium reached by a market satisfies economic efficiency. All units that create more benefit (the buyer's valuation shown by the height of the demand curve) than cost (opportunity cost of production shown by the height of the supply curve) are produced. This maximizes the total gains from trade, the combined area represented by consumer and producer surplus.

trade will result in mutual gain to both parties. When property rights are well defined, and only the buyers and sellers are affected by production and exchange, competitive market forces will automatically guide a market toward an equilibrium level of output that satisfies economic efficiency.

Exhibit 11 illustrates why this is true. Suppliers of bicycles will produce additional bicycles as long as the market price exceeds their opportunity cost of production (shown by the height of the supply curve). Similarly, consumers will continue to purchase additional bikes as long as their benefit (shown by the height of the demand curve) exceeds the market price. Eventually, market forces will result in an equilibrium output level of Q and a price of *P*. At this point all the bicycles providing benefits to consumers that exceed the costs to suppliers will be produced. Economic efficiency is met because all of the potential consumer and producer gains from exchange (shown by the shaded area) have occurred. As you can see, the point of market equilibrium is also the point where the combined area showing consumer and producer surplus is the greatest.

ECONOMICS AT THE MOVIES

TOUCHSTONE/WARNERS/THE KOBAL COLLECTION

Pretty Woman (1990)

In Pretty *Woman* Julia Roberts agrees to spend the week as Richard Gere's companion for $3,000. After Roberts and Gere agree on the price, she tells him that she would have been willing to do it for $2,000. His reply is that he would have been willing to pay $4,000. With this additional information, we know that the exchange netted Roberts $1,000 in producer surplus and Gere $1,000 in consumer surplus. This scene illustrates mutual gains from trade.

When less than Q bicycles are produced, some bicycles valued more by consumers than the opportunity cost of producing them are not being produced. This is inconsistent with economic efficiency. On the other hand, if output is expanded beyond Q, inefficiency will also result because some of the bicycles cost more to produce than consumers are willing to pay for them. Prices in competitive markets eventually guide producers and consumers to the level of output consistent with economic efficiency.

HOW MARKETS RESPOND TO CHANGES IN DEMAND AND SUPPLY

How will a market adjust to a change in demand? **Exhibit 12** shows the market adjustment to an increase in the demand for eggs around Easter time. Demand D_1 and supply S are typical throughout much of the year. During the two weeks before Easter, however, consumer demand for eggs rises because people purchase them to decorate too. This shifts egg demand from D_1 to D, during that time of year. As you can see, the increase in demand pushes the price upwards from P_1 to P_2 (typically by about 20 cents per dozen), and results in a larger equilibrium quantity traded (Q_2 rather than Q,—an increase of typically around 600 million eggs). There is a new equilibrium at point *b* around Easter (versus point a during the rest of the year).

Although consumers may not be happy about paying a higher price for eggs around Easter, the higher price serves two essential purposes. First, it encourages consumers to conserve on their usage of eggs. Some consumers may purchase only two dozen eggs to color, rather than three; other consumers may skip having an omelet for breakfast and have cereal instead. These steps on the consumer side of the market help make the eggs that are available around Easter go further. Second, the higher price is precisely what results in the additional 600 million eggs being supplied to the market to satisfy this increased consumer demand. Without the price increase, excess demand would be present—and many consumers would simply be unable to find eggs to purchase around Easter. If the price remained at P_1 (the equilibrium price throughout most of the year) consumers at Easter time would want to purchase more eggs than producers would be willing to supply. At the higher P, price, however, the quantity suppliers are willing to sell is again in balance with the quantity consumers wish to purchase.

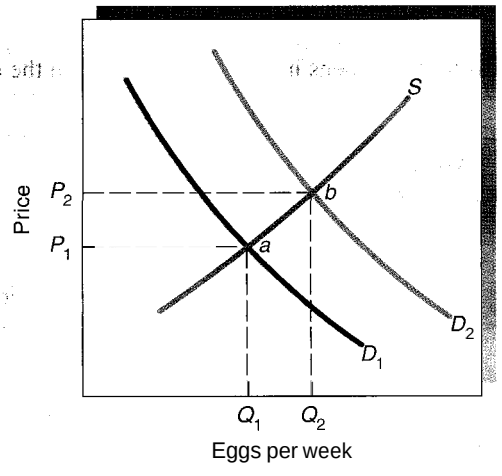
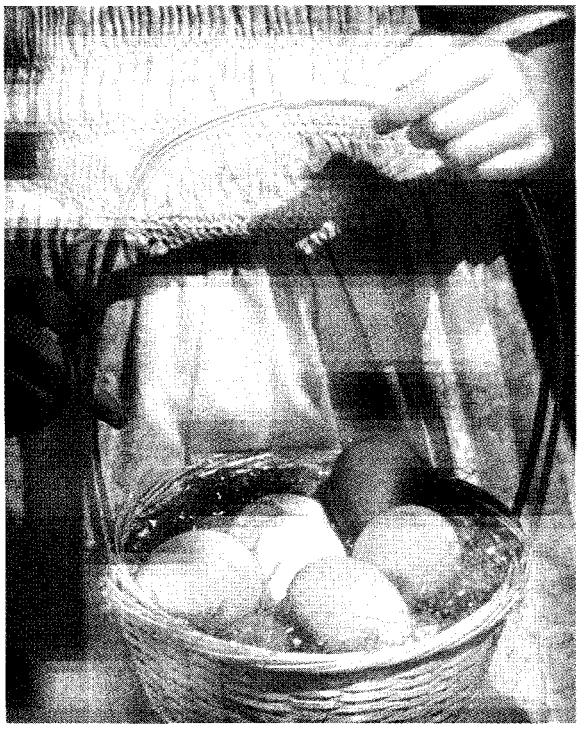


EXHIBIT 12
Market Adjustment to Increase in Demand

Here we illustrate how the market for eggs adjusts to an increase in demand such as generally occurs around Easter time. Initially (before the Easter season), the market for eggs reflects demand D_1 and supply S. The increase in demand (shift from D_1 to D_2) pushes price up and leads to a new equilibrium at a higher price (P_2 rather than P_1) and larger quantity traded (Q_2).



© GETTY IMAGES

The tradition of coloring and hunting for eggs causes an increase in demand for eggs around Easter. As Exhibit 12 illustrates, this leads to higher egg prices and costly actions by producers to supply a larger quantity during this period.

Why were suppliers unwilling to supply the additional 600 million eggs at the original price of P_1? Because at the original equilibrium price of P_1, suppliers were already producing and selling all the eggs that cost less to produce than that price. The additional eggs desired by consumers around Easter all cost more to produce than the old market price of P_1. The higher price of P_2 is what allows suppliers to cover their higher production costs associated with these extra eggs. Around Easter, farmers take costly steps to avoid having the hens molt because hens lay fewer eggs when they are molting. They do this by changing the quantity and types of feed and by increasing the lighting in the birds' sheds—both of which mean higher production costs. Farmers also try to build up larger than normal inventories of eggs before Easter. Eggs are typically about two days old when consumers buy them at the store, but can be up to seven days old around Easter time. Building up and maintaining this additional inventory is costly, too.

In a market economy, when the demand for a good increases, its price will rise, which will (1) motivate consumers to search for substitutes and cut back on additional purchases of the good and (2) motivate producers to supply more of the good. These two forces will eventually bring the quantity demanded and quantity supplied back into balance.

It's important to note that this response on the supply side of the egg market is not a shift in the supply curve. The supply curve remains unchanged. Rather, there is a movement along the original supply curve—a change in *quantity* supplied. The only reason suppliers are willing to alter their behavior (produce more eggs) is because the increased demand has pushed up the price of eggs. Notice that it is the change in demand (a shift of the demand curve) that leads to the change in quantity supplied (a movement along the supply curve). Producers are simply responding to the price movement caused by the change in demand. A movement along one curve (a change in quantity supplied *or* a change in quantity demanded) happens in response to a shift in the other curve (a change in demand or a change in supply).

When the demand for a product declines, the adjustment process sends buyers and sellers just the opposite signals. Take a piece of paper and see if you can diagram a decrease in demand and how it will affect price and quantity in a market. If you've done it correctly, a decline in demand (a shift to the left in the demand curve) will lead to a lower price and a lower quantity traded. What's going on in the diagram is that the lower price (caused by lower consumer demand) is reducing the incentive of producers to supply the good. When consumers no longer want as much of a good, falling market prices signal producers to cut back production. The reduced output allows these resources to be freed up to go into the production of other goods consumers want more.

How will markets respond to changes in supply? **Exhibit 13** shows the market's adjustment to a decrease in the supply of romaine lettuce. Assume that severe rains and flooding destroy a large portion of the romaine lettuce crop. This reduction in supply (shift from S_1 to S_2) will cause the price of romaine to increase sharply (P_1 to P_2). Because of the higher price, consumers will cut back on their consumption of romaine lettuce (the movement along the demand curve from a to *b*). Some will switch to substitutes—in this

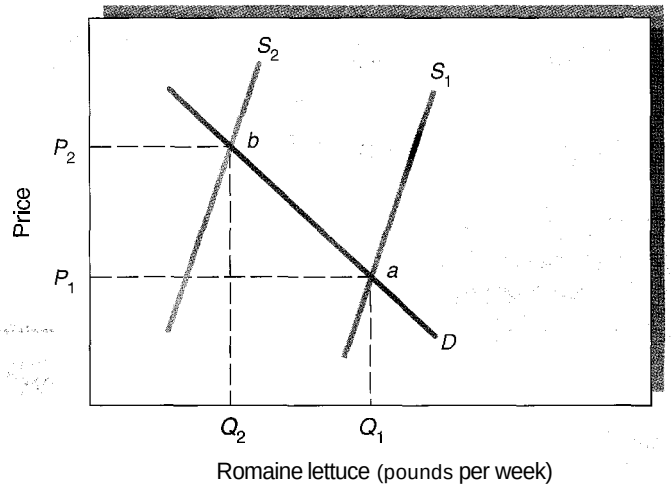


EXHIBIT 13
Market Adjustment to a Decrease in Supply

Here, using romaine lettuce as an example, we illustrate how a market adjusts to a decrease in supply. Assume adverse weather conditions substantially reduce the supply (shift from S_1 to S_2) of romaine. The reduction in supply leads to an increase in the equilibrium price (from P_1 to p_2) and a reduction in the equilibrium quantity traded (from Q to Q_2).

case, probably other varieties of lettuce and leafy vegetables. The higher price also encourages the remaining romaine suppliers to take additional steps—like more careful harvesting techniques or using more fertilizer—that allow them to produce more romaine lettuce than otherwise would be the case. The higher prices will rebalance the quantity demanded and quantity supplied.

As the lettuce example illustrates, a decrease in supply will lead to higher prices and a lower equilibrium quantity. How do you think the market price and quantity would adjust to an increase in supply, as might be caused by a breakthrough in the technology used to harvest the lettuce? Again, try to draw the appropriate supply and demand curves to illustrate this case. If you do it correctly, the graph you draw will show an increase in supply (a shift to the right in the supply curve) leading to a lower market price and a larger equilibrium quantity.

The accompanying **Thumbnail Sketch** summarizes the effect of changes—both increases and decreases—in demand and supply on the equilibrium price and quantity. The cases listed in the thumbnail sketch, however, are for when only a single curve shifts. But sometimes market conditions simultaneously shift both demand and supply. For example, consumer income might increase at the same time that a technological advance in production occurs. These two changes will cause both demand and supply to increase at the same time—both curves will shift to the right. The new equilibrium will definitely be at a larger quantity, but the direction of the change in price is indeterminate. The price may either increase or decrease, depending on whether the increase in demand or increase in supply is larger—which curve shifted the most, in other words.

THUMBNAIL SKETCH

How Changes in Demand and Supply Affect Market Price and Quantity

Changes in Demand

1. An increase in demand—shown by a rightward shift of the demand curve—will cause an increase in both the equilibrium price and quantity.
2. A decrease in demand—shown by a leftward shift of the demand curve—will cause a decrease in both the equilibrium price and quantity.

Changes in Supply

1. An increase in supply—shown by a rightward shift of the supply curve—will cause a decrease in the equilibrium price and an increase in the equilibrium quantity.
2. A decrease in supply—shown by leftward shift of the supply curve—will cause an increase in the equilibrium price and a decrease in the equilibrium quantity.

What will happen if supply increases but demand falls at the same time? Price will definitely fall, but the new equilibrium quantity may either increase or decrease. Draw the supply and demand curves for this case and make sure that you understand why.

INVISIBLE HAND PRINCIPLE

Invisible Hand Principle

Market prices coordinate the actions of self-interested individuals and direct them toward activities that promote the general welfare.

More than 225 years ago, Adam Smith, the father of economics, stressed that personal self-interest *when directed by market prices* is a powerful force promoting economic progress. In a famous passage in his book *The Wealth of Nations,* Smith put it this way:

> Every individual is continually exerting himself to find out the most advantageous employment for whatever [income] he can command. It is his own advantage, indeed, and not that of the society which he has in view. But the study of his own advantage naturally, or rather necessarily, leads him to prefer that employment which is most advantageous to society. . . . He intends only his own gain, and he is in this, as in many other cases, led by an invisible hand to promote an end which was not part of his intention. By pursuing his own interest he frequently promotes that of the society more effectually than when he really intends to promote it.[5]

Smith's fundamental insight was that market forces would tend to align the actions of self-interested individuals with the best interests of society. The tendency of market forces to channel the actions of self-interested individuals into activities that promote the general betterment of society is now known as the **invisible hand principle.** Let's take a closer look at this important principle.

Invisible hand principle
The tendency of market prices to direct individuals pursuing their own interests to engage in activities promoting the economic well-being of society.

Prices and Market Order

The invisible hand principle can be difficult to grasp because there is a natural tendency to associate order with central direction and control. Surely some central authority must be in charge. But this is not the case. ***The pricing system, reflecting the choices of literally millions of consumers, producers, and resource owners, provides the direction.*** The market process works so automatically that most of us give little thought to it. We simply take it for granted.

Perhaps one example from your everyday life will help you better understand the invisible hand principle. Visualize a busy retail store with ten checkout lanes. No individual assigns shoppers to checkout lanes. Shoppers are left to choose for themselves. Nonetheless, they do not all try to get in the same lane. Why? Individuals are always alert for adjustment opportunities that offer personal gain. When one lane gets long or is held up by a price check, some shoppers will shift to other lanes and thereby smooth out the flow among the lanes. Even though central planning is absent, this process of mutual adjustment by self-interested individuals results in order and social cooperation. In fact, the degree of social cooperation is generally well beyond what could be achieved if central coordination were attempted—if, for example, stores hired someone to assign shoppers to checkout lanes in the interest of getting everyone out as quickly as possible. Shoppers *acting in their own interests* promote the most orderly and quickest flow for everyone. A similar phenomenon occurs on busy interstate highways as drivers switch between lanes for personal gain, with the end result being the quickest flow of traffic for everyone.

[5] Adam Smith, *An Inquiry into the Nature and Causer of the Wealth of Nations* (New York Modem Library, 1937), 423

Market participation is a lot like checking out at a retail store or driving on the freeway. Like the number of people in a lane, profits and losses provide market participants with information about the advantages and disadvantages of different economic activities. Losses indicate that an economic activity is congested, and, as a result, producers are unable to cover their costs. In such a case, successful market participants will shift their resources away from such activities toward other, more valuable uses. Conversely, profits are indicative of an open lane, the opportunity to experience gain if one shifts into an activity where the price is high relative to the per-unit cost. As producers and resource suppliers shift away from activities characterized by congestion and into those characterized by the opportunity for profit, they smooth out economic activity and enhance its flow. Remarkably, even though individuals are motivated by self-interest, market prices direct their actions toward activities that promote both order and economic progress. This is precisely the message of Smith's "invisible hand."

Is the concept of the invisible hand really valid? Next time you sit down to have a nice dinner, think about all the people who help make it possible. It is unlikely that any of them, from the farmer to the truck driver to the grocer, was motivated by a concern that you have an enjoyable meal. Market prices, however, bring their interest into harmony with yours. Farmers who raise the best beef or turkeys receive higher prices; truck drivers and grocers earn more money if their products are delivered fresh and in good condition; and so on. An amazing degree of cooperation and order is created by market exchanges—all without the central direction of any government official.

How do markets bring the interests of individuals into harmony with economic progress? Consider the following three vitally important functions performed by market prices.

1. Prices communicate information to decision makers. Suppose a drought in Brazil severely reduces the supply of coffee. Coffee prices will rise. Even if consumers do not know about the drought, the higher prices will provide them with all the information they need to know—it's time to cut back on coffee consumption. ***Market prices register information derived from the choices of millions of consumers, producers, and resource suppliers, and provide them with everything they need to know to make wise decisions.***

Market prices provide producers with up-to-date information about which goods consumers most intensely desire, and with important information about the abundance of the resources used in the production process. The cost of production, driven by the opportunity cost of resources, tells the business decision maker the relative importance others place on the alternative uses of those resources. A boom in the housing market might cause lumber prices to rise. In turn, furniture makers seeing these higher lumber prices will utilize substitute raw materials such as metal and plastic in their production processes. Because of market prices, furniture makers will conserve on their use of lumber, just as if they had known that lumber was now more urgently needed for constructing new housing.

2. Prices coordinate the actions of market participants. Market prices also coordinate the choices of buyers and sellers, bringing their decisions into line with each other. Excess supply will lead to falling prices, which discourage production and encourage consumption until the excess supply is eliminated. Alternatively, excess demand will lead to price increases, which encourage consumers to economize on their uses of the good and suppliers to produce more of it, eliminating the excess demand. Changing market prices induce responses on both sides of the market in the proper direction to help correct these situations.

The combination of product and resource prices will determine profit (and loss) rates for alternative projects and thereby direct entrepreneurs to undertake the production projects that consumers value most intensely (relative to their cost). If consumers really want more of a good—for example, luxury apartments—the intensity of their demand will lead to a market price that exceeds the opportunity cost of constructing the apartments. The profitable opportunity thus created will soon be discovered by entrepreneurs who will undertake the construction, expanding the availability of the apartments. In contrast, if

consumers want less of a good, such as large cars, the sales revenue from their production will be less than the opportunity cost of supplying them, penalizing those who undertake such unprofitable production.

3. Prices motivate economic players. Market prices establish a reward-penalty (profit-loss) structure that encourages people to work, cooperate with others, use efficient production methods, supply goods that are intensely desired by others, and invest for the future. Self-interested entrepreneurs will seek to produce only the goods consumers value enough to pay a price sufficient to cover production cost. Self-interest will also encourage producers to use efficient production methods and adopt cost-saving technologies because lower costs will mean greater profits. Firms that fail to do so will be unable to compete successfully in the marketplace.

At the beginning of this chapter, we asked you to reflect on why the grocery stores in your local community generally have on hand about the right amount of milk, bread, vegetables, and other goods. Likewise, how is it that refrigerators, automobiles, and CD players, produced at different places around the world, make their way to stores near you in approximately the same numbers that they are demanded by consumers? The invisible hand principle provides the answer, and it works without political direction. No government agency needs to tell decision makers to keep costs low or produce those goods most intensely desired by consumers. Similarly, no one has to tell individuals that they should develop skills that are highly valued by others. Once again the profit motive—in this case higher earnings—will do the job. Many of the things we take for granted in our ordinary lives reflect the invisible hand at work.

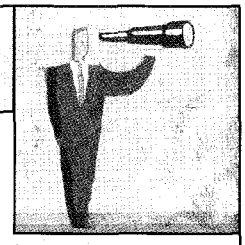
Qualifications: Competition and Property Rights

As we noted earlier in this chapter, our focus so far has been on markets where rival firms can freely enter and exit, and private-property rights are clearly defined and enforced. *The efficiency of market organization is, in fact, dependent on these two things: (1) competitive markets and (2) well-defined and enforced private-property rights.*

Competition, the great regulator, can protect both buyer and seller. It protects consumers from sellers who would charge a price substantially above the cost of production or withhold a vital resource for an exorbitant amount of money. Similarly, it protects employees (sellers of their labor) from the power of any single employer (the buyers of labor). Competition equalizes the bargaining power between buyers and sellers.

When property rights are well defined, secure, and tradable, suppliers of goods and services have to pay resource owners for their use. They will not be permitted to seize and use scarce resources without compensating the owners. Neither will they be permitted to use violence (for example, to attack or invade the property of another) to get what they want. The efficiency of markets hinges on the presence of property rights—after all, people can't easily exchange or compete for things they don't have or can't get property rights to. Without well-defined property rights, markets simply cannot function effectively.

LOOKING AHEAD



Although we incorporated numerous examples designed to enhance your understanding of the supply-and-demand model throughout this chapter, we have only touched the surface. In various modified forms, this model is the central tool of economics. The following chapter will explore several specific applications and extensions of this important model.

KEY POINTS

- The law of demand states that there is an inverse (or negative) relationship between the price of a good or service and the quantity of it that consumers are willing to purchase. The height of the demand curve at any quantity shows the maximum price that consumers are willing to pay for that unit.
- The degree of responsiveness of consumer purchases to a change in price is shown by the steepness of the demand curve. The more responsive buyers are to a change in price, the flatter, or more elastic, the demand curve will be. Conversely, the less responsive buyers are to a change in price, the steeper, or more inelastic, the demand curve will be.
- A movement along a demand curve is called a change in quantity demanded. A shift of the entire curve is called a change in demand. A change in *quantity demanded* is caused by a change in the price of the good (generally in response to a shift of the supply curve). A change in *demand* can be caused by several things, including a change in consumer income or a change in the price of a closely related good.
- The opportunity cost of producing a good is equal to the cost of bidding away the resources needed for its production from alternative uses. Profit indicates that the producer has increased the value of the resources used, whereas a loss indicates that the producer has reduced the value of the resources used.
- The law of supply states that there is a direct (or positive) relationship between the price of a good or service and the quantity of it that producers are willing to supply. The height of the supply curve at any quantity shows the minimum price necessary to induce suppliers to produce that unit—that is, the opportunity cost of producing it.
- A movement along a supply curve is called a change in quantity supplied. A change in *quantity supplied* is caused by a change in the price of the good (generally in response to a shift of the demand curve). A shift of the entire supply curve is called a change in supply. A change in *supply* can be caused by several factors, such as a change in resource prices or an improvement in technology.
- The responsiveness of supply to a change in price is shown by the steepness of the supply curve. The more willing producers are to alter the quantity supplied in response to a change in price, the flatter, or more elastic, the supply curve. Conversely, the less willing producers are to alter the quantity supplied in response to a change in price, the steeper, or less elastic, the supply curve.
- Prices bring the conflicting forces of supply and demand into balance. There is an automatic tendency for market prices to move toward the equilibrium price, at which the quantity demanded equals the quantity supplied.
- Consumer surplus represents the net gain to buyers from market trades. Producer surplus represents the net gain to producers and resource suppliers from market trades. In equilibrium, competitive markets maximize these gains, a condition known as economic efficiency.
- Changes in the prices of goods are caused by changes in supply and demand. An increase in demand will cause the price and quantity supplied to rise. Conversely, a decrease in demand will cause the price and quantity supplied to fall. An increase in supply, on the other hand, will cause the price to fall and quantity demanded to rise. Conversely, a decrease in supply will cause the price to rise and quantity demanded to fall.
- Market prices communicate information, coordinate the actions of buyers and sellers, and motivate decision makers to act. As the invisible hand principle indicates, market prices are generally able to bring the self-interest of individuals into harmony with the general welfare of society. The efficiency of the system is dependent on two things, however: (1) competitive market conditions and (2) well-defined and secure property rights.

CRITICAL ANALYSIS QUESTIONS

***1.** Which of the following do you think would lead to an increase in the current demand for beef?
- **a.** higher pork prices
- **b.** higher consumer income
- **c.** higher prices of feed grains used to feed cattle
- **d.** widespread outbreak of mad cow or foot-and-mouth disease
- **e.** an increase in the price of beef

2. What is being held constant when a demand curve for a specific product (shoes or apples, for example) is constructed? Explain why the demand curve for a product slopes downward to the right.

3. What is the law of supply? How many of the following "goods" do you think conform to the general law of supply? Explain your answer in each case.
   a. gasoline
   b. cheating on exams
   c. political favors from legislators
   d. the services of heart specialists
   e. children
   f. legal divorces

*4. Are prices an accurate reflection of a good's total value? Are prices an accurate reflection of a good's marginal value? What's the difference? Can you think of a good that has high total value but low marginal value? Use this concept to explain why professional wrestlers earn more than nurses, despite the fact that nurses probably create more total value to society.

5. What is being held constant when the supply curve is constructed for a specific good like pizza or automobiles? Explain why the supply curve for a good slopes upward to the right.

6. Define consumer and producer surplus. What is meant by economic efficiency, and how does it relate to consumer and producer surplus?

7. Recent tax reforms make college tuition partially tax deductible for certain families. This should motivate more people to attend college. How will this higher demand for a college education affect tuition prices? How will it affect the cost of college for families who don't qualify for the tax deduction?

*8. "The future of our industrial strength cannot be left to chance. Somebody has to develop notions about which industries are winners and which are losers." Is this statement by a newspaper columnist true? Who is the "somebody"?

9. What does the cost of a good or service reflect? Will producers continue to supply a good or service if consumers are unwilling to pay a price sufficient to cover the cost?

*10. "Production should be for people and not for profit." Answer the following questions concerning this statement:
   a. If production is profitable, are people helped or harmed? Explain.
   b. Are people helped more if production results in a loss than if it leads to profit? Is there a conflict between production for people and production for profit?

11. What must an entrepreneur do to earn a profit? How do the actions of firms earning profits influence the value of resources? What happens to the value of resources when losses are present? If a firm making losses goes out of business, is this bad? Why or why not?

*12. What's wrong with this way of thinking? "Economists claim that when the price of something goes up, producers increase the quantity supplied to the market. But last year, the price of oranges was really high and the supply of them was really low. Economists are wrong!"

13. What is the invisible hand principle? Does it indicate that self-interested behavior within markets will result in actions that are beneficial to others? What conditions are necessary for the invisible hand to work well? Why are these conditions important?

*14. What's wrong with this way of thinking? "Economists argue that lower prices will result in fewer units being supplied. However, there are exceptions to this rule. For example, in 1972, a very simple ten-digit electronic calculator sold for $120. By 2000, the price of the same type of calculator had declined to less than $5. Yet business firms produced and sold many more calculators in 2000 than they did in 1972. Lower prices did not result in less production or in a decline in the number of calculators supplied."

15. What is the difference between substitutes and complements? Indicate two goods that are substitutes for each other. Indicate two goods that are complements.

16. How is the market price of a good determined? When the market for a product is in equilibrium, how will consumers value an additional unit compared to the opportunity cost of producing that unit? Why is this important?

*17. Do business firms operating in competitive markets have a strong incentive to serve the interest of consumers? Are they motivated by a strong desire to help consumers? Are "good intentions" necessary if individuals are going to engage in actions that are helpful to others? Discuss.

*Asterisk denotes questions for which answers are given in Appendix B.

Supply and Demand: Applications and Extensions

CHAPTER 4

Chapter Focus

- How broadly can the supply and demand framework be used?
- What happens when prices are set by law above or below the market equilibrium level?
- How do rent controls affect the maintenance and quality of rental housing? How do minimum-wage rates influence the job opportunities of low-skilled workers?
- What are "black markets"? How does the lack of a well-structured legal environment affect their operation?
- How does a tax or subsidy affect a market? What determines the distribution of the tax burden (or subsidy benefit) between buyers and sellers?
- What is the Laffer curve? What does it indicate about the relationship between tax rates and tax revenues?

The division of labour, from which so many advantages are derived, is not originally the effect of any human wisdom, which foresees and intends that general opulence to which it gives occasion. It is the necessary, though very slow and gradual consequence of a certain propensity in human nature . . . ; the propensity to truck, barter, and exchange one thing for another.

—Adam Smith[1]

Nations stumble upon establishments, which are indeed the result of human action, but not the execution of any human design.

—Adam Ferguson[2]

[1]Adam Smith, *An Inquiry into the Nature and Causes of the Wealth of Nations* (New York: Modern Library, 1937), 13.
[2]Adam Ferguson, *An Essay on the History of Civil Society* (Edinburgh: A. Millar and T. Caddel, London, 1767), 187.

Markets are everywhere. They exist in many different forms and degrees of sophistication. In elementary schools, children trade Pokémon cards; in households, individuals trade chores ("I'll clean the bathroom, if you'll clean the kitchen"); and in the stock market, individuals who have never met exchange shares of corporate stock and other financial assets worth billions of dollars each business day. Even making an activity illegal does not eliminate the market for it. Instead, the market is merely pushed underground. The exchange of illegal drugs or tickets to a big game at illegal prices illustrates this point.

Trading with other individuals is a natural part of human behavior that exists regardless of legal and societal conditions. As Adam Smith put it more than 225 years ago, human beings have a natural propensity "to truck, barter, and exchange one thing for another" (see the quotation at the chapter opening). We all want to improve our standard of living, and trade with others helps us achieve this goal—by allowing us to get the goods and services we really want and giving us the opportunity to earn the income necessary to buy them. Further, as Adam Ferguson points out, markets are a result of human action, not human design.[3] They reflect the desire of people to trade things with one another.

Market prices coordinate the actions of buyers and sellers, but sometimes the "price" of a good or service in a particular market is called something different. For example, in the labor market, the price is often called the "wage rate." In the loanable funds market, the price is generally referred to as the "interest rate." However, as Juliet observes in Shakespeare's *Romeo and Juliet,* "What's in a name? That which we call a rose by any other name would smell as sweet." The same is true for prices. When the price of something is referred to by another term, such as the wage or interest rate, it will still play the same role. Therefore, when these special terms are used, we put them along the vertical axes of supply and demand diagrams, just as we do "price"—becausethat's what they are.

In the previous chapter, we saw how the forces of supply and demand determine market prices and coordinate the actions of buyers and sellers in the absence of government intervention. In this chapter, we turn our attention to using the supply and demand model to understand more fully what happens when governments intervene in markets by implementing price controls, taxes, and subsidies.

THE LINK BETWEEN RESOURCE AND PRODUCT MARKETS

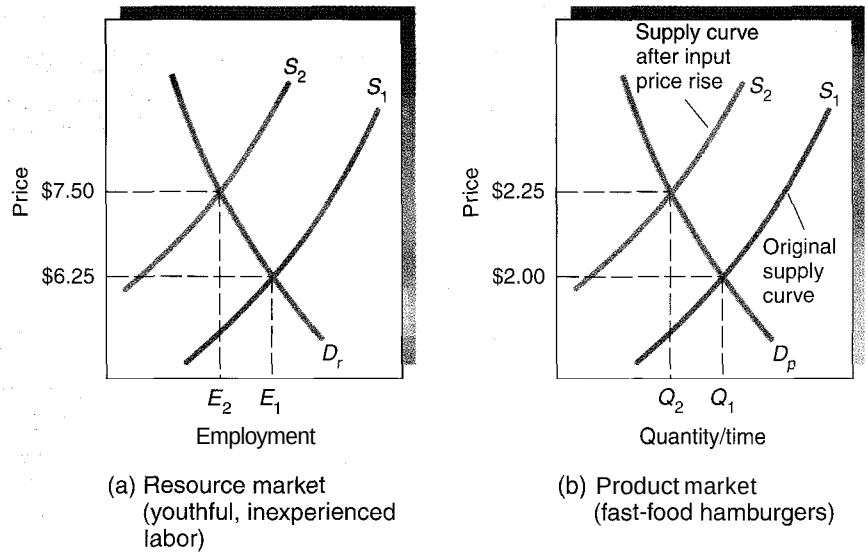
The interrelationship among markets is vitally important. A change in one market will also lead to changes in other markets. Understanding these links is important. This section addresses the important link between the labor and product markets.

The production process generally involves (a) purchasing resources—like raw materials, labor services, tools, and machines; (b) transforming these resources into products (goods and services); and (c) selling the goods and services in a product market. Production is generally undertaken by business firms. Typically, business firms will demand resources, while households will supply them. Firms demand resources *because* they contribute to the production of goods and services. In turn, households supply them in order to earn income.

Resource market
The market for inputs used to produce goods and services.

Just as in product markets, the demand curve in a **resource market** is typically downward-sloping and the supply curve upward-sloping. An inverse relationship will exist between the amount of a resource demanded and its price because businesses will substitute away from a resource as its price rises. In contrast, there will be a direct relationship between the amount of a resource supplied and its price because a higher price will make that resource more attractive to provide. As in product markets, prices will coordinate the choices of buyers and sellers in resource markets, bringing the quantity demanded into balance with the quantity supplied.

[3] This theme was a focus of much of the work of Nobel Prize–winning economist Fnedrich Hayek



(a) Resource market (youthful, inexperienced labor)

(b) Product market (fast-food hamburgers)

EXHIBIT 1
Resource Prices, Opportunity Cost, and Product Markets

When the supply of young, inexperienced workers falls, it pushes the wage rates of fast-food employees upward (a). In the product market (b), the higher wage rates will increase the opportunity costs of restaurants, reducing supply (shifting from S_1 to S_2) and increasing hamburger prices.

The labor market is a large component of the broader resource market. Actually, there is not just one market for labor, but rather there are many labor markets, one for each different skill-experience-occupational category. Let's look at the labor market for low-skilled, inexperienced workers. **Exhibit 1** shows how resource and product markets are linked. The supply of young, inexperienced workers has declined in recent years in many areas of the United States. This lower supply has pushed the wages of young workers upward (for example, from $6.25 to $7.50 in Exhibit 1a). The higher price of this resource increases the opportunity cost of goods and services that young workers help produce. In turn, the higher cost reduces the supply (shifting S_1 to S,) of products like hamburgers at McDonald's and other fast-food restaurants, pushing their price upward (Exhibit 1b). When the price of a resource increases, it will lead to higher production costs, lower supply, and higher prices for the goods and services produced with the resource.

Of course, lower resource prices have the opposite effect. Lower resource prices reduce costs and expand the supply of consumer goods made with the lower-priced resources (shifting the supply curve to the right). The increase in supply will lead to a lower price in the product market. ***Thus, when the price of a resource—such as labor—changes, the prices of goods and services produced with that resource will change in the same direction.***

Changes in product markets will also influence resource markets. ***There is a close relationship between the demand for products and the demand for the resources required for their production.*** An increase in demand for a consumer good—automobiles, for example—will lead to higher auto prices, which will increase the profitability of producing automobiles and give automakers an incentive to expand output. But the expansion in automobile output will require additional resources, causing an increase in the demand for. and prices of, the resources required for their production (steel, rubber, plastics, and the labor services of autoworkers, for example). The higher prices of these resources will cause other industries to conserve on their use, freeing them up for more automobile production.

Of course, the process will work in reverse if demand for a product falls. **A** decrease in demand will not only reduce the price of the product but will also reduce the demand for and prices of the resources used to produce it. ***Thus, when the demand for a product changes, the demand for (and prices of) the resources used to produce it will change in the same direction.***

THE ECONOMICS OF PRICE CONTROLS

Price controls
Government-mandated prices that are generally imposed in the form of maximum or minimum legal prices.

Buyers often complain that prices are too high, while sellers complain that they are too low. Unhappy with the prices established by market forces, various groups might try to persuade the government to intervene and impose **price controls**. Price controls force buyers or sellers to alter the prices of certain products. Price controls may be either price ceilings, which set a maximum legal price for a product, or price floors, which impose a minimum legal price. Imposing price controls might seem like a simple, easy way for the government to help buyers at the expense of sellers (or vice versa). The problem is that doing so frequently creates secondary effects that make *both* sides worse off.

Despite good intentions, price controls can, in fact, harm the very people they were intended to help because they undermine the exchange process and reduce the gains from trade. The regulation of automated teller machine (ATM) surcharge fees is one example. Many states, after being lobbied by consumer groups, enacted regulations forbidding or severely restricting the ability of ATM owners to charge fees for using their machines. The unintended consequence of these regulations is that there are now fewer ATMs available to consumers in these states because there's less incentive to own and operate them. Consumers in these states benefit by paying lower ATM fees, but they also bear the cost of reduced ATM access.

The Impact of Price Ceilings

Price ceiling
A legally established maximum price sellers can charge for a good or resource.

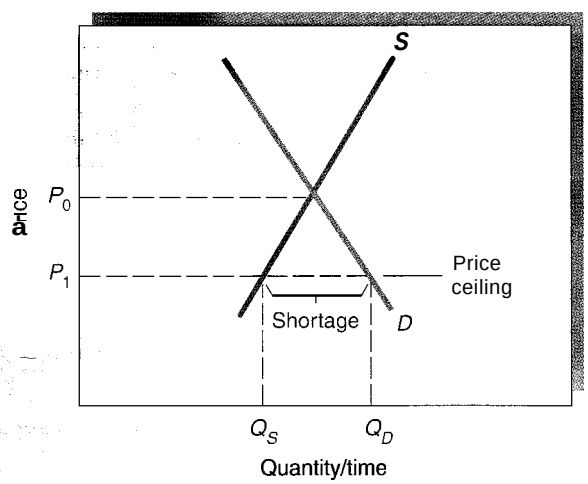
Shortage
A condition in which the amount of a good offered for sale by producers is less than the amount demanded by buyers at the existing price An increase in price would eliminate the shortage.

Exhibit 2 shows the impact of imposing a **price ceiling** $(P,)$ for a product below its equilibrium level $(P,)$. At the lower price, the quantity supplied by producers decreases along the supply curve to Q_S, while the quantity demanded by consumers increases along the demand curve to Q. A **shortage** $(Q_D - Q,)$ of the good will result because the quantity demanded by consumers exceeds the quantity supplied by producers at the new controlled price. After the price ceiling is imposed, the quantity of the good exchanged declines from the equilibrium quantity to Q_S, and the gains from trade (consumer and producer surplus) fall as well.

Normally, a higher price would ration the good to the buyers most willing to pay for it. Because the price ceiling keeps this from happening, though, other means must be used to allocate the smaller quantity Q, among consumers wanting to purchase Q. Predictably, nonprice factors will become more important in the rationing process. Sellers will be forced to discriminate on some basis other than willingness to pay as they ration their sales to eager buyers. They will be more inclined to sell their products to their friends, to buyers who do them favors, and even buyers willing to make illegal "under-the-table" payments. (The accompanying Applications in Economics box, "The Imposition of Price Ceilings During Hurricane Hugo," highlights this point.) Time might also be used as the rationing

EXHIBIT 2
The Impact of a Price Ceiling

When a price ceiling like P_1 pushes the price of a product (rental housing, for example) below the market equilibrium, a shortage will develop. Because prices are not allowed to direct the market to equilibrium, nonprice elements will become more important in rationing the good.



APPLICATIONS IN ECONOMICS

The Imposition of Price Ceilings During Hurricane Hugo

In the fall of 1989, Hurricane Hugo struck the coast of South Carolina, causing massive property damage and widespread power outages lasting for weeks. The lack of electric power meant that gasoline pumps, refrigerators, cash registers, ATMs, and many other types of electrical equipment did not work. In the hardest-hit coastal areas such as Charleston, the demand for items such as lumber, gasoline, ice, batteries, chain saws, and electric generators increased dramatically. A bag of ice that sold for $1 before the hurricane went up in price to as much as $10; the price of plywood rose to about $200 per sheet; chain saws soared to the $600 range; and gasoline sold for as much as $10.95 per gallon. At these higher prices, individuals from other states were renting trucks, buying supplies in their home state, driving them to Charleston, and making enough money to pay for the rental truck and the purchase of the goods and to compensate them for taking time off from their regular jobs.

In response to consumer complaints of "price gouging," the mayor of Charleston signed emergency legislation making it a crime, punishable by up to *30* days in jail and a $200 fine, to sell goods at prices higher than their prehurricane levels in the city. The price ceilings kept prices down, but also stopped the flow of goods into the area almost immediately. Shippers of items like ice would stop and sell their goods outside the harder-hit Charleston area to avoid the price controls. Shipments that actually made it into Charleston were often greeted by long lines of consumers, many of whom would end up without goods after waiting in line for up to five hours. Many of the people at the front of the line who were able to buy the goods before supplies ran out would then drive those goods back out of the city to sell them at the higher, noncontrolled prices in areas outside of Charleston to obtain money to pay for repairs to their home. Shortages became so bad that military guards were needed to protect shipments of the goods and maintain order when a shipment did arrive.

The price controls resulted in serious misallocations of resources. Electric generators provide one of the best examples. Grocery stores couldn't open because there was no electricity. Inside stores, food was spoiling—thousands of dollars worth, in many cases. Although gas stations had gasoline in their underground storage tanks, it couldn't be pumped out without electricity. ATMs and banks couldn't operate without electricity, so people couldn't get their hands on their money—which was critical because almost all transactions in post hurricane Charleston were made with cash.

Hardware stores that sold gasoline-powered electric generators before the hurricane typically had only a few in stock, but suddenly hundreds of businesses and residents wanted to buy them. In the absence of price controls, the price of these generators would have risen to thousands of dollars to allocate the limited supply. Individual homeowners would have been outbid by businesses, which could have put the generators to use operating stores and gas stations and ATMs. It would have been these uses that could have generated enough revenue to cover the high price of the generators. People with generators at home would have found it attractive to sell them to businesses for the high sums of money involved.

However, the price ceilings prevented the generators from being allocated to those most willing to pay for them. Instead, people kept their generators at home, and it was commonplace for hardware store owners with a few generators on hand to take one home for their family and then sell the others to their close friends, neighbors, and relatives. In the absence of price rationing, nonprice factors played a larger role in the allocation process. The electric generators so critically needed for grocery stores, gasoline stations, and banks to open were instead being used by households for tasks such as running television sets, lighting, electric razors, hair dryers, and so on. As a result, hundreds of thousands of consumers couldn't get goods they urgently wanted. Moreover, the flow of new generators into the city effectively stopped, and many generators were actually taken out of the city to be sold in the less-damaged, outlying areas where price controls were not in effect. If price controls hadn't been imposed, the price of generators would quickly have been bid up to the point where they would have been (1) purchased by those with the most urgent uses for them, and (2) imported into the city fairly rapidly because of the high prices they commanded.

This dramatic example shows how important prices are. Market prices will both allocate goods to those who will put them to the highest valued uses and motivate people to supply more of the items in short supply. The secondary effects, or unintended consequences, of the price controls imposed in the Charleston area actually magnified people's suffering and retarded the recovery from the hurricane.[1] In recent years, communities hit by hurricanes and other disasters have been reluctant to impose price controls, perhaps because of what happened in Charleston.

[1]See David N. Laband, "In Hugo's Path, a Man-Made Disaster," *Wall Street Journal*, September 27, 1989, A22; and Tim Smith, "Economists Spurn Price Restrictions," *Greenville News*, September 28, 1989, C1.

device, with those willing to wait in line the longest being the ones able to purchase the good. In addition, the below-equilibrium price reduces the incentive of sellers to expand the future supply of the good. At the lower price, suppliers will direct resources away from production of the good and into other, more profitable areas. As a result, the product shortage will worsen through time.

What other secondary effects can be expected? ***In the real world, there are two ways that sellers can raise prices. First, they can raise their money price, holding quality constant. Or, second, they can hold the money price constant while reducing the quality of the good.*** (The latter might also include reducing the size of the product, say, for example, a candy bar or a loaf of bread.) Faced with a price ceiling, sellers will use quality reductions as a way to raise their prices. Because of the government-created shortage, many consumers will buy the lower quality good rather than do without it.

It is important to note that a shortage is not the same as scarcity. ***Scarcity is inescapable.*** Scarcity exists whenever people want more of a good than nature has provided. This means, of course, that almost everything of value is scarce. ***Shortages, on the other hand, are a result of prices being set below their equilibrium values — situation that is avoidable if prices are permitted to rise.*** Removing the price ceiling will allow the price to rise back to its equilibrium level (P_0 rather than P_1 in Exhibit 2). This will stimulate additional production, discourage consumption, and increase the incentive of entrepreneurs to search for and develop substitute goods. This combination of forces will eliminate the shortage.

Rent Control: A Closer Look at a Price Ceiling

Rent controls are a price ceiling intended to protect residents from high housing prices. Rent controls are currently in place in many U.S. cities, including New York, Washington, D.C., Newark, New Jersey, and San Jose, California. Most of these measures were enacted during either World War II or the 1970s, when inflation was high. Rent controls peaked in the mid-1980s. At that time, more than 200 cities, encompassing about 20 percent of the nation's population, imposed rent controls.

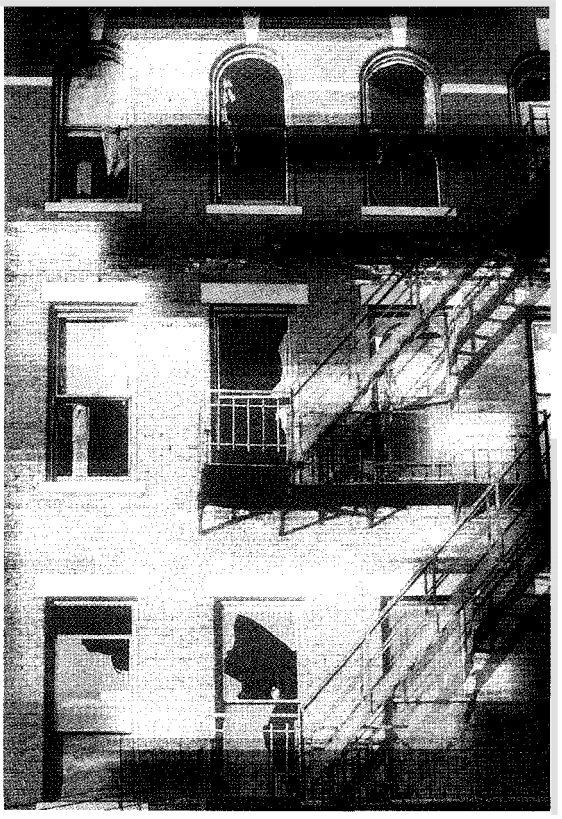
Because rent controls push the price of rental housing below the equilibrium level, the amount of rental housing demanded by consumers will exceed the amount landlords will make available. Initially, if the mandated price is only slightly below equilibrium, the impact of rent controls may be barely noticeable. Over time, however, the effects will worsen. Inevitably, rent controls will lead to the following results.

1. Shortages and black markets will develop. Because the quantity of housing demanded will exceed the quantity supplied, some people who value rental housing highly will be unable to find it. Frustrated by the shortage, they will try to induce landlords to rent to them. Some will agree to prepay their rent, including a substantial damage deposit. Others might agree to rent or buy the landlord's furniture at exorbitant prices in order to get an apartment. Still others will make under-the-table (black market) payments to secure housing.

2. The future supply of rental houses will decline. The below-equilibrium price will discourage entrepreneurs from constructing new rental housing units, and private investment will flow elsewhere. In the city of Berkeley, rental units available to students at the University of California dropped by 31 percent in the first five years after the city adopted rent controls in 1978.[4] In Boston and some of its suburbs, housing and apartment construction rose dramatically following the repeal of rent controls in the late 1990s. Similar results were observed in Santa Monica, California, following removal of rent controls.

[4] William Tucker, ***The Excluded Americans*** (Washington, D.C.: Regnery Gateway, 1990), 162. For additional information on rent controls, see William Tucker, "Rent Control Drives Out Affordable Housing," in ***USA Today Magazine*** (July 1998) and Walter Block, "Rent Controls," in ***Fortune Encyclopedia of Economics,*** ed. David Henderson (New York Warner Books, 1993). The latter publication can also be found online at http://www.econlib.org.

3. The quality of rental housing will deteriorate. When apartment owners are not allowed to raise their prices, they will use quality reductions to achieve this objective. Normal maintenance and repair service will deteriorate. Tenant parking lots will be eliminated (or rented out). Eventually, the quality of the rental housing will reflect the controlled price. Cheaper housing will be of cheaper quality.



© JOHN GRIFFIN/THE IMAGE WORKS

Rent controls lead to shortages, poor maintenance, and deterioration in the quality of rental housing.

4. Nonprice methods of rationing will become more important. Because price no longer rations rental housing, other forms of competition will develop. Landlords will rely more heavily on nonmonetary discriminating devices. They will favor friends, people of influence, and those whose lifestyles resemble their own. In contrast, applicants with many children or unconventional lifestyles, and perhaps racial minorities, will find fewer landlords who will rent to them. Since the cost to landlords of discriminating against those they do not like is lower, discrimination will become more prevalent in the rationing process. In New York City, where rent controls are in force, a magazine article suggested that "joining a church or synagogue" could help people make the connections they need to get an apartment. Can you imagine having to devote this amount of effort to finding an apartment? If your city enacts rent controls, you just might have to.

5. Inefficient use of housing space will result. The tenant in a rent-controlled apartment will think twice before moving. Why? Even though the tenant might want a larger or smaller space or an apartment closer to work, he or she will be less likely to move because it will be much more difficult to find a unit that's vacant. Turnover will be lower, and many people will find themselves in locations and in apartments not well suited to their needs. In a college town, students who live in the local area will have an advantage over newcomers. Local students and their parents will be more likely to have connections with apartment owners in the area. Many students from farther away, including those who value the apartments more highly, will find it extremely difficult to locate a place to rent.

6. Long-term renters will benefit at the expense of newcomers. People who stay for lengthy periods of time in the same apartment often pay rents substantially below market value, while newcomers are faced with exorbitant prices for units sublet from other tenants. Distortions and inequities result. A book on housing and the homeless by William Tucker reports several examples: Actress Ann Turkel spends only two months each year in New York. Turkel pays $2,350 per month for a seven-room, four-and-a-half bathroom duplex she has rented for many years. Identical apartments in Turkel's building sublet for $6,500 per month.[5] Former mayor Edward Koch pays $441.49 a month for a large one-bedroom apartment that would probably rent for $1,200 in the absence of rent controls.

[5] Tucker, ***The Excluded Americans,*** 248.

Koch kept the apartment the entire twelve years he lived in Gracie Mansion (the official mayor's residence) because had he given up the apartment, it would have been extremely difficult to find another one.

Imposing rent control laws may sound like a simple way to deal with high housing prices. However, the secondary effects are so damaging that many cities have begun repealing them. In the words of Swedish economist Assar Lindbeck: "In many cases, rent control appears to be the most efficient technique presently known to destroy a city—except for bombing."[6] Though this may overstate the case, Lindbeck definitely has a valid point.

The Impact of Price Floors

Price floor
A legally established minimum price buyers must pay for a good or resource.

A **price floor** establishes a minimum price that can legally be charged. The government imposes price floors on some agricultural products, for example, in an effort to artificially increase the prices that farmers receive. When a price floor is imposed above the current market equilibrium price, it will alter the market's operation. **Exhibit 3** illustrates the impact of imposing a price floor $(P,)$ for a product above its equilibrium level $(P,)$. At the higher price, the quantity supplied by producers increases along the supply curve to Q_S, while the quantity demanded by consumers decreases along the demand curve to Q. A **surplus** $(Q_S - Q.)$ of the good will result, as the quantity supplied by producers exceeds the quantity demanded by consumers at the new controlled price. Just like a price ceiling, a price floor reduces the quantity of the good exchanged, and reduces the gains from trade.

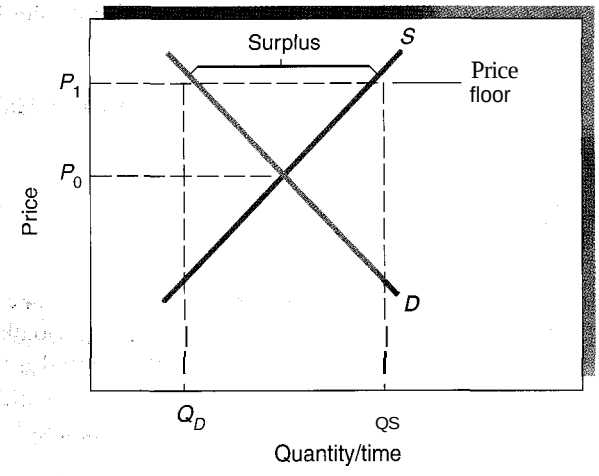
Surplus
A condition in which the amount of a good offered for sale by producers is greater than the amount that buyers will purchase at the existing price. A decline in price would eliminate the surplus.

As in the case of the price ceiling, nonprice factors will play a larger role in the rationing process. But because there is a surplus rather than a shortage, this time buyers will be in a position to be more selective. Buyers will purchase from sellers willing to offer them nonprice favors—better service, discounts on other products, or easier credit terms, for example. When it's difficult to alter the product's quality—in this case, improve it to make it more attractive for the price that must be charged—some producers will be unable to sell it.

It is important to note that a surplus doesn't mean the good is no longer scarce. People still want more of the good than is freely available from nature, even though they want less of it *at the controlled price* than sellers want to bring to the market. A decline in price would eliminate the surplus, but the item will be scarce in either case.

EXHIBIT 3
The Impact of a Price Floor

When a price floor such as P_1 keeps the price of a good or service above the market equilibrium, a surplus will result.



[6]Assar Lindbeck, *The Political Economy of the New Left* (New York Harper & Row, 1972), 39.

Minimum Wage: A Closer Look at a Price Floor

In 1938 Congress passed the Fair Labor Standards Act, which mandated a national **minimum wage** of 25 cents per hour. During the past 65 years, the minimum wage has been increased many times. The current federal minimum wage is $5.15 per hour, although some states have their own higher minimum-wage rates ranging up to just over $7 per hour.

Minimum wage
Legislation requiring that workers be paid at least the stated minimum hourly rate of pay.

A minimum wage is a price floor. Because most employees in the United States earn wages in excess of the minimum, their employment opportunities are largely unaffected by the minimum wage law. However, low-skilled and inexperienced workers whose equilibrium wage rates are lower than the minimum wage will be affected. **Exhibit 4** shows the direct effect of a $5.15-per-hour minimum wage on the employment opportunities of a group of low-skilled workers.

Without a minimum wage, the supply of and demand for these low-skilled workers would be in balance at a wage rate of $4.00. Because the minimum wage makes low-skilled labor more expensive, employers will substitute machines and more highly skilled workers for the now more expensive low-skilled employees. Fewer low-skilled workers will be hired when the minimum wage pushes their wages up. Graphically, this is reflected in the movement up along the demand curve in Exhibit 4 from the equilibrium point to the point along the demand curve associated with the higher, $5.15 wage rate (point **A**). The result will be a reduction in employment of low-skilled workers from E_0 to E_1.

On the supply side of the market, as the wages of low-skilled workers are pushed above equilibrium, there will be more unskilled workers looking for jobs. Graphically, this is reflected in the movement up along the supply curve in Exhibit 4 from the equilibrium point to the point along the supply curve associated with the higher, $5.15 wage rate (point ***B***). At the $5.15 wage rate, the quantity of workers searching for jobs will exceed the quantity of jobs available, causing excess supply. Economists generally use the term unemployment when referring to excess supply in a labor market.

In summary, economic analysis indicates that minimum-wage legislation increases the rate of unemployment among low-skilled workers. The exceedingly high unemployment rate of teenagers in the United States (one of the groups most affected by the minimum wage) is consistent with this analysis. In the United States, the unemployment rate for teenagers is more than three times the average for all workers, and the unemployment rate for black youth has generally exceeded 30 percent in recent years.

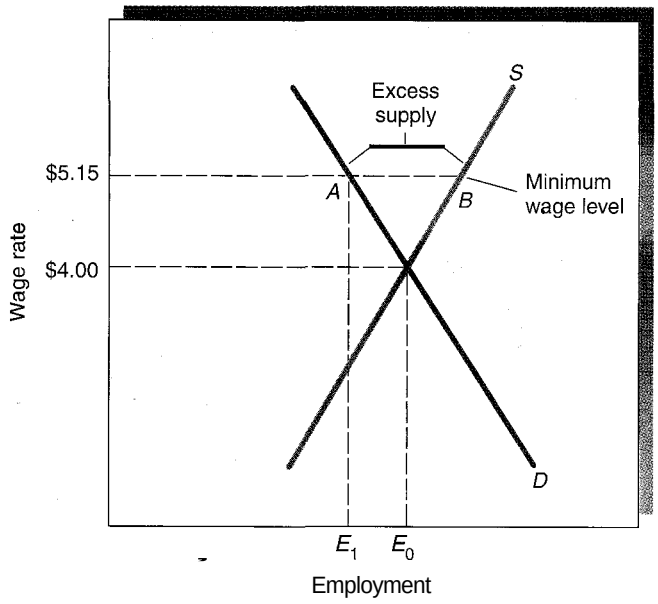


EXHIBIT 4
Employment and the Minimum Wage

If the market wage of a group of employees were $4.00 per hour, a $5.15-per-hour minimum wage will increase the earnings of workers able to retain their jobs, but reduce the employment opportunities of others as the number of jobs available shrinks from E_0 to E_1.

It is important to remember that the market price—the wage rate—is only one dimension of the transaction. When a price floor pushes the wage rate above equilibrium, employers will have less of an incentive to offer nonwage benefits to employees because they will have no trouble hiring low-skilled workers. Predictably, a higher minimum wage will lead to a deterioration of the nonwage qualities of minimum-wage jobs, and so workers in these jobs will experience less convenient working hours, fewer training opportunities, and less continuous employment.

The adverse impact of the minimum wage on the opportunity of youthful workers to acquire work experience and on-the-job training is a particularly important unintended consequence of minimum-wage laws. Low-paying, entry-level jobs can provide workers with experience that will help them move up the job ladder to higher-paying positions. Employment experience obtained at an early age, even on menial tasks, can help a person acquire self-confidence, good work habits, and skills that make them more valuable to future employers. The minimum wage makes this more difficult. Not only does the minimum wage make it harder for low-skilled workers to find jobs, it also reduces their on-the-job training opportunities. In order to pay the higher wage rate required by the law, employers will have to find other ways to cut employment costs, like reducing the amount of job training. Not surprisingly, most minimum-wage jobs are dead-end positions with little opportunity for future advancement.[7]

Workers who are able to maintain their employment at the higher minimum-wage rate—most likely the better qualified among those with low skill levels—gain from a minimum wage. But other low-skilled workers are harmed by the minimum wage, particularly those with the lowest skill levels, who will find it more difficult to get jobs.

How many fewer low-skilled workers are hired because of the minimum wage? Studies indicate that a 10 percent increase in the minimum wage reduces the employment of low-skilled workers by 1 to 3 percent. Minimum-wage supporters argue that the higher wages for low-skilled workers are worth this reduction in employment and job-training opportunities. Critics argue, however, that the reduced job opportunities for the lowest-skilled workers are reason enough to eliminate the minimum wage.

Does the minimum wage help the poor? Most minimum-wage workers belong to families with an income substantially above the poverty line. In fact, about 40 percent of minimum-wage workers are members of a family with an income in the top half of the income distribution. The typical minimum-wage worker is a spouse or a teenage member of a household with an income well above the poverty level. Therefore, even if the effects of a higher minimum wage on employment and nonwage forms of compensation were small, a higher minimum wage does little to help the poor, making it a much less attractive antipoverty program than other alternatives.[8]

BLACK MARKETS AND THE IMPORTANCE OF THE LEGAL STRUCTURE

When price controls are imposed, exchanges at prices outside of the range set by the government are illegal. Governments may also make it entirely illegal to buy and sell certain products. This is the case with drugs like marijuana and cocaine in the United States. Similarly, prostitution is illegal in all states except Nevada. However, controlling prices and making a good or service illegal doesn't eliminate market forces. When demand is strong and gains from trade can be had, markets will develop and exchanges will occur in spite of the restrictions. People will also engage in illegal exchanges in order to evade taxes. For example, the $3.39 per-pack cigarette tax in New York City has made cigarette smuggling in that city a thriving business.

[7]For evidence that the minimum wage limits training opportunities, see David Neumark and William Wascher, "Minimum Wages and Training Revisited," *Journal of Labor Economics* 19 (July 2001): 563–95.

[8]See William E. Even and David A. Macpherson, "Consequences of Minimum Wage Indexing," *Contemporary Economic Policy* 14 (October 1996): 67–77; David Neumark and William Wascher, "Do Minimum Wages Fight Poverty?" *Economic Inquiry* 40 (July 2002): 315–33; and David Neumark and William Wascher, "The Effects of Minimum Wages Throughout the Wage Distribution," *Journal of Human Resources* 39 (April 2004): 425–50, for evidence on this point.

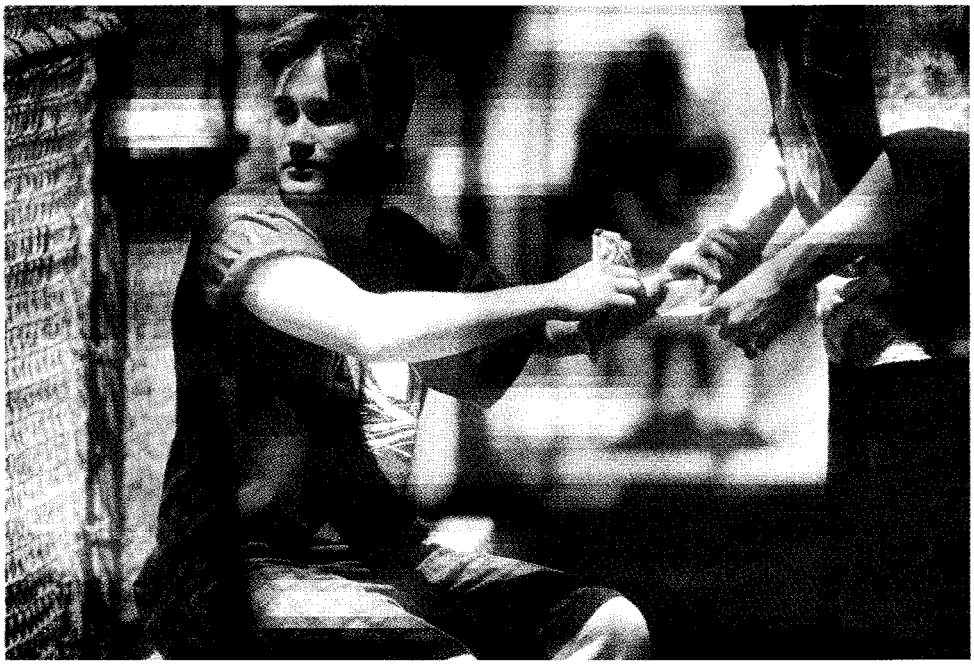
Markets that operate outside the legal system are called **black markets**. How do black markets work? Can markets function without the protection of the law? As in other markets, supply and demand will determine prices in black markets, too. However, because black markets operate outside the official legal structure, enforcement of contracts and the dependability of quality will be less certain. Furthermore, participation in black markets involves greater risk, particularly for suppliers. Prices in these markets will have to be higher than they otherwise would be to compensate suppliers for the risks they are taking—the threat of arrest, possibility of a fine or prison sentence, and so on. Perhaps most important, in black markets there are no legal channels for the peaceful settlement of disputes. When a buyer or seller fails to deliver, it is the other party who must try to enforce the agreement, usually through the use or threat of physical force.

Black market
A market that operates outside the legal system, where either illegal goods are sold or legal goods are sold at illegal prices or terms.

Compared to normal markets, black markets are characterized by a higher incidence of defective products, higher profit rates (for those who do not get caught), and more violence. The incidence of phony tickets purchased from street dealers selling them at illegal prices, and deaths caused by toxic, illicit drugs, are a reflection of the high presence of defective goods in these markets. Certainly the expensive clothes and automobiles of many drug dealers suggest that monetary profits are high in black-markets. Evidence of violence as a means of settling disputes arising from black market transactions is widespread. Crime statistics in urban areas show that a high percentage of the violent crimes, including murder, are associated with illegal trades gone bad and competition among dealers in the illegal drug market.

The prohibition of alcohol in the United States from 1920 to 1933 vividly illustrates how violence, deception, and fraud plague markets that operate outside the law. When the production and sale of alcohol was illegal during the Prohibition era, gangsters dominated the alcohol trade, and the murder rate soared to record highs. There were also problems with product quality (tainted or highly toxic mixtures, for example) similar to the ones present in modern-day illegal-drug markets. When Prohibition was repealed and the market for alcoholic beverages began operating once again within the legal framework, these harmful secondary effects disappeared.

The operation of black markets highlights a point often taken for granted: a legal system that provides for secure private-property rights, contract enforcement, and access to an unbiased court system for settling disputes is vitally important for the smooth operation of markets. Markets will exist in any environment, but they can be



GETTY IMAGES

Black markets like those for illegal drugs are characterized by less dependable product quality and the greater use of violence to settle disputes between buyers and sellers.

counted on to function efficiently only when property rights are secure and contracts are enforced in an evenhanded manner.

The analysis of black markets also provides insights into the economies of Russia, Ukraine, and other parts of the former Soviet Union. Following the collapse of communism, the legal systems in these areas reflected the prior socialist nature of these economies. Both the protection of private property and the enforcement of contracts between private parties were highly uncertain. People with political connections were often able to escape their contractual responsibilities and obtain favorable rulings from legal and regulatory authorities. As a result, markets in these countries operated much like black markets. Fraud and deception were commonplace, and the incidence of violence related to business dealings was widespread. A market economy—like any other—does not work well in such an environment. Many foreign businesspeople who were initially attracted to markets in these countries soon began packing their bags and returning home. Financial capital fled, investment shrank, and these economies performed poorly. During the last few years, there has been some improvement in the legal environment, but the enforcement of contracts and conduct of business activities in the former Soviet Union still involve considerable risk. Without sound legal systems, these countries will be unable to reap the full benefits of a market economy.

THE IMPACT OF A TAX

Tax incidence
The way the burden of a tax is distributed among economic units (consumers, producers, employees, employers, and so on). The actual tax burden does not always fall on those who are statutorily assigned to pay the tax.

How do taxes affect market exchange? When governments tax goods, who bears the burden? Economists use the term **tax incidence** to indicate how the burden of a tax is *actually* shared between buyers (who pay more for what they purchase) and sellers (who receive less for what they sell). When a tax is imposed, the government can make either the buyer or the seller legally responsible for payment of the tax. The legal assignment is called the *statutory incidence* of the tax. However, the person who writes the check to the government—that is, the person statutorily responsible for the tax—is not always the one who bears the tax burden. The *actual incidence* of a tax may lie elsewhere. If, for example, a tax is placed statutorily on a seller, the seller might simply increase the price of the product. In this case, the buyers end up bearing some, or all, of the tax burden through the higher price.

To illustrate, **Exhibit 5** shows how a $1,000 tax placed on the sale of used cars would affect the market. (To simplify this example, let's assume all used cars are identical.) Here, the tax has statutorily been placed on the seller. When a tax is imposed on the seller, it shifts the supply curve upward by exactly the amount of the tax—$1,000 in this example. To understand why, remember that the height of the supply curve at a particular quantity shows the minimum price required to cause enough sellers to offer that quantity of cars for sale. Suppose you were a potentiai seller, willing to sell your car for any price over $6,000, but you were unwilling to sell it unless you could pocket at least $6,000 from the sale. Because you now have to pay a tax of $1,000 when you sell your car, the minimum price you will accept *from the buyer* will rise to $7,000, so that after paying the tax, you will retain $6,000. Other potential sellers will be in a similar position. The tax will push the minimum price each seller is willing to accept upward by $1,000. Thus, the after-tax supply curve will shift vertically by this amount.

Sellers would prefer to pass the entire tax on to buyers by raising prices by the full amount of the tax, rather than paying any part of it themselves. However, as sellers begin to raise prices, customers respond by purchasing fewer units. At some point, to avoid losing additional sales, sellers will find it more profitable to accept part of the tax burden themselves (in the form of a lower price net of tax), rather than to raise the price by the full amount of the tax. This process is shown in Exhibit 5.

Before the tax was imposed, used cars sold for a price of $7,000 (at the intersection of the original supply and demand curves shown by point **A**). After the $1,000 tax is imposed, the equilibrium price of used cars will rise to $7,400 (to point *B*, the intersection

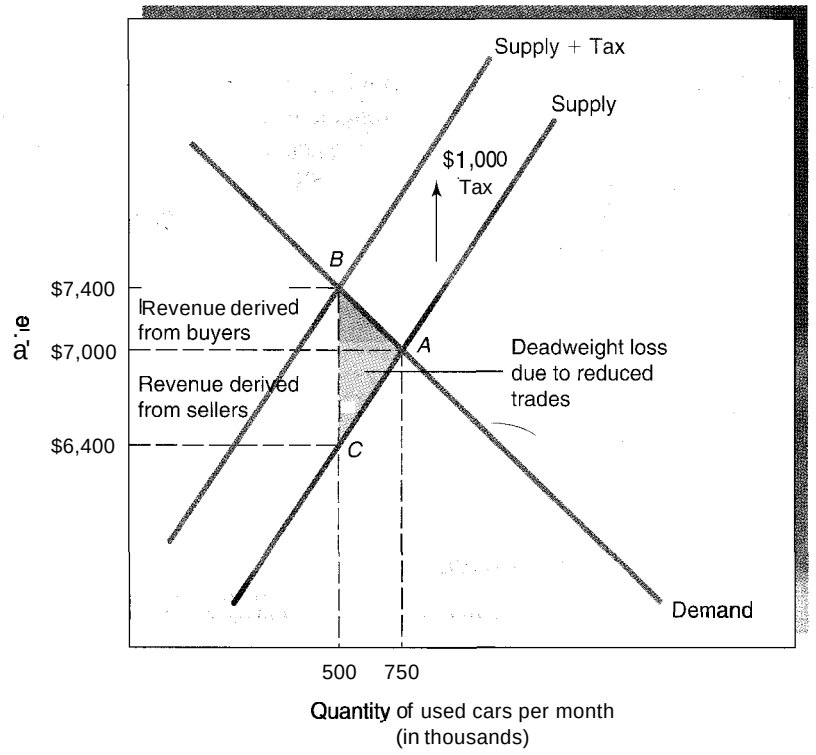


EXHIBIT 5
The Impact of a Tax Imposed on Sellers

When a $1,000 tax is imposed statutorily on the sellers of used cars, the supply curve shifts vertically upward by the amount of the tax. The price of used cars to buyers rises from $7,000 to $7,400, resulting in buyers bearing $400 of the burden of this tax. The price received by a seller falls from $7,000 to $6,400 ($7,400 minus the $1,000 tax), resulting in sellers bearing $600 of the burden.

of the new supply curve including the tax, and the demand curve). Thus, despite the tax being statutorily imposed on sellers, the higher price shifts some of the tax burden to buyers. Buyers will now pay $400 more for used cars. Sellers now receive $7,400 from the sale of their used cars. However, after sending $1,000 in taxes to the government, they retain only $6,400. This is exactly $600 less than the seller would have received had the tax not been imposed. Because the distance between the supply curves is exactly $1,000, this net price can be found in Exhibit 5 by following the vertical line down from the new equilibrium (point *B*) to the original supply curve (point C) and over to the price axis. In this case, each $1,000 of tax revenue transferred to the government imposes a burden of $400 on buyers (in the form of higher used-car prices) and a $600 burden on sellers (in the form of lower net receipts from a car sale)—even though sellers are responsible for actually sending the $1,000 tax payment to the government.

The tax revenue derived from a tax is equal to the **tax base** (in this case, the number of used cars exchanged) multiplied by the **tax rate.** After the tax is imposed, the quantity exchanged will fall to 500,000 cars per month because some buyers will choose not to purchase at the $7,400 price, and some sellers will decide not to sell when they are able to net only $6,400. Given the after-tax quantity sold, the monthly revenue derived from the tax will be $500 million (500,000 cars multiplied by $1,000 tax per car).

Tax base
The level or quantity of an economic activity that is taxed. Higher tax rates reduce the level of the tax base because they make the activity less attractive.

Tax rate
The per-unit amount of the tax or the percentage rate at which the economic activity is taxed.

The Deadweight Loss Caused by Taxes

As Exhibit 5 shows, a $1,000 tax on used cars causes the number of units exchanged to fall from 750,000 to 500,000. It reduces the quantity of units exchanged by 250,000 units. Remember, trade results in mutual gains for both buyers and sellers. The loss of the mutual benefits that would have been derived from these additional 250,000 units also imposes a cost on buyers and sellers. But this cost—the loss of the gains from trade eliminated by the tax—does not generate any revenue for the government. Economists call this the **deadweight loss** of taxation. In Exhibit 5, the size of the triangle *ABC* measures the deadweight loss. The deadweight loss is a burden imposed on buyers and sellers over

Deadweight loss
The loss of gains from trade to buyers and sellers that occurs when a tax is imposed. The deadweight loss imposes a burden on both buyers and sellers over and above the actual payment of the tax.

Excess burden of taxation Another term for deadweight loss. It reflects losses that occur when beneficial activities are forgone because they are taxed.

and above the cost of the revenue transferred to the government. Sometimes it is referred to as the **excess burden of taxation**. It is composed of losses to both buyers (the lost consumer surplus consisting of the upper part of the triangle *ABC)* and sellers (the lost producer surplus consisting of the lower part of the triangle *ABC)*.

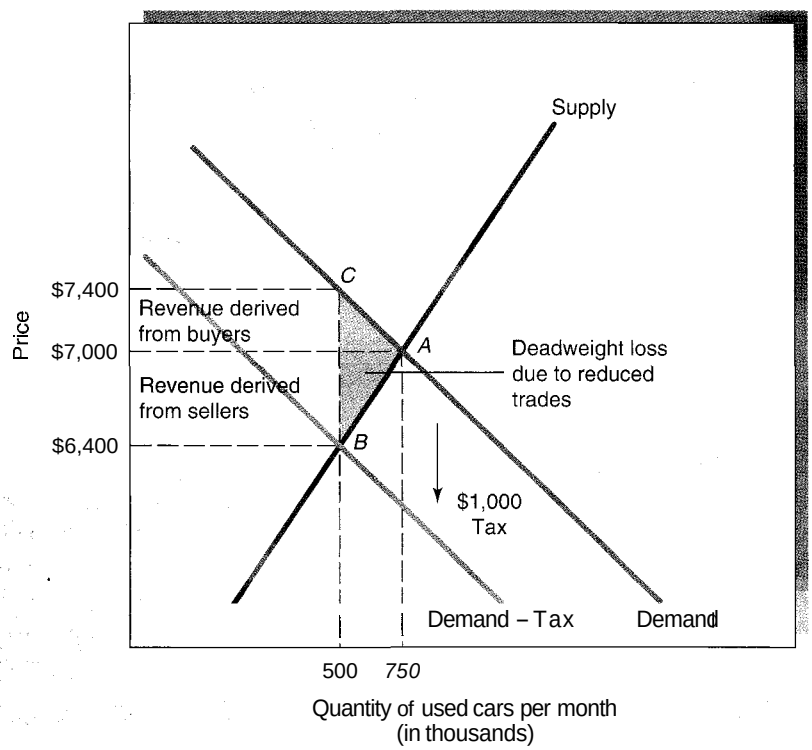
The deadweight loss to sellers includes an indirect cost imposed on the people who supply resources to that industry (such as its suppliers and employees). The 1990 luxury-boat tax provides a good example. Supporters of the luxury-boat tax assumed the tax burden would fall primarily on wealthy yacht buyers. The actual effects were quite different, though. Because of the tax, luxury-boat sales fell sharply and thousands of workers lost their jobs in the yacht-manufacturing industry. The deadweight loss triangle might seem like an abstract concept, but it wasn't so abstract to the employees in the yacht industry who lost their jobs! Their losses are part of what is reflected in the triangular area. Moreover, because luxury-boat sales declined so sharply, the tax generated only a meager amount of revenue. The large deadweight loss (or excess burden) combined with meager revenue for the government eventually led to the repeal of the tax.

Actual Versus Statutory Incidence

Economic analysis indicates that the actual burden of a tax—or more precisely, the split of the burden between buyers and sellers—does not depend on whether the tax is statutorily placed on the buyer* or *the seller. To see this, we must first look at how the market responds to a tax statutorily placed on the buyer. Continuing with the auto tax example, let's suppose that the government places the $1,000 tax on the buyer of the car, rather than the seller. After making a used-car purchase, the buyer must send a check to the government for $1,000. Imposing a tax on buyers will shift the demand curve downward by the amount of the tax, as shown in **Exhibit 6.** This is because the height of the demand curve represents the maximum price a buyer is willing to pay for the car. If a particular buyer is willing and able to pay only $5,000 for a car, the $1,000 tax would mean that the most the

EXHIBIT 6
The Impact of a Tax Imposed on Buyers

When a $1,000 tax is imposed statutorily on the buyers of used cars, the demand curve shifts vertically downward by the amount of the tax. The price of used cars falls from $7,000 to $6,400, resulting in sellers bearing $600 of the burden. The buyer's total cost of purchasing the car rises from $7,000 to $7,400 ($6,400 plus the $1,000 tax), resulting in buyers bearing $400 of the burden of this tax. The incidence of this tax on used cars is the same regardless of whether it is statutorily imposed on buyers or sellers.



The actual burden of a tax is independent of whether it is imposed on buyers or sellers.

BY JOHN TREVER, *ALBUQUERQUE JOURNAL*. REPRINTED BY PERMISSION.

"THIS NEW TAX PLAN SOUNDS PRETTY GOOD... WE GET A 9% CUT AND BUSINESS PICKS UP THE BURDEN...."

buyer would be willing to pay ***to the seller*** would be $4,000. This is because the total cost to the buyer is now the purchase price plus the tax.

As Exhibit 6 shows, the price of used cars falls from $7,000 (point **A**) to $6,400 (point *B*) when the tax is statutorily placed on the buyer. Even though the tax is placed on buyers, the reduction in demand that results causes the price received by sellers to fall by $600. Thus, $600 of the tax is again borne by sellers, just as it was when the tax was placed statutorily on them. From the buyer's standpoint, a car now costs $7,400 ($6,400 paid to the seller plus $1,000 in tax to the government). Just as when the tax was imposed on the seller, the buyer now pays $400 more for a used car.

A comparison of Exhibits 5 and 6 makes it clear that the actual burden of the $1,000 tax is independent of its statutory incidence. In both cases, buyers pay a total price of $7,400 for the car (a $400 increase from the pretax level), and sellers receive $6,400 from the sale (a $600 decrease from the pretax level). Correspondingly, the revenue derived by the government, the number of sales eliminated by the tax, and the size of the deadweight loss are identical whether the law requires payment of the tax by the sellers or by the buyers. A similar phenomenon occurs with any tax. The 15.3 percent Social Security payroll tax, for example, is statutorily levied as 7.65 percent on the employee and 7.65 percent on the employer. The impact is to drive down the net pay received by employees and raise the employers' cost of hiring workers. Economic analysis tells us that the actual burden of this tax will probably differ from its legal assignment, and that it will be the same regardless of how the tax is statutorily assigned. Because market prices (here, workers' gross wage) will adjust, the incidence of the tax will be identical regardless of whether the 15.3 percent is levied on employees or on employers or is divided between the two parties.

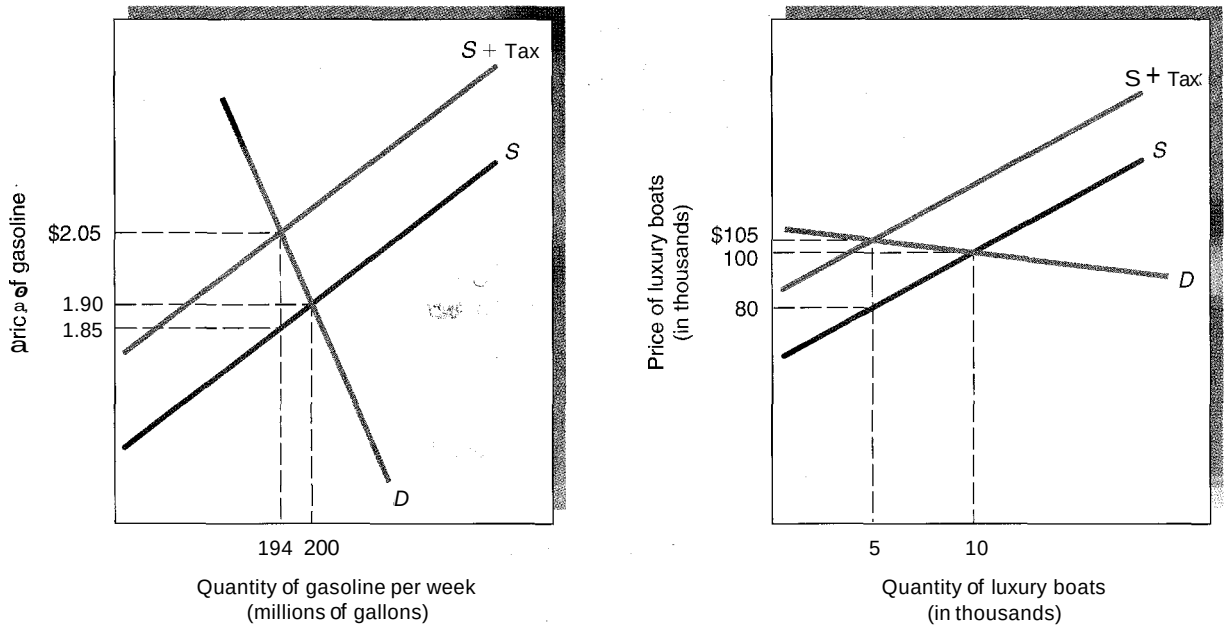
Elasticity and the Incidence of a Tax

If the actual incidence of a tax is independent of its statutory assignment, what does determine the incidence? The answer: ***The incidence of a tax depends on the responsiveness of buyers and of sellers to a change in price.*** When buyers respond to even a small increase in price by leaving the market and buying other things, they will not be willing to accept a price that is much higher than it was prior to the tax. Similarly, if sellers respond to a small reduction in what they receive by shifting to the production of other goods or going out of business, they will not be willing to accept a much smaller payment. net of tax. The burden of a tax—its incidence—tends to fall more heavily on whichever side of the market has the least attractive options elsewhere—the side of the market that is less sensitive to price changes, in other words.

In the preceding chapter, we saw that the steepness of the supply and demand curves reflects how responsive producers and consumers are to a price change. Relatively inelastic

EXHIBIT 7
How the Burden of a Tax Depends on the Elasticities of Demand and Supply

In part (a), when demand is relatively more inelastic than supply, buyers bear a larger share of the burden of the tax. In part (b), when supply is relatively more inelastic than demand, sellers bear a larger share of the tax burden.



demand or supply curves are steeper (more vertical), indicating less responsiveness to a change in price. Relatively elastic demand or supply curves are flatter (more horizontal), indicating a higher degree of responsiveness to a change in price.

Using gasoline as an example, part (a) of **Exhibit 7** illustrates the impact of a tax when demand is relatively inelastic and supply is relatively elastic. It will not be easy for gasoline consumers to shift—particularly in the short run—to other fuels in response to an increase in the price of gasoline. The inelastic demand curve shows this. When a 20-cent tax is imposed on gasoline, buyers end up paying 15 cents more per gallon ($2.05 instead of $1.90), while the net price of sellers is 5 cents less ($1.85 instead of $1.90). ***When demand is relatively inelastic, or supply is relatively elastic, buyers will bear the larger share of the tax burden.***

Conversely, when demand is relatively elastic and supply is inelastic, more of the tax burden will fall on sellers and resource suppliers. The luxury-boat tax illustrates this point. As we mentioned earlier, Congress imposed a tax on the sale of luxury boats in 1990. Later, the tax was repealed because of its adverse impact on sales and employment in the industry. There are many things wealthy potential yacht owners can spend their money on other than luxury boats *sold* in *the United States*. For one thing, they can buy a yacht someplace else, perhaps in Mexico, England, or the Bahamas. Or they can spend more time on the golf course, travel to exotic places, or purchase a nicer car or more expensive home. Because there are attractive substitutes, the demand for domestically produced luxury boats is relatively elastic compared to supply. Therefore, as Exhibit 7b illustrates, when a $25,000 tax is imposed on luxury boats, prices rise by only $5,000 (from $100,000 to $105,000), but output falls substantially (from 10,000 to 5,000 boats). The net price received by sellers falls by $20,000 (from $100,000 to $80,000 per boat). ***When demand is relatively elastic, or supply is relatively inelastic, sellers (including resource suppliers) will bear the larger share of the tax burden.***

Elasticity and the Deadweight Loss

We have seen that the elasticities of supply and demand determine how the burden of a tax is distributed between buyer and seller. They also influence the size of the deadweight loss caused by the tax because they determine the total reduction in the quantity exchanged. When either demand or supply is relatively inelastic, fewer trades will be eliminated by the tax, so the deadweight loss will be smaller. From a policy perspective, the excess burden of a tax system will therefore be lower if taxes are levied on goods and services for which either demand or supply is highly inelastic.

TAX RATES, TAX REVENUES, AND THE LAFFER CURVE

It is important to distinguish between the average and marginal rates of taxation. They can be very different, and both provide important information. The average tax rate is generally used to examine how different income groups are burdened by a tax, whereas the marginal tax rate is the key to understanding the negative economic effects created by a tax. Both can be computed with simple equations. The **average tax rate** (ATR) can be expressed as follows:

Average tax rate (ATR)
Tax liability divided by taxable income. It is the percentage of income paid in taxes.

$$\text{ATR} = \text{Tax liability} / \text{Taxable income}$$

For example, if a person's tax liability was $3,000 on an income of $20,000, her average tax rate would be 15 percent ($3,000 divided by $20,000). The average tax rate is simply the percentage of income that is paid in taxes.

In the United States, the personal income tax provides the largest single source of government revenue. This tax is particularly important at the federal level. You may have heard that the federal income tax is "progressive." A **progressive tax** is defined as a tax in which the average tax rate rises with income. In other words, people with higher income pay a larger *percentage of their income* in taxes. Alternatively, taxes can be proportional or regressive. A **proportional tax** is defined as a tax in which the average tax rate remains the same across income levels. Under a proportional tax, everyone pays the same percentage of their income in taxes. Finally, a **regressive tax** is defined as a tax in which the average tax rate falls with income. If someone making $100,000 per year paid $30,000 in taxes (an ATR of 30 percent) while someone making $30,000 per year paid $15,000 in taxes (an ATR of 50 percent), the tax code would be regressive. Note that a regressive tax merely means that the *percentage* paid in taxes declines with income; the actual dollar amount of the tax bill might still be higher for those with larger incomes.

Progressive tax
A tax in which the average tax rate rises with income. People with higher incomes will pay a higher percentage of their income in taxes.

Proportional tax
A tax in which the average tax rate is the same at all income levels. Everyone pays the same percentage of income in taxes.

Regressive tax
A tax in which the average tax rate falls with income. People with higher incomes will pay a lower percentage of their income in taxes.

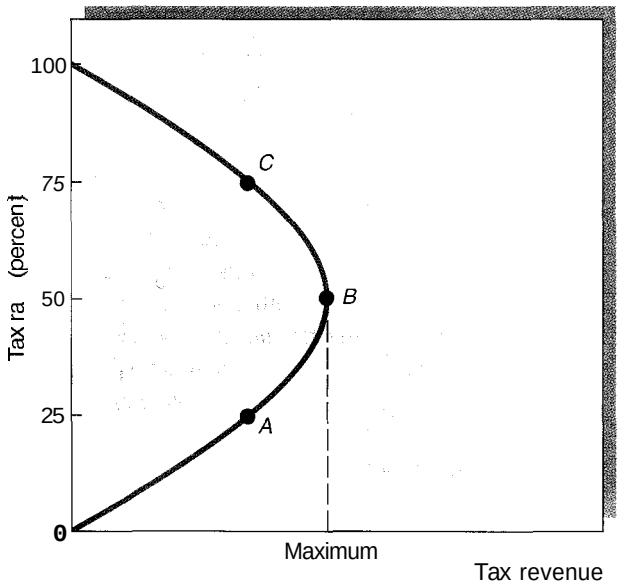
Although the average tax rate is useful in determining whether a tax is progressive, proportional, or regressive, it is the marginal tax rate that concerns individuals when they are making decisions. It is the marginal tax rate that determines how much of an additional dollar of income must be paid in taxes (and thus, also, how much one gets to keep). An individual's marginal tax rate can be very different from his or her average tax rate. The **marginal tax rate (MTR)** can be expressed as follows:

Marginal tax rate (MTR)
The additional tax liability a person faces divided by his or her additional taxable income. It is the percentage of an extra dollar of income earned that must be paid in taxes. It is the marginal tax rate that is relevant in personal decision making.

$$\text{MTR} = \text{Change in tax liability} / \text{Change in taxable income}$$

The MTR reveals both how much of one's *additional* income must be turned over to the tax collector and how much is retained by the individual taxpayer. For example, when the MTR is 25 percent, $25 of every $100 of additional earnings must be paid in taxes. The individual is permitted to keep only $75 of his or her additional income, in other words. The marginal tax rate is vitally important because it affects the incentive to earn additional income. The higher the marginal tax rate, the less incentive individuals have to earn more income. At high marginal rates, for example, many spouses will choose to stay home rather than take a job, and others will choose not to take on second jobs or extra work. **Exhibit 8** shows the calculation of both the average and marginal tax rates within the framework of the 2003 income tax tables provided to taxpayers.

EXHIBIT 9
Laffer Curve



Because taxing an activity affects the amount of it people will do, a change in tax rates will not lead to a proportional change in tax revenues. As the Laffer curve indicates, beyond some point (*B*), an increase in tax rates will cause tax revenues to fall. At high tax rates, revenue can be increased by lowering tax rates. The tax rate that maximizes tax revenue is higher than the ideal tax rate for the economy as a whole because of the large deadweight loss of taxation as tax rates increase toward point *B*.

from zero, tax revenue expands. Clearly, at some rate greater than zero but less than 100 percent, tax revenue will be maximized (point *B* in Exhibit 9). This is not to imply that the tax rate that maximizes revenue is the ideal, or optimal, tax rate from the standpoint of the economy as a whole. Although it might be the tax rate that generates the most revenue for government, we must also consider the welfare reductions imposed on individuals by the deadweight loss created by the tax. As rates are increased and the maximum revenue point *(B)* is approached, relatively large tax rate increases will be necessary to expand tax revenue by even a small amount. In this range, the deadweight loss of taxation in the form of reductions in gains from trade will be exceedingly large relative to the additional tax revenue. Thus, the ideal tax rate will be well below the rate that maximizes revenue.

The Laffer curve shows that it is important to distinguish between changes in tax rates and changes in tax revenues. Higher rates will not always lead to more revenue for the government. Similarly, lower rates will not always lead to less revenue. When tax rates are already high, a rate reduction may increase tax revenues. Correspondingly, increasing high tax rates may lead to less tax revenue.

Evidence from the sharp reduction in marginal tax rates imposed on those with high incomes during the 1980s supports the Laffer curve. The top marginal rate was reduced from 70 percent at the beginning of the decade to 33 percent by the end of the decade. A person in this tax bracket who used to bring home 30 cents for each additional dollar earned now was able to bring home 67 cents for each dollar earned. Focusing on this sharp seduction in the top marginal rate, critics charged that the tax cuts of the 1980s were a bonanza for the rich. When considering the validity of this charge, however, it is important to distinguish between tax rates and tax revenues. Even though the top sates were cut sharply, tax revenues and the share of the personal income tax paid by high-income earners actually rose as a result. During the decade, revenue collected from the top 1 percent of earners rose a whopping 51.4 percent (after adjusting for inflation). In 1980, 19 percent of the personal income tax was collected from the top 1 percent of earners. By 1990 at the

The Laffer Curve and Mountain-Climbing Deaths

The Laffer curve can be used to illustrate many other relationships besides just tax rates and tax revenues. Economists J. R. Clark and Dwight Lee have used it to analyze the relationship between the safety of mountain climbing and mountain-climbing deaths on Mt. McKinley, North America's highest peak. As the risk of dying from climbing Mt. McKinley fell due to greater search-and-rescue efforts by National Park personnel, the number of people seeking *to* "conquer the mountain" rose significantly. The increase in the number of climbers attempting to conquer the mountain offset the lower risk, leading to a Laffer curve–type relationship. In other words, greater search-and-rescue efforts led to a *higher* number of total deaths on the mountain.

Let's look at the problem numerically. Assume that if the probability of death from an attempted climb were 90 percent, only 100 people would attempt to climb the mountain each year, leading to an annual death rate of 90. Now suppose that greater search-and-rescue efforts lower the probability of death to 50 percent. Because incentives matter, the increased safety will result in an increase in the number of people attempting to climb the mountain. Suppose that the number of climbers increases from 100 to 200. With 200 climbers and a 50 percent probability of death, the annual number of fatalities would increase to 100, 10 more than before rescue efforts were improved. The total number of mountain-climbing deaths is actually lowest when there is both a very high and a very low probability of death—just as the Laffer curve predicts. The number of deaths is largest in the middle probability ranges. Making a very risky mountain safer can therefore result in more rather than fewer fatalities.

Clark and Lee have also explored a similar relationship between average lengths of prison sentences and total prison space occupied. Other economists have explored the Laffer curve relationship between the minimum-wage and the earnings of minimum-wage workers, as well as the regulatory costs of protecting endangered species and the habitat acres available to them.[1]

[1]See J R Clark and Dwight R. Lee, "Too Safe to Be Safe: Some Implications of Short- and Long-Run Rescue Laffer Curves," *Eastern Economic Journal* 23 no. 2 (Spring 1997) 127–37; Russell S. Sobel, "Theory and Evidence on the Political Economy of the Minimum Wage,' *Journal of Political Economy* 107 no 4 (August 1999): 761–85; and Richard L Stroup, 'The Endangered Species Act The Laffer Curve Strikes Again," *The Journal* of *Private Enterprise*, vol XIV (Special Issue 1998): 48–62.

lower tax rates, the top 1 percent of earners accounted for more than 25 percent of income tax revenues. The top 10 percent of earners paid just over 49 percent of total income taxes in 1980, but by 1990 the share paid by these earners had risen to 55 percent. Thus, the reduction in the exceedingly high rates increased the revenue collected from high-income taxpayers.

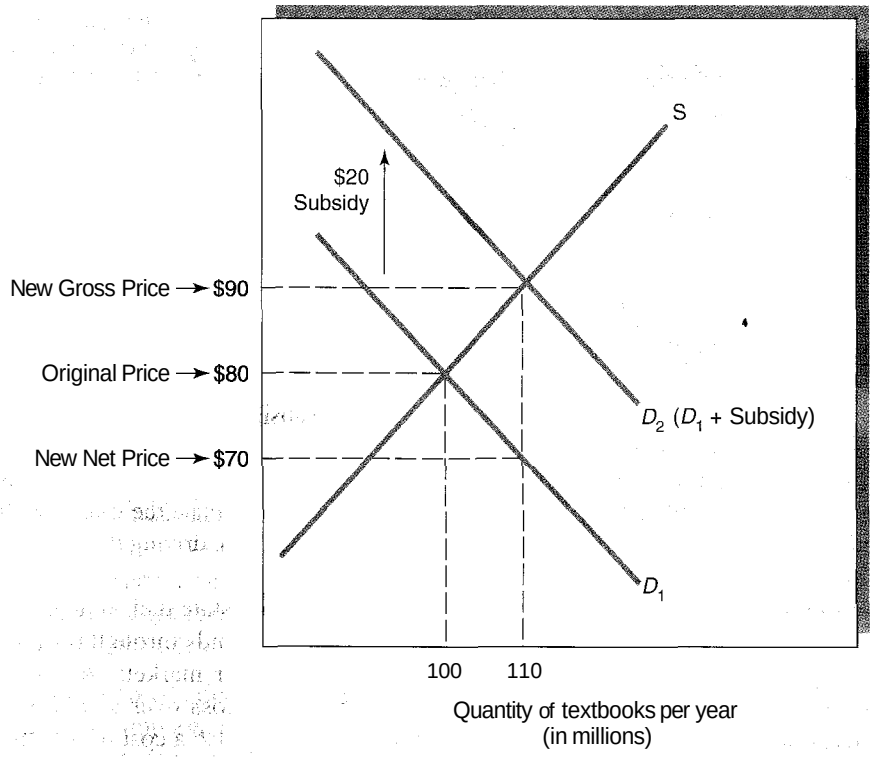
THE IMPACT OF A SUBSIDY

Subsidy
A payment the government makes to either the buyer or seller, usually on a per-unit basis, when a good or service is purchased or sold.

The supply and demand framework can also be used to analyze the impact of a government **subsidy. A** subsidy is a payment to either the buyer or seller of a good or service, usually on a per-unit basis. Subsidies are often granted in an effort to help buyers afford a good or service, or to increase the profitability of producers in an industry. As we have seen in other cases, however, the effect of a government program often differs substantially from its original intent. Because prices change when subsidies are imposed (just as when taxes are imposed), the benefit of a subsidy can be partially, or totally, shifted from buyer to seller, or vice versa.

Suppose that the government, in an effort to make textbooks more affordable, gives college students (buyers) a $20 subsidy for each book they buy. **Exhibit 10** shows the effect of the program. Before the subsidy was instituted, 100 million textbooks were sold each year at an average price of $80 per book. The $20-per-book subsidy paid to the buyers will increase demand by the amount of the subsidy (shift from D_1 to D_2). As the result

EXHIBIT 10
The Impact of a Subsidy Granted to Buyers



When a $20-per-textbook subsidy is given to students, the demand curve for textbooks shifts vertically upward by the amount of the subsidy. The market price of textbooks rises from $80 to $90 (new gross price). With the $20 subsidy, buyers now pay a net price of $70 per textbook (the new $90 price minus the $20 subsidy), which is $10 less than before. Textbook buyers get only $10 of the benefit of the subsidy; the remaining $10 benefit accrues to the supply side of the market (sellers and resource suppliers) in the form of higher textbook prices. The distribution of the benefit from the subsidy between buyers and sellers would be the same, whether it was granted to buyers or sellers.

of the subsidy, the equilibrium price will increase from $80 to $90, and the total quantity purchased will expand to 110 million textbooks per year.

The subsidy program reduces the students' out-of-pocket cost of a textbook (from $80 to $70), but the net gain to them is less than the amount of the subsidy. Why? Even though the textbook subsidy is granted to buyers, substantial benefits also accrue to sellers. Because the subsidy program increases the demand for textbooks, pushing their price upward by $10, half of the benefits are captured by sellers (including resource suppliers like copy editors, authors, and paper suppliers).

Alternatively, if textbook supplier? had been granted a $20 payment from the government for each textbook sold, the supply curve would have shifted downward by the amount of the subsidy. This would cause the market price of textbooks to decline to $70. In this case, buyers pay $10 less than before the subsidy program, while the sellers receive $10 more (the sellers now get $90 for each book sold—the $70 market price plus the $20 government subsidy). Just like a tax, a subsidy results in the same outcome, regardless of whether the subsidy is granted to buyers or sellers.

Elasticity and the Benefit of Government Subsidy Programs

In this example, the benefit of the $20-per-textbook subsidy was split evenly between buyers and sellers. However, the actual distribution of this benefit will depend on the elasticities of supply and demand—just as it does with a tax. The benefit of the subsidy will always be shifted toward the more inelastic side of the market. Thus, the more inelastic the supply, the larger the share of the benefit that will accrue to sellers. On the other hand,

the more inelastic the demand, the larger the share of the benefit that will accrue to buyers. Using our earlier examples from the section on taxation, consumers would be the main beneficiary of a subsidy on gasoline (a good for which the demand is relatively inelastic, and supply elastic), while suppliers would be the main beneficiary of a subsidy on luxury boats (a good for which demand is relatively elastic, and supply inelastic). Economic analysis indicates that the true benefit of a subsidy will: (1) be the same regardless of whether the subsidy is granted to the buyers or sellers in a market, and (2) will depend on the elasticities of supply and demand.

The Cost of Government Subsidy Programs

Policy makers and citizens alike often complain that the cost of government subsidy programs almost invariably exceeds initial projections. One reason for this is the increase in the quantity of the good purchased resulting from the subsidy. Prior to the enactment of the textbook subsidy, 100 million textbooks were sold annually. With a subsidy of $20 per textbook, one might be inclined to think that the annual cost of the program will be $2 billion ($20 × 100 million). This figure, however, will underestimate the true cost. Once the subsidy is in place, textbook sales will increase to 110 million, driving the overall cost of the program up to $2.2 billion ($20 × 110 million).

Furthermore, the expenditures on the subsidies will understate their total costs. To finance the subsidies, the government will have to raise the funds through taxation. A subsidy granted in one market will require taxation in other markets. As we have previously discussed, the taxes will generate a deadweight loss over and above the revenues transferred to the government. This excess burden is also a cost of the subsidy payments.

Real-World Subsidy Programs

The United States has a vast array of subsidy programs. Spending on these programs and the taxes that finance them are major items in the government budget. Some subsidy programs, such as Medicare and food stamps, provide payments to buyers. Others, such as the subsidies to the arts, public broadcasting, and sports stadiums, are directed toward suppliers. As we discussed, however, the party granted the subsidy may not be the one who captures the larger share of the actual benefit from the subsidy.

Still other subsidy programs are combined with price controls. Many agriculture subsidies fall into this category. The government fixes the prices of products like wheat, corn, cotton, and tobacco above the market equilibrium. To maintain the above-equilibrium price, the government purchases any amount produced that cannot be sold at the artificially high price. The government also restricts the acreage farmers are permitted to plant for these crops. If it were not for the planting restrictions, huge surpluses of these products would develop.

Sometimes government subsidies are granted only to a select group, or subset, of buyers (or sellers). Consider the structure of health-care subsidies in the United States. The Medicare program subsidizes the health-care purchases of senior citizens, and the Medicaid program provides subsidies to low-income households. These subsidies increase the demand for health care and drive up the prices of medical service for all consumers, including those ineligible for either program. In cases where only some of the buyers in a market are subsidized, groups that are ineligible for the subsidies will generally be harmed because they will have to pay higher prices for the good or service than they would otherwise have to, even though they do not receive a subsidy.

Subsidy programs are often highly complex, and it is sometimes difficult to determine whom they really benefit. As we proceed, we will analyze several of these programs in more detail. The supply and demand model presented here will facilitate our analysis.

LOOKING AHEAD

This chapter focused on how government-mandated price controls, taxes, subsidies, and prohibitions affect market outcomes. The next two chapters will apply the basic tools of economics to the political process more generally. In the chapters that follow, we will consider when intervention by the government is likely to enhance the well-being of citizens, and when it is likely to make them worse off. We will also analyze how the political process works and explain why it is sometimes a source of economic inefficiency.

KEY POINTS

- Resource markets and product markets are closely linked. A change in one will generally result in changes in the other.
- Legally imposed price ceilings result in shortages, and legally imposed price floors will cause surpluses. Both also cause other harmful secondary effects. Rent controls, for example, will lead to shortages, less investment, poor maintenance, and deterioration in the quality of rental housing.
- The minimum wage is a price floor for low-skilled labor. It increases the earnings of some low-skilled workers but also reduces employment and leads to fewer training opportunities and nonwage job benefits for other low-skilled workers.
- Because black markets operate outside the legal system, they are often characterized by deception, fraud, and the use of violence as a means of enforcing contracts. A legal system that provides secure private-property rights and unbiased enforcement of contracts enhances the operation of markets.
- The division of the actual tax burden between buyers and sellers is determined by the relative elasticities of demand and supply rather than on whom the tax is legally imposed.
- In addition to the cost of the tax revenue transferred to the government, taxes will reduce the level of the activity taxed, eliminate some gains from trade, and thereby impose an excess burden, or deadweight loss.
- As tax rates increase, the size of the tax base will shrink. Initially, rates and revenues will be directly related—revenues will expand as rates increase. However, as higher and higher rates are imposed, eventually an inverse relation will develop—revenues will decline as rates are increased further. The Laffer curve illustrates this pattern.
- The division of the benefit from a subsidy is determined by the relative elasticities of demand and supply rather than to whom the subsidy is actually paid.

CRITICAL ANALYSIS QUESTIONS

*1. How will a substantial increase in demand for housing affect the wages and employment of carpenters, plumbers, and electricians?

2. Suppose that college students in your town persuaded the town council to enact a law setting the maximum price for rental housing at $200 per month. Will this help or hurt college students who rent housing? In your answer, address how this price ceiling will affect (a) the quality of rental housing, (b) the amount of rental housing available, (c) the incentive of landlords to maintain their properties, (d) the amount of racial, gender, and other types of discrimination in the local rental housing market, (e) the ease with which students will be able to find housing, and finally, (f) whether a black market for housing would develop.

3. What is the difference between a price ceiling and a price floor? What will happen if a price ceiling is imposed below the market equilibrium? If a price ceiling for a good is set below the market equilibrium, what will happen to the quality and future availability of the good? Explain.

*4. To be meaningful, a price ceiling must be below the market price. Conversely, a meaningful price floor must be above the market price. What impact will a meaningful price ceiling have on the quantity exchanged? What impact will a meaningful price floor have on the quantity exchanged? Explain.

5. Congress recently passed a new program that will subsidize the purchase of prescription drugs by the elderly. What impact will this program have on the demand for and price of prescription drugs? How will people who are not elderly be affected by this program? Explain.

*6. Analyze the impact of an increase in the minimum wage from the current level to $10 per hour. How would the following be affected?
   a. Employment of people previously earning less than $10 per hour
   b. The unemployment rate of teenagers
   c. The availability of on-the-job training for low-skilled workers
   d. The demand for high-skilled workers who are good substitutes for low-skilled workers

7. What is a black market? What are some of the main differences in how black markets operate relative to legal markets?

8. How do you think the markets for organ donation and child adoption would be affected if they were made fully legal with a well-functioning price mechanism? What would be the advantages and disadvantages relative to the current system?

9. What is meant by the incidence of a tax? Explain why the statutory and actual incidence of a tax can be different.

10. What conditions must be met for buyers to bear the full burden of a tax? What conditions would cause sellers to bear the full burden? Explain.

*11. What is the nature of the deadweight loss accompanying taxes? Why is it often referred to as an "excess burden"?

12. The demand and supply curves for unskilled labor in a market are given in the accompanying table.
   a. Find the equilibrium wage and number of workers hired.
   b. Suppose that a new law is passed requiring employers to pay an unemployment insurance tax of $1.50 per hour for every employee. What happens to the equilibrium wage rate and number of workers hired? How is this tax burden distributed between employers and workers?
   c. Now suppose that, rather than being paid by employers, the tax must be paid by workers. How does this affect the equilibrium wage rate and number of workers hired? How is this tax burden distributed between employers and workers?
   d. Does it make a difference who is statutorily liable for the tax?

Demand		Supply	
Wage	Quantity Demanded	Wage	Quantity Supplied
$6.50	1,000	$6.50	1,900
$6.00	1,200	$6.00	1,800
$5.50	1,400	$5.50	1,700
$5.00	1,600	$5.00	1,600
$4.50	1,800	$4.50	1,500
$4.00	2,000	$4.00	1,400

13. Currently, the Social Security payroll tax is legally imposed equally on workers and employers: 7.65 percent for employees and 7.65 percent for employers. Show this graphically, being careful to distinguish between the total cost to the employer of hiring a worker, the employee's gross wage, and the employee's net wage. Show how the outcome would differ if all 15.3 percent were imposed on the employee or if all 15.3 percent were imposed on the employer.

*14. Suppose Congress passes legislation requiring that businesses employing workers with three or more children pay these employees at least $10 per hour. How would this legislation affect the employment level of low-skilled workers with three or more children? Do you think some workers with large families might attempt to conceal the fact? Why?

15. "We should impose a 20 percent luxury tax on expensive automobiles (those with a sales price of $50,000 or more) in order to collect more tax revenue from the wealthy." Will the burden of the proposed tax fall primarily on the wealthy? Why or why not?

*16. Should policy makers seek to set the tax on an economic activity at a rate that will maximize the revenue derived from the tax? Why or why not? Explain.

"17. During the summer of 2001, the combination of city and state taxes on cigarettes sold in New York City rose from $1.19 to more than $3.00 per pack. How will this tax increase affect (a) the quantity of cigarettes sold in New York City, (b) the revenue derived by the city and state from the tax, (c) the Internet purchases of cigarettes by New Yorkers, and (d) the incidence of smoking by New Yorkers?

*Asterisk denotes questions for which answers are given in Appendix B.

Difficult Cases for the Market, and the Role of Government

CHAPTER 5

Chapter Focus

- What is economic efficiency and how can it be used to evaluate markets?
- Why is it generally undesirable to pursue any goal to perfection?
- What is the role of government in a market economy?
- What are externalities? What are public goods?
- Why might markets fail to allocate goods and services efficiently?
- If the market has shortcomings, does this mean the government should intervene?

The principal justification for public policy intervention lies in the frequent and numerous shortcomings of market outcomes.

—*Charles Wolf, Jr.*[1]

[1]Charles Wolf, Jr., *Markets or Government* (Cambridge, Mass.: MIT Press, 1988), 17.

As we previously discussed, markets and government planning are the two main alternatives for the organization of economic activity. Chapters 3 and 4 introduced you to how markets work and demonstrated how the invisible hand of the market process directs the self-interest of individuals toward activities in the best interest of society. Throughout, we noted that some qualifications were in order, in terms of both the "rules of the game" that must be in place for markets to work well and the existence of special cases, in which the invisible hand might not function effectively. In this chapter, we turn our attention to discussing these potential problem areas for the market and consider their implications with regard to the role of government. In the following chapter, we will analyze how the political process works more directly.

A CLOSER LOOK AT ECONOMIC EFFICIENCY

Economic efficiency
A situation that occurs when (1) all activities generating more benefit than cost are undertaken, and (2) no activities are undertaken for which the cost exceeds the benefit.

Economists use the standard of **economic efficiency** to assess the desirability of economic outcomes. We briefly introduced the concept in Chapter 3. We now want to explore it in more detail. The central idea of economic efficiency is straightforward. For any given level of cost, we want to obtain the largest possible benefit. Alternatively, we want to obtain any particular benefit for the least possible cost. Economic efficiency means getting the most value from the available resources—making the largest pie from the available set of ingredients, so to speak.

Economists acknowledge that individuals generally do not regard the efficiency of the entire economy as a primary goal for themselves. Rather, each person is interested in enlarging the size of his or her own slice. But if resources are used more efficiently, the overall size of the pie will be larger, and therefore, at least potentially, *everyone* could have a larger slice. For an outcome to be consistent with ideal economic efficiency, two conditions are necessary:

Rule 1. ***Undertaking an economic action is efficient if it produces more benefits than costs.*** To satisfy economic efficiency, all actions generating more benefits than costs must be undertaken. Failure to undertake all such actions implies that a potential gain has been forgone.

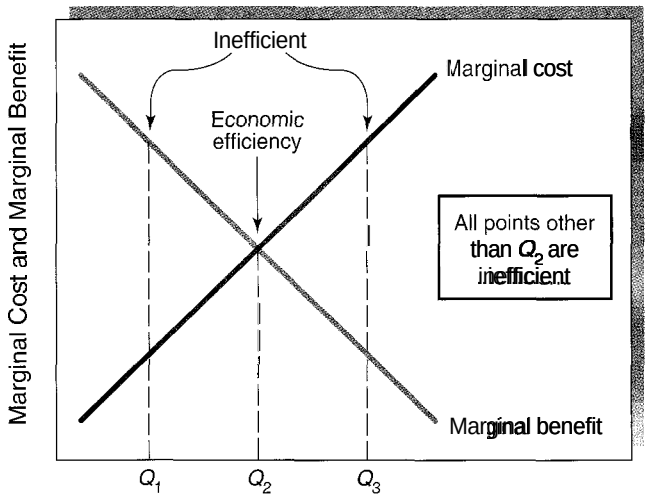
Rule 2. ***Undertaking an economic action is inefficient if it produces more costs than benefits.*** To satisfy economic efficiency, no action that generates more costs than benefits should be undertaken. When such counterproductive actions are taken, society is worse off because even better alternatives were forgone.

Economic efficiency results only when both of these conditions have been met. ***Either failure to undertake an efficient action (Rule 1) or the undertaking of an inefficient action (Rule 2) will result in economic inefficiency.*** To illustrate, consider **Exhibit 1,** which shows the benefits and costs associated with expanding the amount of any particular activity. We have avoided using a specific example here to ensure you understand the general idea of efficiency without linking it to a specific application. As we will show, the concept has wide-ranging applications — from the evaluation of government policy to how long you choose to brush your teeth in the morning.[2]

In Exhibit 1, the marginal benefit curve shows the additional benefit associated with expanding the activity. The marginal cost curve shows the cost — including any opportu-

[2]Note to students who may pursue advanced study in economics: Using the concept of efficiency to compare alternative policies typically requires that the analyst estimate costs and benefits that are difficult or impossible to measure. Costs and benefits are the values of opportunities forgone or accepted by individuals, ***as*** evaluated ***by those*** individuals. Then these costs and benefits must be added up across all individuals and compared. But does a dollar's gain for one individual really compensate for a dollar's sacrifice by another? Some economists simply reject the validity of making such comparisons. They say that neither the estimates by the economic analyst of subjectively determined costs and benefits nor the adding up of these costs and benefits across individuals is meaningful. Their case may be valid, but most economists today nevertheless use the concept of efficiency as we present it. No other way to use economic analysis to compare policy alternatives has been found.

EXHIBIT 1
Economic Efficiency



As we use more time and resources to expand the level of an activity, the marginal benefits will generally decline and the marginal costs rise. From the viewpoint of efficiency, the activity should be expanded as long as the marginal benefits exceed the marginal cost. Thus, quantity Q_2 is the economically efficient level of this activity. Q_1 is inefficient because some production that could generate more benefits than costs is not undertaken. Q_3 is also inefficient because some units are produced even though their costs exceed the benefits they create. Thus, both too much and too little of an activity will result in inefficiency.

nity costs—of spending additional time, effort, and resources on the activity. At Q_1, the height of the marginal benefit curve exceeds the height of the marginal cost curve. Thus, at that point, the additional benefits of expanding the activity past Q, exceed the additional costs. According to Rule 1 of economic efficiency, we should continue to expand the activity until we reach Q_2. Beyond Q_2 (at Q_3, for example), the height of the marginal benefit curve is less than the height of the marginal cost curve. The additional benefits from expanding activity to that point are smaller than the additional costs. According to Rule 2, at Q_3, we have gone too far and should cut back on the activity. Q_2 is the only point consistent with both rules of economic efficiency.

IF IT'S WORTH DOING, IT'S WORTH DOING IMPERFECTLY

Eliminating pollution. Earning straight As. Being completely organized Cleaning your apartment until it sparkles. Making automobiles completely safe. Making airplanes fully secure against terrorist attacks. All of these are worthwhile goals, right? Well, they are until you consider the costs of actually achieving them. The heading for this section is. of course, a play on the old saying, "If it's worth doing, it's worth doing to the best of your ability." Economics suggests, however, that this is not a sensible guideline At some point. the gains from doing something even better will not be worth the cost It will make more sense to stop short of perfection.

Exhibit 1 can also be used to illustrate this point. As more resources are dedicated to an activity, the marginal improvements (benefits) will become smaller and smaller, while the marginal costs will rise. The optimal time and effort put into the activity will be achieved at Q_2, and this will nearly always be well below one's best effort. Note that inefficiency results when either too little (for example, Q,) or too much (for example, Q,) time and effort are put into the activity.

Do you make decisions this way? Last time you cleaned your car or apartment, why did you decide to leave some things undone? Once the most important areas were clean, you likely began to skip over other areas (like on top of the refrigerator or under the bed), figuring that the benefits of cleaning these areas were simply not worth the cost. Very few

ECONOMICS AT THE MOVIES

Along Came Polly (2004)

Ben Stiller and Jennifer Aniston struggle to find the efficient (optimal) amount of organization in their lives. In one scene they use a knife to destroy pillows after Aniston convinces Stiller that the eight minutes a day he spends arranging decorative pillows on his bed (that nobody else sees) isn't worth the effort. In the next scene, Aniston's inefficiently low level of organization is illustrated when she spends a lot of time searching for her car keys. At the margin, Stiller's time spent arranging his pillows isn't generating enough benefit to justify the cost. Meanwhile, if Aniston were to spend just a little more time getting organized, the benefits to her would exceed the costs. The two of them would both be more efficient, and they'd probably get along better, too!

OBALCO LE TION/BENNETT TRACY

people live in a perfectly organized and clean house, wash their hands enough to prevent all colds, brush their teeth long enough to prevent all cavities, or make their home as safe as Fort Knox. They recognize that the benefit of perfection in these, and many other areas, is simply not worth the cost.

Economics is about trade-offs; it is possible to pursue even worthy activities beyond the level that is consistent with economic efficiency. People seem to be more aware of this in their personal decision making than when evaluating public policy. It is not uncommon to hear people say things like, "We ought to eliminate all pollution" or "No price is too high to save a life."

If we want to get the most out of our resources, we need to think about both marginal benefits and marginal costs and recognize that there are alternative ways of pursuing objectives. Consequently, economists do not ask whether eliminating pollution or saving lives is worth the cost *in terms of dollars* per se, but whether it is worth the cost in terms of giving up other things that could have been done with those dollars—the opportunity cost. Spending an extra $10 billion on worker safety requirements to save 100 lives isn't efficient if the funds could have been spent differently and saved 500 lives. Furthermore, it makes no more sense to have the government pursue perfection than it does for each of us personally to pursue it. Regardless of sector, achievement of perfection is virtually never worth the cost.

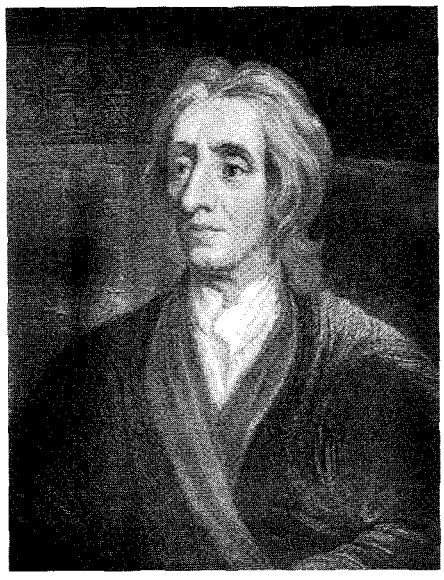
THINKING ABOUT THE ECONOMIC ROLE OF GOVERNMENT

For centuries, philosophers, economists, and other scholars have debated the proper role of government. While the debate continues, there is substantial agreement that at least two functions of government are legitimate: (1) protecting individuals and their property against invasions by others and (2) providing goods that cannot easily be provided through private markets. These two functions correspond to what Nobel laureate James Buchanan conceptualizes as the protective and productive functions of government.

Protective Function of Government

The most fundamental function of government is the protection of individuals and their property against acts of aggression. As John Locke wrote more than three centuries ago, individuals are constantly threatened by "the invasions of others." Therefore, each individual

"is willing to join in society with others, who are already united, or have a mind to unite, for the mutual preservation of their lives, liberties, and estates."[3] ***The protective function of government involves the maintenance of a framework of security and order—an infrastructure of rules within which people can interact peacefully with one another.*** Protection of person and property is crucial. It entails providing police protection and prosecuting aggressors who take things that do not belong to them. It also involves providing for a national defense designed to protect against foreign invasions. The legal enforcement of contracts and rules against fraud are also central elements of the protective function. People and businesses that write bad checks, violate contracts, or knowingly supply others with false information, for example, are therefore subject to legal prosecution.



© BETTMANN/CORBIS

The English philosopher John Locke argued that people own themselves and, as a result of this self-ownership, they also own the fruits of their labor. Locke stressed that individuals are not subservient to governments. On the contrary, the role of governments is to protect the "natural rights" of individuals to their person and property. This view, also reflected in the "unalienable rights" section of the U.S. Declaration of Independence, is the basis for the protective function of government.

It is easy to see the economic importance of the protective function. When it is performed well, the property of citizens is secure, freedom of exchange is present, and contracts are legally enforceable. When people are assured that they will be able to enjoy the benefits of their efforts, they will be more productive. In contrast, when property rights are insecure and contracts unenforceable, productive behavior is undermined. Plunder. fraud, and economic chaos result. Governments set and enforce the "rules of the game" that enable markets to operate smoothly.

Productive Function of Government

The nature of some goods makes them difficult to provide through markets. Sometimes it is difficult to establish a one-to-one link between the payment and receipt of a good. If this link cannot be established, the incentive of market producers to supply these goods is weak. In addition, high transaction costs—particularly, the cost of monitoring use and collecting fees—can sometimes make it difficult to supply a good through the market. When either of these conditions is present, it may be more efficient for the government to supply the good and impose taxes on its citizens to cover the cost.

One of the most important productive functions of government is providing a stable monetary and financial environment. If markets are going to work well, individuals have to know the value of what they are buying or selling. For market prices to convey this information, a stable monetary system is needed. This is especially true for the many market exchanges that involve a time dimension. Houses, cars, consumer durables, land, buildings, equipment, and many other items are often paid for over a period of months or even years. When the purchasing power of money fluctuates wildly, previously determined prices do not represent their intended values. Under these circumstances, exchanges involving long-term commitments are hampered, and the smooth operation of markets is undermined.

The government's tax, spending, and monetary policies exert a powerful influence on the stability of the overall economy. If properly conducted, these policies contribute to economic stability, full and efficient utilization of resources, and stable prices. However, improper stabilization policies can cause massive unemployment, rapidly rising prices, or—both. For those pursuing a course in macroeconomics, these issues will be central to that analysis.

[3] John Locke, *Treatise of Civil Government,* 1690, ed Charles Sherman (New York Appleton-Century-Croft?, 1937), 82

POTENTIAL SHORTCOMINGS OF THE MARKET

As we previously discussed, the invisible hand of market forces generally gives resource owners and business firms a strong incentive to use their resources efficiently and undertake projects that create value. Will this always be true? The answer to this question is "No." There are four maior factors that can undermine the invisible hand and reduce the efficiency of markets: (1) lack of competition, (2) externalities, (3) public goods, and (4) poorly informed buyers or sellers. We will now consider each of these factors and explain why they may justify government intervention.

Lack of Competition

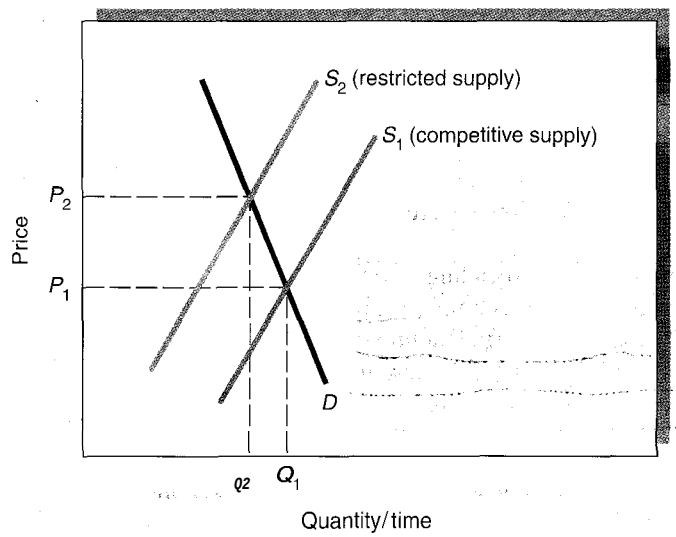
Competition is vital to the proper operation of the pricing mechanism. The existence of competing buyers and sellers reduces the power of both to rig or alter the market in their own favor. Although competition is beneficial from a social point of view, individually each of us would prefer to be loosened from its grip. Students do not like stiff competitors in their social or romantic lives, at exam time, or when they're trying to get into graduate school. Buyers on eBay hope for few competing bidders so they can purchase the items they're bidding on at lower prices. Similarly, sellers prefer fewer competing sellers so they can sell at higher prices.

Exhibit 2 illustrates how sellers can gain from restricting competition. In the absence of any restrictions on competition in the market, the price P_1 and output Q_1 associated with the competitive supply curve (S,) will prevail. Here, Q_1 is the level of output consistent with economic efficiency. If a group of sellers is able to restrict competition, perhaps by forcing some firms out of the market and preventing new firms from entering, the group would be able to gain by raising the price of the product. This is illustrated by the price P, and output Q_2 associated with the restricted supply (S_2). Even though the output is smaller, the total revenue (price P, times quantity Q,) derived by the sellers at the restricted output level is greater than at the competitive price P_1. Clearly, the sellers gain because, at the higher price, they are being paid more to produce less.

The restricted output level, however, is clearly less efficient. At the competitive output level Q_1, all units that were valued more than their cost are produced and sold. But this is not the case at Q_2. The additional units between Q_2 and Q_1 are valued more than their cost. Nonetheless, they will not be produced if suppliers are able to limit competition and restrict output. When competition is absent, there is a potential conflict between the interests of sellers and the efficient use of resources.

EXHIBIT 2
Lack of Competition and Problems for the Market

If a group *of* sellers can restrict competition, it may be able to gain by reducing supply (to S_2, for example) and raising the price (to P_2, for example) rather than charging the competitive market price of P_1. Under these circumstances, output will be less than the economically efficient level.



What can the government do to ensure that markets are competitive? The first guideline might be borrowed from the medical profession: Do no harm. ***A productive government will refrain from using its powers to impose licenses, discriminatory taxes, price controls, tariffs, quotas, and other entry and trade restraints that lessen the intensity of competition.*** In the vast majority of markets, sellers will find it difficult or impossible to limit the entry of rival firms (including rival producers from other countries). The only means by which they can limit competition is lobbying the government to impose restrictions or controls that limit competition on their behalf. In the interest of efficiency, governments should refrain from giving in to these demands.

When entering a market is very costly and there are only a few existing sellers, it may be possible for these sellers by themselves to restrict competition. In an effort to deal with cases like this, the United States has enacted a series of "antitrust laws," most notably the Sherman Antitrust Act (1890) and the Clayton Act (1914), making it illegal for firms to collude or attempt to monopolize a market.

For the most part, economists favor the general principle of government action to ensure and promote competitive markets. There is considerable debate, however, about the effectiveness of actual government policy in this area. Many economists believe that, by and large, government policy in this area has been ineffective. Others stress that government policies have often been misused to actually limit competition, rather than promote it. These laws have been used as a basis for restricting entry into markets, protecting existing producers from competitors, and limiting price competition. This is counterproductive. For those taking a microeconomics course, noncompetitive markets and related policy alternatives will be analyzed in greater detail later.

Externalities—A Failure to Account for All Costs and Benefits

When property rights are unclear or poorly enforced, the actions of an individual or group may "spill over" onto others and thereby affect their well-being without their consent. These spillover effects are called **externalities.** You are probably familiar with externalities. For example, when your neighbor's loud stereo makes it hard for you to study, you are experiencing an externality firsthand. Although your neighbors do not have a right to come in to your apartment and turn on your stereo, they do have a right to listen to their own stereo, and their listening may interfere with the quietness in your apartment. Their actions impose a cost on you, and they also raise an issue of property rights. Do your neighbors have a property right to play their stereo as loudly as they please? Or do you have a property right to quietness in your own apartment? When questions like these arise, how should the boundaries of property rights be determined, and what steps should be taken to ensure adequate enforcement? Although the volume of your neighbor's stereo may not be a major economic issue, it nonetheless illustrates the nature of the problems that arise when property rights are unclear and externalities are present.

Externalities
Spillover effects of an activity that influence the well-being of nonconsenting third parties.

The spillover effects may either impose a cost or create a benefit for third parties—people not directly involved in the transaction, activity, or exchange. Economists use the term **external cost** to describe a situation in which the spillover effects harm third parties. If the spillover effects enhance the welfare of the third parties, an **external benefit** is present. We will analyze both external costs and external benefits and consider why both of them can lead to problems.

External costs
Spillover effects that reduce the well-being of nonconsenting third parties.

External benefits
Spillover effects that generate benefits for nonconsenting third parties.

External Costs Economists worry about external costs because they may result in economic inefficiency. For example, resources may be used to produce goods that are valued less than their production costs, including the costs imposed on the nonconsenting third parties. Consider the production of paper. The firms in the market operate mills and purchase labor, trees, and other resources to produce the paper. But they also emit pollutants into the atmosphere that impose costs on residents living around the mills. The pollutants cause paint on buildings to deteriorate more rapidly. They make it difficult for some people to breathe normally, and perhaps cause other health hazards. If the residents living near a pulp mill can prove they have been harmed, they could take the mill to court and force the paper producer to cover the cost of their damages. But it might

be difficult to prove that they were harmed and that the pulp mill is responsible for the damage. As you can see, the residents' property rights to clean air may be difficult to enforce, particularly if there are many parties emitting pollutants into the air.

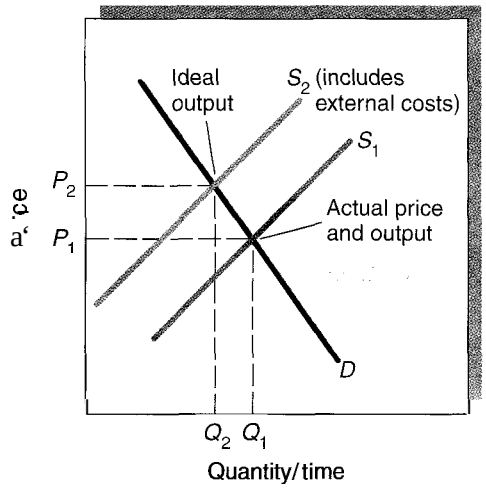
If the residents are unable to enforce their property rights, the production of paper will generate an external cost that will be ignored by markets. **Exhibit 3** illustrates the implications of these external costs within the supply and demand framework. As the result of the external cost, the market supply curve S_1 will understate the true cost of producing paper. It reflects only the cost actually paid by the firms, and ignores the uncompensated costs imposed on the nearby residents. Under these circumstances, the firm will expand output to Q, (the intersection of the demand curve D and supply curve S_1) and the market price P_1 will emerge. Is this price and output consistent with economic efficiency? The answer is clearly "No." If all of the costs of producing the paper, including those imposed on third parties, were taken into account, the supply curve S_2 would result. From an efficiency standpoint, only the smaller quantity Q_2 should be produced. The units beyond Q, on out to Q, cost more than their value to consumers. People would be better off if the resources used to produce those units (beyond Q,) were used to produce other things. Nonetheless, profit-maximizing firms will expand output into this range. Thus, when external costs are present, the market supply curve will understate production costs, and output will be expanded beyond the quantity consistent with economic efficiency. Moreover, resources for which property rights are poorly enforced will be overutilized and sometimes polluted. This is often the case with air and water when the property rights to these resources are poorly enforced.

What should be done about external costs? These costs arise because property rights are poorly defined or imperfectly enforced. Initially, therefore, it makes sense to think seriously about how property rights might be better defined and enforced. However, the nature of some goods will make the defining and enforcement of property rights extremely difficult. This will certainly be the case for resources like clean air and many fish species in the ocean. In cases that involve a relatively small number of people, the parties involved may be able to agree to rules and establish procedures that will minimize the external effects. For example, property owners around a small lake will generally be able to control access to the lake and prevent each other, as well as outsiders, from polluting or overfishing the lake.

However, in cases that involve large numbers of people, the transaction costs of arriving at an agreement will be prohibitively high, so it is unrealistic to expect that private contracts among the parties will handle the situation satisfactorily. For example, this will be the case when a large number of automobiles and firms emit pollutants into the atmosphere. In these "large number" cases, government regulations may be the best approach. At this

EXHIBIT 3
External Costs and Output That Is Greater Than the Efficient Level

When an activity such as paper production imposes external costs on nonconsenting third parties, these costs will not be registered by the market supply curve (S_1). As a result, output will be beyond the economically efficient level. The units between Q_2 and Q_1 will be produced, even though their cost exceeds the value they provide to consumers.



GETTY IMAGES

External costs resulting from poorly defined and enforced property rights underlie the problems of excessive air and water pollution.

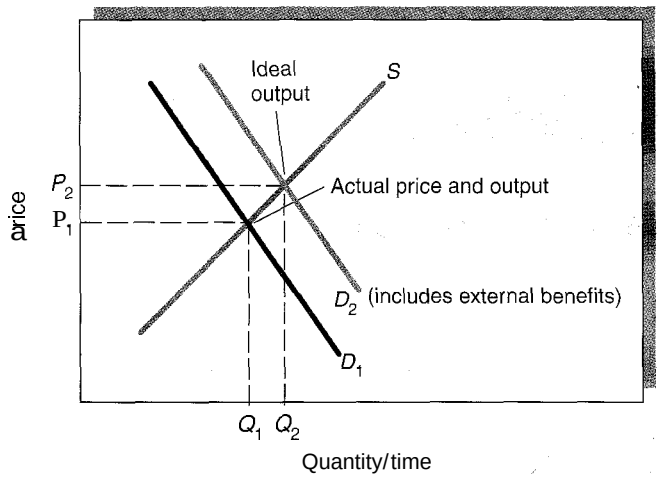
point, we want you to see the nature of the problem when external costs are present. As we proceed, we will analyze a number of problems in this area in detail and consider alternative approaches that might improve economic efficiency.

External Benefits As we mentioned, sometimes the actions of individuals and firms generate external benefits for others. The homeowner who keeps a house in good condition and maintains a neat lawn improves the beauty of the entire community. A flood-control dam built by upstream residents for their benefit might also generate gains for those who live downstream. Scientific theories benefit their authors, but the knowledge can also help others who did not contribute to the development of them.

From the standpoint of efficiency, why might external benefits be a problem? Here, inefficiency may arise because potential producers are unable to capture fully the benefits that their actions create for others. Suppose a pharmaceutical company develops a vaccine protecting users against a contagious virus or some other communal disease. Of course, the vaccine can easily be marketed to users who will benefit directly from it. However, because of the communal nature of the virus, as more and more people take the vaccine, nonusers will also be less likely to get the flu. But it will be very difficult for the pharmaceutical companies to capture any of the benefits derived by the nonusers. As a result, too little of the vaccine may be supplied.

Exhibit 4 illustrates the impact of external benefits like those generated by the vaccine within the framework of supply and demand. The market demand curve reflects the benefits derived by the users of the vaccine, while the supply curve reflects the opportunity cost of providing it. Market forces result in an equilibrium price of P_1 and output of Q_1. Is this outcome consistent with economic efficiency? Again, the answer is "No." The market demand curve D_1 will register only the benefits derived by the users. Those benefits that accrue to nonusers, who are now less likely to contract the flu, will not be taken into account by decision makers. The producer of the vaccine makes it more likely that these people will not get sick, but it doesn't derive any benefit (sales revenue) from having done so. Thus, market demand D_1 understates the total benefits derived from the production and use of the vaccine. Demand D_2 provides a measure of these total benefits. including those that accrue to the nonusers. The units between Q_1 and Q, are valued more highly than what it costs to produce them. Nonetheless, they will not be supplied because the suppliers of the vaccine will be unable to capture the benefits that accrue to the nonusers. Thus, when external benefits are present, market forces may supply less than the amount consistent with economic efficiency.

EXHIBIT 4
External Benefits and Output That Is Less Than the Efficient Level



A vaccine that protects users against the flu will also help nonusers by making it less likely that they will catch it. But this benefit will not be registered by the market demand curve (D_1). In cases where external benefits like this are present, output will be less than the economically efficient level. Even though the units between Q_1 and Q_2 generate more benefits than costs, they will not be supplied because sellers are unable to capture the value of these external benefits.

While external benefits are a potential source of inefficiency, entrepreneurs have a strong incentive to figure out ways to capture more fully the gains their actions generate for others. In some cases, they are able to capture what would otherwise be external benefits by extending the scope of the firm. The accompanying Application in Economics, "Capturing External Benefits: The Case of Walt Disney World," provides an interesting and informative illustration of this point.

Public Goods and Why They Pose a Problem for the Market

Public goods
Goods for which rivalry among consumers is absent and exclusion of nonpaying customers is difficult.

What are public goods? **Public goods** have two distinguishing characteristics: (1) nonrivalry in consumption and (2) nonexcludability. Let's take a closer look at both of these characteristics.

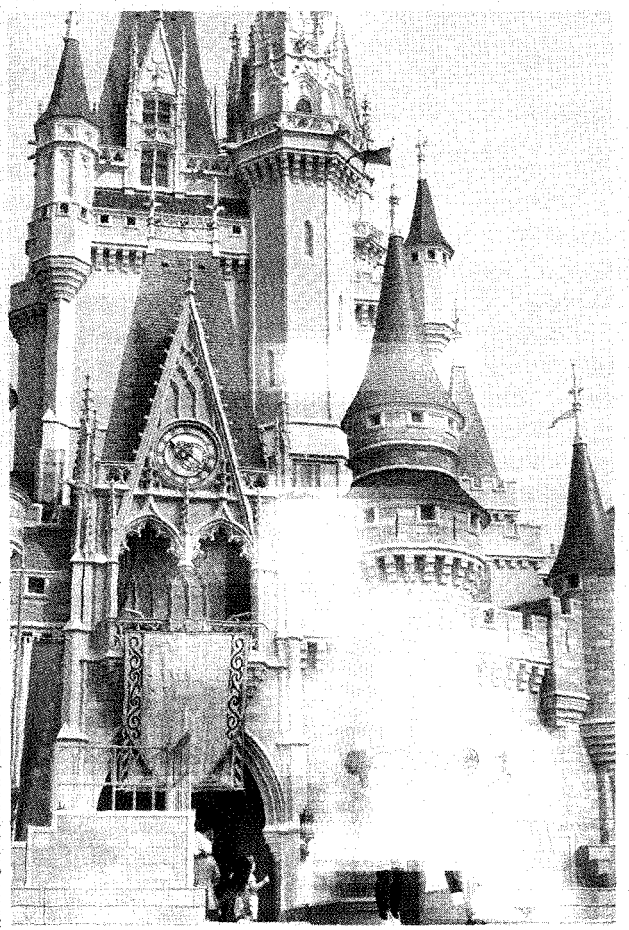
Nonrivalry in consumption means that making the good available to one consumer does not reduce its availability to others. In fact, providing it to one person simultaneously makes it available to other consumers. Thus, a consumer has no reason to compete with others for the good. A radio broadcast signal provides an example. The same signal can be shared by everyone within the listening range. Having additional listeners tune in does not detract from the availability of the signal. Clearly, most goods do not have this shared consumption characteristic, but are instead rivals-in-consumption. For example, two individuals cannot simultaneously consume the same pair of jeans. Further, if one person purchases a pair of jeans, there is one less pair available for someone else.

The second characteristic of a public good—nonexcludability—means that it is impossible (or at least very costly) to exclude nonpaying customers from receiving the good. Suppose an antimissile system were being built around the city in which you live. How could some people in the city be protected by the system and others excluded? Most people will realize there is no way the system can protect their neighbors from incoming missiles without providing similar protection to other residents. Thus, the services of the antimissile system have the nonexcludability characteristic.

It is important to note that it is the characteristic of the good, not the sector in which it is produced, that determines whether it qualifies as a public good. There is a tendency to think that if a good is provided by the government, then it is a public good. This is not the case. Many of the goods provided by governments clearly do not have the characteristics

APPLICATIONS IN ECONOMICS

Capturing External Benefits: The Case of Walt Disney World



MATT STROSHANE / BLOOMBERG NEWS / LANDOV

Sometimes projects that generate more benefits than cost are still unattractive because a substantial share of the benefits is external and therefore difficult to capture. If an entrepreneur could figure out a way to capture more of these benefits, an otherwise unprofitable project might be transformed into a profitable one. Sometimes this can be done by extending the scope of a project.

The development of golf courses is an example. Because of the beauty and openness of the courses, many people find it attractive to live nearby. Thus, constructing a golf course typically generates an external benefit—an increase in the value of the nearby property. In recent years, golf course developers have figured out how to capture this benefit. Now, they typically purchase a large tract of land around the planned course before *it is built.* This places them in a position to resell the land at a higher price after the golf course has been completed and the surrounding land has increased in value. By extending the scope of their activities to include real estate as well as golf course development, they are able to capture what would otherwise be external benefits.

Florida's Walt Disney World is an interesting case study in entrepreneurial ingenuity designed to capture external benefits more fully. When Walt Disney developed Disneyland in California, the market value of the land in the immediate area soared as a result of the increase in demand for services (food, lodging, gasoline, and so on). Because the land in the area was owned by others, the developers of Disneyland were unable to capture these external benefits. In addition, Disney felt as if some of the adult nightclubs that had opened around his existing Disneyland park were imposing external costs on him by detracting from the family image his park was trying to attain.

Because of his experience with these externalities, when Walt Disney World was developed outside of Orlando, Florida, in the mid-1960s, Walt Disney purchased far more land than was needed for the amusement park, This enabled him to capture the increased land value surrounding his development (when he resold the land for a higher price), and reduce the negative externalities imposed on him via his control of the surrounding property.

The purchases were made as secretly as possible to prevent speculators from driving up the land prices if Disney's actions were detected. Disney even created a handful of smaller companies, with names like the Latin-American Development and Managers Corporation and the Reedy Creek Ranch Corporation, to purchase the land. After his first major land purchase of 12,400 acres, Walt Disney was at a meeting at which he was offered an opportunity to purchase an additional 8,500 acres. Walt Disney's assistant was rumored to have said, "But Walt, we already own 12,000 acres, enough *to* build the park." Disney replied, "How would you like to own 8,000 acres around our existing Disneyland facility right now?" His assistant immediately responded, "Buy it!"

After another major acquisition of 1,250 acres, Disney began concentrating on buying smaller land parcels around his main property. By June 1965, Disney had purchased 27,400 acres, or about 43 square miles—an area 150 times larger than his existing Disneyland park, and about twice as big as Manhattan. In October 1965, when an Orlando newspaper finally broke the story that Disney was behind the land purchases, the remaining land prices around his property jumped from $183 an acre to $1,000 an acre overnight. But by then, except for several small parcels he was unable to acquire, Walt Disney had purchased all of the land he wanted.

Florida eventually gave Walt Disney permission to create an autonomous Reedy Creek Improvement District, outside the authorrty of any local government in Florida. In a very real sense, Walt Disney World is a jurisdiction of its own, separate from any other local government authority. Because of this, Walt Disney World can write its own zoning

(continued)

APPLICATIONS IN ECONOMICS

(continued)

restrictions and building codes. It can also plan its own roadways, lakes, security, sidewalks, airports, and recreational areas. Walt Disney World is able to provide goods and services like these—that might normally be considered public goods—by charging general admission fees to its park. This helped Disney overcome the potential free-rider problems sometimes associated with producing these goods.

Just as Disney expected, the value of the land surrounding Walt Disney World soared as the demand for hotels, restaurants, and other businesses increased along with the development of the the amusement park. Through the years, the resale of land near the park has been a major source of revenue for the company. To a large degree, the success of the Disney Corporation reflects Walt Disney's entrepreneurial ability to deal with externality and public-good problems.

of public goods. Medical services, education, mail delivery, trash collection, and electricity come *to* mind. Although these goods are often supplied by governments, they do not have either nonrivalry or nonexcludability characteristics. Thus, they are not public goods.

Why are public goods difficult for markets to allocate efficiently? The nonexcludability characteristic provides the answer. Since those who do not pay cannot be excluded, sellers are generally unable to establish a one-to-one link between the payment and receipt of these goods. Realizing they cannot be excluded, potential consumers have little incentive to pay for these goods. Instead, they have an incentive to become **free riders**, people who receive the benefits of the good without helping to pay for its cost. But, when a large number of people become free riders, not very much of the good is supplied. This is precisely the problem: markets will tend to undersupply public goods, even when the population in aggregate values them highly relative to their cost.

Free rider
A person who receives the benefit of a good without paying for it. Because of their nonexcludable nature, public goods are subject to free-rider oroblems.

Suppose national defense were provided entirely through the market. Would you voluntarily help to pay for it? Your contribution would have little impact on the total supply of defense available to each of us, even if you made a large personal contribution. Many citizens, even though they might value defense highly, would become free riders, and few funds would be available to finance national defense.

For most goods, it is easy to establish a link between payment and receipt. If you do not pay for a gallon of ice cream, an automobile, a television set, a DVD player, and literally thousands of other items, suppliers will not provide them to you. Thus, there are very few public goods. National defense is the classic example of a public good. Radio and TV signals, software programs, flood-control projects, mosquito abatement programs, and perhaps some scientific theories also have public good characteristics. But beyond this short list, it is difficult to think of additional goods that qualify.

Just because a good is a public good does not necessarily mean that markets will fail to supply it. When the benefit of producing these goods is high, entrepreneurs will attempt to find innovative ways to gain by overcoming the free-rider problem. For example, radio and television broadcasts, which have both of the public good characteristics, are still produced well by the private sector. The free-rider problem is overcome through the use of advertising (which generates indirect revenue from listeners), rather than directly charging listeners. Private entrepreneurs have developed things like scrambling devices (so nonpaying customers can't tune into broadcasts free of charge), copy protection on DVDs, and tie-in purchases (for example, tying the purchase of a software instruction manual to the purchase of the software itself) to overcome the free-rider problem. The marketing of computer software provides an interesting illustration. Since the same software program can be copied without reducing the amount available, and it is costly to prevent consumption by nonpayers, software clearly has public good characteristics. Nonetheless, Bill Gates became the richest man in the world by producing and marketing it!

In spite of the innovative efforts of entrepreneurs, however, the quantity of public goods supplied strictly through market allocation might still be smaller than the quantity consistent with economic efficiency. This creates a potential opportunity for government action *to* improve the efficiency of resource allocation.

Potential Information Problems

Like other goods, information is scarce. Thus, when making purchasing decisions, people are sometimes poorly informed about the price, quality, durability, and side effects of alternative products. Imperfect knowledge is not the fault of the market. In fact, the market provides consumers with a strong incentive to acquire information. If they mistakenly purchase a ''lemon,''they will suffer the consequences. Furthermore, sellers have a strong incentive to inform consumers about the benefits of their products, especially in comparison to competing products. However, circumstances will influence the incentive structure confronted by both buyers and sellers.

The consumer's information problem is minimal if the item is purchased regularly. Consider the purchase of soap. There is little cost associated with trying different brands. Since soap is a regularly purchased product, trial and error is an economical means of determining which brand is most suitable to one's needs. Regularly purchased items such as toothpaste, most food products, lawn service, and gasoline provide additional examples of **repeat-purchase items.** When purchasing items like these, the consumer can use past experience to acquire accurate information and make wise decisions.

Repeat-purchase item
An item purchased often by the same buyer.

Furthermore, the sellers of repeat-purchase items also have a strong incentive to supply consumers with accurate information about them because failing to do so will adversely affect future sales. Because future demand is directly related to the satisfaction level of current customers, sellers of repeat-purchase items will want to help their customers make satisfying long-run choices. This helps harmonize the interests of buyers and sellers.

But harmony will not always occur. Conflicting interests, inadequate information, and unhappy customers can arise when goods are either (1) difficult to evaluate on inspection and seldom repeatedly purchased from the same producer. or (2) potentially capable of serious and lasting harmful side effects that cannot be predicted by a typical consumer. Under these conditions, consumers might make decisions they will later regret.

When customers are unable to distinguish between high-quality and low-quality goods, business entrepreneurs have an incentive to cut costs by reducing quality. Businesses that follow this course may survive and even prosper. Consider the information problem when an automobile is purchased. Are consumers capable of properly evaluating the safety equipment? Most are not. Of course, some consumers will seek the opinion of experts, but this information will be costly and difficult to evaluate. In this case, it might be more efficient to have the government regulate automobile safety and require certain safety equipment.

Similar issues arise with regard to product effectiveness. Suppose a new wonder drug promises to reduce the probability a person will be stricken by cancer or heart disease. Even if the product is totally ineffective, many consumers will waste their money trying it. Verifying the effectiveness of the drug will be a complicated and lengthy process. Consequently, it may be better to have experts certify its effectiveness. The federal Food and Drug Administration was established to perform this function. However, letting the experts decide is also a less than ideal solution. The certification process is likely to be costly and lengthy. As a result, the introduction of products that are effective may be delayed for years, and they are likely to be more costly than they would be otherwise.

Information as a Profit Opportunity

Consumers are willing to pay for information that will help them make better decisions. This presents a profit opportunity. ***Entrepreneurial publishers and other providers of information help consumers find what they seek by offering product evaluations by experts.*** For example, dozens of publications provide independent expert opinions about automobiles and computers at a low cost to potential purchasers. Laboratory test results and detailed product evaluations on a wide variety of goods are provided by *Consumer Reports, Consumer Research,* and other publications.

Franchise
A right or license granted to an individual to market a company's goods or services or use its brand name. The individual firms are independently owned but must meet certain conditions to continue to use the name.

Franchises are another way entrepreneurs have responded to the need of consumers for more and better information. A **franchise** is a right or license granted to an individual

to market a company's goods or services (or use their brand name). Fast-food restaurants like McDonald's, Wendy's, and Burger King are typically organized as franchises. The individual restaurants are independently owned, but the owner pays for the right to use the company name and must offer specific products and services in a manner specified by the franchiser. Franchises help give consumers reliable information. The tourist traveling through an area for the first time with very little time to search out alternatives may find that eating at a franchised restaurant and sleeping at a franchised motel are the cheapest ways to avoid annoying and costly mistakes that might come from patronizing an unknown local establishment. The franchiser sets the standards for all firms in the chain and establishes procedures, including continuous inspections designed to maintain the standards. Franchisers have a strong incentive to maintain their reputation for quality, because if it declines, their ability to sell new franchises and to collect ongoing franchise fees is adversely affected. Even though the tourist may visit a particular establishment only once, the franchise turns that visit into a "repeat purchase," since the reputation of the entire national franchise operation is at stake.

Similarly, advertising a brand name nationally puts the brand's reputation at stake each time a purchase is made. How much would the Coca-Cola Company pay to avoid the sale of a dangerous bottle of Coke? Surely, it would be a large sum. Interbrand, a branding consulting agency that evaluates and ranks the top brand names in the world, estimates that Coke's brand name is worth $67.4 billion. The value of that brand name is a hostage to quality control. The firm would suffer enormous damage if it failed to maintain the quality of its product. For example, in 2000 and 2001, Firestone's brand name suffered an immense reduction in value after only a few Firestone tires were suspected of being defective. Firestone is still attempting to recover from its loss in brand name value.

Enterprising entrepreneurs have found ways to assure buyers that products meet high standards of quality, even when the producer is small and not so well known. Consider the case of Best Western Motels.[4] Best Western owns no motels; however, building on the franchise idea, it publishes rules and standards with which motel owners must comply if they are to use the Best Western brand name and the reservation service that the company also operates. To protect its brand name, Best Western sends out inspectors to see that each Best Western Motel meets these standards. Every disappointed customer harms the reputation and reduces the value of the Best Western name, and reduces the willingness of motel owners to pay for use of the name. The standards are designed to keep customers satisfied. Even though each motel owner has only a relatively small operation, renting the Best Western name provides the small operator with the kind of international reputation formerly available only to large firms. In effect, Best Western acts as a regulator of all motels bearing its name. It profits by requiring efficient standards—those that produce maximum visitor satisfaction for every dollar spent by the motels utilizing the franchise name. As it does so, it helps eliminate problems in the market that result from imperfect information.

Underwriters Laboratories, Inc., is another example of private-sector regulation aimed at overcoming potential information problems. UL, as it is better known, is a private-sector corporation that has been testing and certifying products for more than 100 years based on its own set of quality standards. You have probably seen the UL mark on many of your household appliances. Sellers pay a fee to have UL evaluate their products for possible certification. The value of the UL brand depends on their careful evaluation of every product they certify. If UL allows defective products to carry its mark, its brand value will diminish.

Information published by reliable sources, franchising, and brand names can help consumers make better-informed decisions. Although these options are effective, they will not always provide an ideal solution. Government regulation may sometimes be able to improve the situation, but this, too, has its shortcomings. As with other things, there is no general solution to imperfect information problems.

[4]This section draws from Randall G Holcombe and Lora P. Holcombe, "The Market for Regulation," *Journal of Institutional and Theoretical Economics* 142, no. 4 (1986): 684–96.

COCA COLA BLOOMBERG NEWS/LANDOV

© GLINN WO F/PHOTOEDIT

© JEFF GREENBERG/THE IMAGE WORK

COPYRIGHT © 1995–2004 UNDERWRITERS LABORATORIES INC. ALL RIGHTS RESERVED

Brand names (like Coca-Cola). franchises (like McDonald's), consumer-ratings magazines (like *Consumer Reports*), and private-sector certification firms (like Underwriters Laboratories, Inc), are ways the private sector helps buyers overcome potential information problems.

PULLING THINGS TOGETHER

Throughout this textbook, we have stressed that a sound legal system—one that protects individuals and their property and provides access to evenhanded courts for the enforcement of contracts and settlement of disputes—is vitally important for the smooth operation of markets. So, too, is a monetary regime that provides people with access to a sound currency—money that maintains its value across time periods. Beyond these functions, however, there is little justification for government action when there is reason to expect that markets will allocate resources efficiently. But a lack of competition, externalities, public goods, and information problems often pose challenges and sometimes undermine the efficient operation of markets. Market shortcomings due to these factors raise the possibility that government intervention beyond the protective function might improve things. But before jumping to that conclusion, we need better knowledge about how the political process works. We are now ready to move on to that topic.

LOOKING AHEAD

Political decision making is complex, but the tools of economics can enhance our understanding of how it works. This is the subject matter of the next chapter:

KEY POINTS

- Economists use the standard of economic efficiency to assess the desirability of economic outcomes. Efficiency requires both: (1) that all actions generating more benefit than cost be undertaken, and (2) that no actions generating more cost than benefit be undertaken.
- Although perfection is a noble goal, it is rarely worth achieving because additional time and resources devoted to an activity generally yield smaller and smaller benefits and cost more and more. Inefficiency can result when either too little or too much effort is put into an activity.
- Governments can enhance economic well-being by performing both protective and productive functions. The protective function involves (1) the protection of individuals and their property against aggression and (2) the provision of a legal system for the enforcement of contracts and settlement of disputes. The productive function of government can help people obtain goods that would be difficult to supply through markets.
- When markets fail to meet the conditions for ideal economic efficiency, the problem can generally be traced to one of four sources: absence of competition, externalities, public goods, or poor information.
- Externalities reflect a lack of fully defined and enforced property rights. When external costs are present, output can be too large—units are produced even though their costs exceed the benefits they generate. In contrast, external benefits can lead to an output that is too small—some units are not produced even though the benefits of doing so would exceed the cost.
- Public goods are goods for which (1) rivalry in consumption is absent and (2) it is difficult to exclude those who do not pay. Because of the difficulties involved in establishing a one-to-one link between payment and receipt of such goods, the market supply of public goods will often be less than the economically efficient quantity.
- Entrepreneurs in markets have an incentive to find solutions to each market problem, and new solutions are constantly being discovered. But problems remain that can potentially be improved through government action.

CRITICAL ANALYSIS QUESTIONS

*1. Why is it important for producers to be able to prevent nonpaying customers from receiving a good?

2. In response to the terrorist attacks of September 11, 2001, airline security screening has increased dramatically. As a result, travelers must now spend considerably more time being screened before flights. Would it make economic sense to devote enough resources to completely prevent any such future attacks? Why or why not?

3. What are the distinguishing characteristics of "public goods"? Give two examples of a public good. Why are public goods difficult for markets to allocate efficiently?

*4. Which of the following are public goods? Explain, using the definition of a public good.
   a. an antimissile system surrounding Washington, D.C.
   b. a fire department
   c. tennis courts
   d. Yellowstone National Park
   e. elementary schools

5. Explain in your own words what is meant by external costs and external benefits. Why may market outcomes be less than ideal when externalities are present?

6. English philosopher John Locke argued that the protection of each individual's person and property (acquired without the use of violence, theft, or fraud) was the primary function of government. Why is this protection important to the efficient operation of an economy?

7. "If it's worth doing, it's worth doing to the best of your ability." What is the economic explanation for why this statement is frequently said but rarely followed in practice? Explain.

8. "Unless quality and price are regulated by government, travelers would have no chance for a fair deal. Local people would be treated well, but the traveler would have no way to know, for example, who offers a good night's lodging at a fair price." Is this true or false? Explain.

***9.** If sellers of toasters were able to organize themselves, reduce their output, and raise their prices, how would economic efficiency be affected? Explain.

10. What are external costs? When are they most likely to be present? When external costs are present, what is likely to be the relationship between the market output of a good and the output consistent with ideal economic efficiency?

***11.** "Elementary education is obviously a public good. After all, it is provided by the government." Evaluate this statement.

12. What are the necessary conditions for economic efficiency? In what four situations might a market fail to achieve ideal economic efficiency?

13. Suppose that Abel builds a factory next to Baker's farm, and air pollution from the factory harms Baker's crops. Is Baker's property right to the land being violated? Is an externality present? What if the pollution invades Baker's home and harms her health? Are her property rights violated? Is an externality present? Explain.

***14.** Apply the economic efficiency criterion *to* the role of government. When would a government intervention be considered economically efficient? When would a government intervention be considered economically inefficient?

***Asterisk** denotes questions for which answers are given in **Appendix** B

CHAPTER 6

The Economics of Collective Decision Making

[Public choice] analyzes the motives and activities of politicians, civil servants and government officials as people with personal interests that may or may not coincide with the interest of the general public they are supposed to serve. It is an analysis of how people behave in the world as it is.

—*Arthur Seldon*[1]

It does not follow that whenever laissez faire falls short government interference is expedient; since the inevitable drawbacks of the latter may, in any particular case, be worse than the shortcomings of private enterprise.

—*Harry Sidgwick, 1887*[2]

Chapter Focus

- **How large is the government sector, and what are the main activities undertaken by government?**
- **What are the differences and similarities between market and government actions?**
- **What insights can economics provide about the behavior of voters, politicians, and bureaucrats? How will their actions affect political outcomes?**
- **When is democratic representative government most likely to lead to economic efficiency?**
- **Why will there sometimes be a conflict between winning politics and economic efficiency?**
- **How does economic organization influence the efficiency of resource use?**

[1]Preface to Gordon Tullock, *The Vote Motive* (London: Institute of Economic Affairs, 1976), x.
[2]Quoted in Charles Wolf, Jr., *Markets or Government* (Cambridge, Mass.: MIT Press, 1988), 17.

As we have previously discussed, the protection of property rights, evenhanded enforcement of contracts, and provision of a stable monetary environment are vital for the smooth and efficient operation of markets. Governments that perform these functions well will help their citizens prosper and achieve higher levels of income. Governments may also help allocate goods difficult for markets to handle. However, it is crucially important to recognize that government is simply an alternative form of economic organization. In most industrialized nations, the activities of governments are directed by the democratic political process. In this chapter, we will use the tools of economics to analyze how this process works.

THE SIZE AND GROWTH OF THE U.S. GOVERNMENT

What exactly does government do? Has its role in the economy shrunk or grown over time? Data on government spending shed light on these questions. As **Exhibit 1** illustrates, total government expenditures (federal, state, and local combined) were only 9.4 percent of the U.S. economy in 1930. *(Note:* GDP is generally how economists measure the size of the economy. The term will be explained more fully in a macroeconomics course.) In that year, federal government spending by itself was only 3 percent of the economy. At the time, this made the federal government about half the size of all state and local governments combined.

However, between 1930 and 1980, the size of government grew very rapidly. By 1980, government expenditures had risen to 32.8 percent of the economy, *more than three times* the level of 1930. Moreover, the federal government grew to about twice the size of all state and local governments combined—despite the fact that they were growing rapidly, too. Over the last two decades, total government spending as a share of the economy has been relatively constant at approximately one-third of GDP.

Exhibit 2 shows the major categories of government spending for both the federal government and state and local governments. The major categories of federal spending are health care, national defense, Social Security, and other income transfers. Education, administration, and public welfare and health constitute the largest areas of spending for state and local governments.

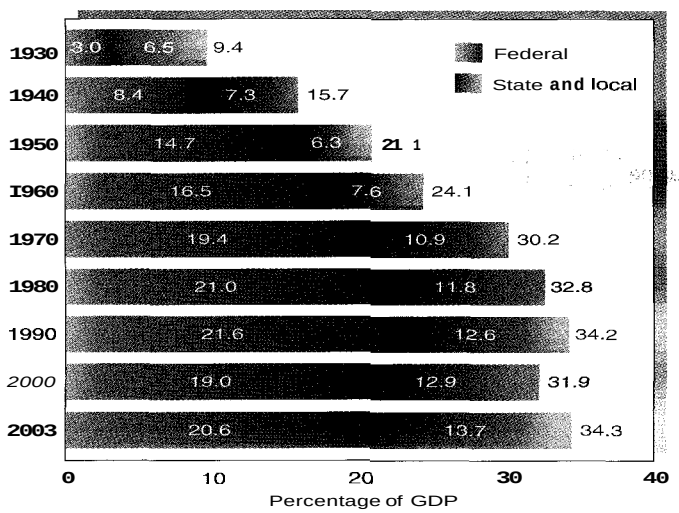


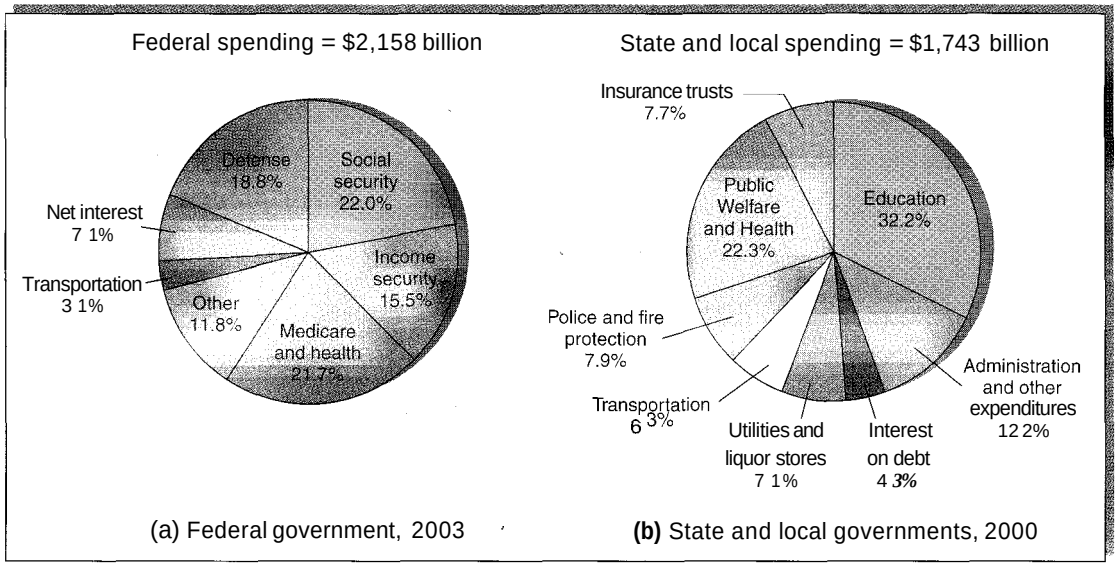
EXHIBIT 1
The Growth of Government Spending between 1930 and 2003

U.S. government expenditures as a share of the economy's gross domestic product have risen dramatically over the past seventy years.

Source: Bureau of Economic Analysis, http://www.bea.gov. Grants to state and local governments are included in federal expenditures. Individual data may not add to total due to rounding.

EXHIBIT 2
Government Spending by Category

The major categories of federal government spending are health care, Social Security, national defense, and income security (welfare programs). The major categories of state and local government spending are education, health care, and welfare programs.



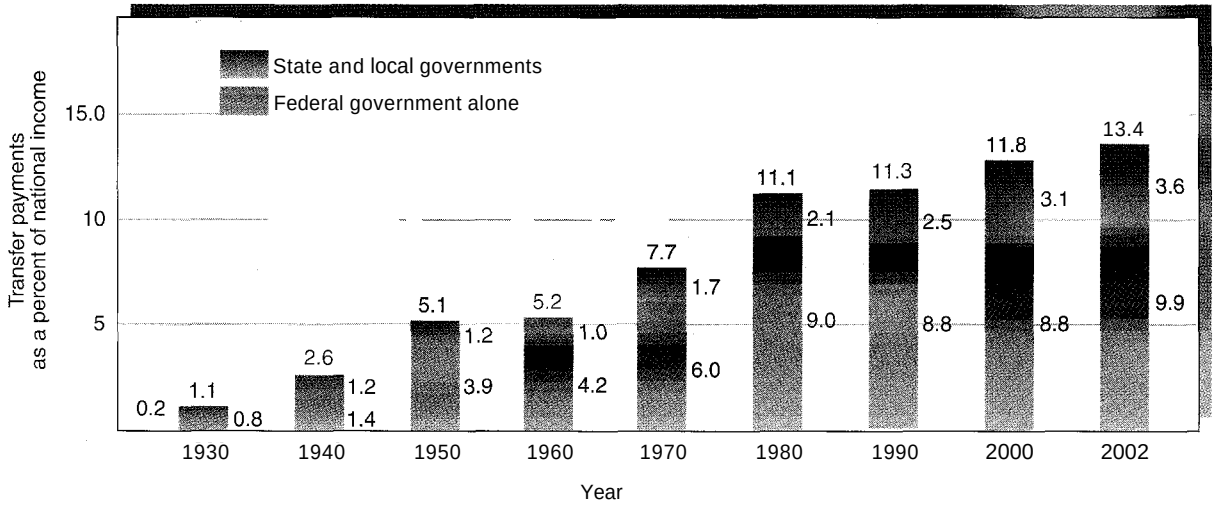
Source: Economic Report *of* the President, 2004, and Statistical Abstract *of* the United States, 2003.

Transfer payments Payments to individuals or institutions that are not linked to the current supply of a good or service by the recipient.

Transfer payments are transfers of income from some individuals (who pay taxes) to others (who receive government payments). Social Security, unemployment benefits, and welfare are examples of transfer payments. Direct income transfers now account for almost 40 percent of the total spending of the government. As **Exhibit 3** illustrates, government spending on income transfers has grown rapidly. In 1930, income transfers summed to only

EXHIBIT 3
The Growth of Government Transfer Payments

The government taxes approximately 13 percent of national income away from some people and transfers it to others. Means-tested income transfers — those directed toward the poor — account for only about one-sixth of all income transfers. Government income-transfer activities have grown substantially over the past seventy years.



Source: Bureau of Economic Analysis, http,//www.bea.gov.

1.1 percent of total income. By 1970, the figure had jumped to 7.7 percent, and by 2002 it had risen to 13.4 percent of national income. Obviously, the government has become much more involved in tax-transfer activities during the past seventy years.

Given the size and growth of government, analyzing how the political process works and what impact it is likely to have on the economy is a vitally important topic. The remainder of this chapter will address this issue.

THE DIFFERENCES AND SIMILARITIES BETWEEN GOVERNMENTS AND MARKETS

When political decisions are made democratically, the choices of individuals will influence outcomes in the government sector—just as they do in the market sector. Therefore, when we analyze the political process, we focus on individuals and how incentives influence their choices, just as we do when we analyze markets. There are both differences and similarities between political and market decision making. Let's take a look at six of them.

1. Competitive behavior is present in both the market and public sectors. The nature of the competition and the criteria for success differ between the two sectors, but people compete in both. Politicians compete for elective office. Bureau chiefs and agency heads compete for taxpayer dollars and the authority to regulate others to meet their bureau or agency goals. Public-sector employees compete for promotions, higher incomes, and additional power, just as they do in the private sector. Lobbyists compete for program funding, for favorable bureaucratic rulings, and for legislation favorable to the interest groups they represent—including both private and government clients. The nature of the competition may differ between the two sectors, but it is present in both. (See Applications in Economics: Perspectives on the Cost of Political Competition.)

2. Public-sector organization can break the individual consumption-payment link. In the market sector, goods are allocated to those who are willing to pay the price: there is

APPLICATIONS IN ECONOMICS

Perspectives on the Cost of Political Competition: What Does It Cost to Get Elected?

Competition for elective office is fierce and campaigns are expensive. For example, in recent years, candidates for U.S. House and Senate positions raised and spent more than $1 billion. This amounts to approximately $2 million per congressional seat! Highly contested seats are often far more expensive.

During and after an election, lobbying groups compete for the attention of elected officials. In fact, the greatest portion of campaign funds raised by incumbents is not raised at election time; rather, it accrues over their entire term in office. A large campaign contribution may not be able to "buy" a vote, but it certainly enhances the lobbyist's chance to sit down with the elected official to explain the power and "beauty" of the contributor's position. In the competitive world of politics, the politician who does not at least listen to helpful "friends of the campaign" is less likely to survive.

The U.S. Congress controls approximately $2 trillion in spending annually and imposes regulations that cost another $800 billion. That's a huge amount of money. As long as Congress wields the power to spend these sums, huge expenditures designed to influence the policies representatives make will continue.[1]

[1] More details on campaign finance can he found in Michael Barone dnd Grant Ujifusa *The Almanac of American Politic,* (Washington, D C National Journal, annual), or at the Federal Election Commission's Web site

a one-to-one relationship between a person's payment and receipt of a good. This is often not the case when decisions are made politically. Sometimes people receive very large benefits from the government even though they do not pay much of the cost to cover them. In other cases, individuals are required to pay dearly for a government program even though they derive few, if any, benefits.

3. Scarcity imposes the aggregate consumption-payment link in both sectors. Although the government can break the link between a person's payment for a good and the right to consume it, the reality of the *aggregate consumption–aggregate payment link* remains. Resources used by the government have alternative uses. Therefore, it is costly to provide goods and services through the government. This is true even if the good is provided "free of charge" to certain consumers.

4. Private-sector action is based on mutual agreement; public-sector action is based on majority rule. In the market sector, when two parties engage in trade, they do so voluntarily. Corporations like General Motors and Microsoft, no matter how large or powerful, cannot take income from you or force you to buy their products. On the other hand, when collective action occurs in a democratic setting, majority rule is the key, either through direct voting or through legislative procedures involving elected representatives. If a legislative majority decides on a particular policy, the minority must accept the policy and help pay for it, even if they strongly disagree. Similarly, if government regulators mandate that private parties must provide a wildlife habitat, wetlands, or housing at below-market prices, for example, both providers and potential buyers must comply. Although market action is based on mutual benefit, government action through the political process generates losers as well as winners.

5. When collective decisions are made legislatively, voters must choose among candidates who represent a bundle of positions on issues. On election day, the voter cannot choose the views of one politician on poverty and business welfare and simultaneously choose the views of a different politician on national defense and tariffs. This greatly limits the voter's power to make his or her preferences count on specific issues. Since the average representative is asked to vote on roughly 2,000 different issues during a two-year term, the size of the problem is obvious. The situation in markets, however, is quite different. A buyer can purchase some groceries or clothing from one store, while choosing related items from different suppliers. There is seldom a bundle-purchase problem in markets.

6. Income and power are distributed differently in the two sectors. People who supply more highly valued resources in the marketplace have larger incomes. The number of these dollar "votes" earned by a person in the marketplace will reflect his or her abilities, ambitions, skills, past savings, inheritance, good fortune, and willingness to produce for others, among other things. Bill Gates is a good example. Many people have "voted" for his products. Consequently, Gates has become quite wealthy. This process results in an unequal distribution of income and power in the market sector.

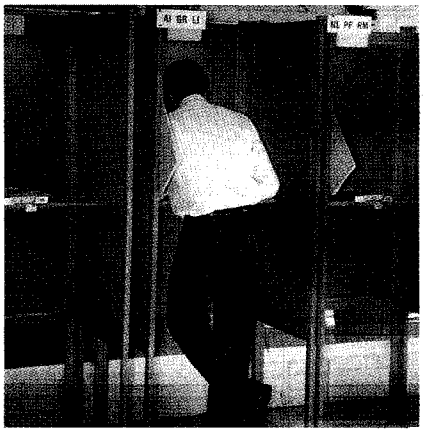
On the other hand, in a democratic government, one citizen, one vote is the rule. But there are ways other than voting to influence political outcomes. People can donate both their money and their time to help a campaign. They can also try to influence friends and neighbors, write letters to legislators, and speak in public on behalf of a candidate or cause. The greatest rewards of the political process go to those best able and most willing to use their time, persuasive skills, organizational abilities, and financial contributions to help politicians get votes. People who have more money and skills of this sort—and are willing to spend them in the political arena—can expect to benefit more handsomely for themselves and their favorite causes. Thus, while the sources of success and influence differ, there is an unequal distribution of influence and power in both sectors.

POLITICAL DECISION MAKING: AN OVERVIEW

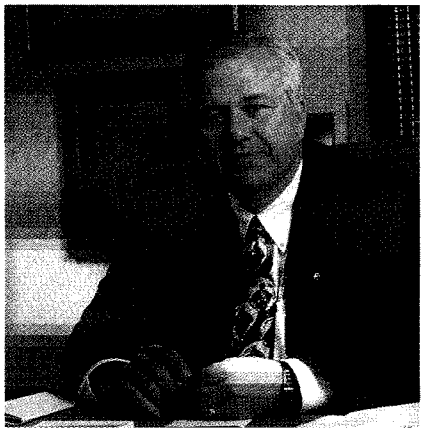
Public-choice analysis is a branch of economics that applies the principles and methodology of economics to the operation of the political process. Public-choice analysis links the theory of *individual* behavior to political action, analyzes the implications of the theory, and tests them against events in the real world. Over the past fifty years, research in this area has greatly enhanced our understanding of political decision-making.[1] Just as economists have used the idea of self-interest to analyze markets, public-choice economists use it to analyze political choices and the operation of government. After all, the same people make decisions in both sectors. If self-interest and the structure of incentives influence market choices, there is good reason to expect that they will also influence choices in a political setting.

Public-choice analysis
The study of decision making as it affects the formation and operation of collective organizations, such as governments. In general, the principles and methodology of economics are applied to political science topics.

The collective decision-making process can be thought of as a complex interaction among voters, legislators, and bureaucrats. Voters elect a legislature, which levies taxes and allocates budgets to various government agencies and bureaus. The bureaucrats in charge of these agencies utilize the funds to supply government services and income transfers. In a representative democracy, voter support determines who is elected to the legislature. A majority vote of the legislature is generally required for the passage of taxes, budget allocations, and regulatory activities. Let's take a closer look at the incentive structure confronting the three primary political players — voters, legislators, and bureaucrats — and consider how they affect the operation of the political process.



© BOB DAEMMRICH/THE IMAGE WORKS



AP PHOTO/J. SCOTT APPLEWHITE

REUTERS/JASON REED/LANDOV

Voters, politicians, and bureaucrats are the primary decision makers in the political arena.

[1]The contributions of Kenneth Arrow, Jamer Buchanan, Duncan Black, Anthony Downr, Mancur Olson, Robert Tollison and Gordon Tullock have been particularly important Public choice is something of a cross between economics and political science Thus, advanced courses are generally offered in both departments

Incentives Confronted by the Voter

How do voters decide whom to support? Self-interest dictates that voters, like market consumers, will ask, "What can you do for me and my goals, and how much will it cost me?" The greater the voter's perceived net personal gain from a particular candidate's election, the more likely it is that the voter will favor that candidate. In contrast, the greater the perceived net economic cost imposed on the voter by the positions of a candidate, the less inclined the voter will be to support the candidate. Other things being equal, voters will tend to support those candidates whom they believe will provide them the most government services and transfer benefits, net of personal costs.

How well will voters be informed about political issues and candidates? When decisions are made collectively, the choices of a single person will not be decisive. The probability that an individual vote will decide a city, state, or national election is virtually zero. Realizing that their votes will not affect the outcome, individual voters have little incentive to spend much effort seeking the information needed to cast an informed ballot. Economists refer to this lack of incentive as the **rational ignorance effect**.

Rational ignorance effect
Because it is highly unlikely that an individual vote will decide the outcome of an election, a rational individual has little or no incentive to search for and acquire the information needed to cast an informed vote.

As the result of the rational ignorance effect, most voters simply rely on information supplied to them freely by candidates (via political advertising) and the mass media, as well as conversations with friends and coworkers. Surveys, in fact, indicate that huge numbers of voters are unable to even identify their own congressional representatives, much less know where they stand on issues like Social Security reform, tariffs, and agricultural price supports. Given that voters gain little from casting a more informed vote, their meager knowledge of political candidates and issues is not surprising.

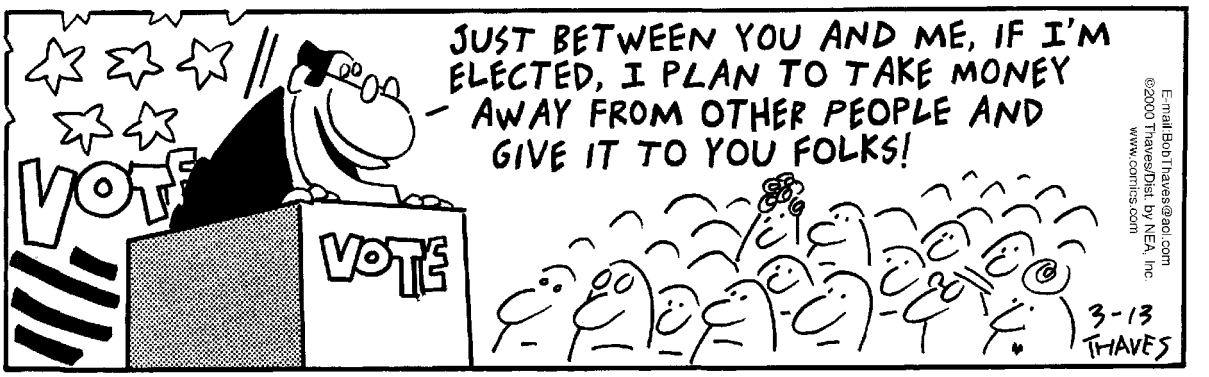
On the other hand, when people can put information to good use, they will put forth the effort to acquire it. Consider the incentive of an automobile purchaser to make a well-informed choice. The model, the dealer, and the financial terms are a matter of personal preference. If a bad choice is made, the individual consumer will bear the consequences. As a result, auto consumers have a strong incentive to make informed decisions. Thus, they often take different models for test drives, review consumer publications, and consult with various car experts about them. On the other hand, the voter gains little or nothing in terms of a changed result from a more informed political choice. Because the person is not in a position to decide the outcome of an election, if he or she makes a mistake by casting an uninformed ballot, it won't make much difference. Thus, it is actually *reasonable* to expect people to be far better informed when choosing a car than a senatorial, congressional, or other political candidate.

OUTSTANDING ECONOMIST

James Buchanan (1919–)

James Buchanan is a key figure in the development of public-choice theory. Buchanan's most famous work, *The Calculus of Consent* (1962), coauthored with Gordon Tullock, argues that unless constitutional rules are structured in a manner that will bring the self-interests of the political players into harmony with the wise use of resources, government action will often be counterproductive.[1] This and related contributions won him the 1986 Nobel Prize in economics. Buchanan is the founder of the Center for the Study of Public Choice and a longtime professor of economics at George Mason University.

[1] J. M Buchanan and G Tullock, *The Calculus of Consent* (Ann Arbor: University of Michigan Press, 1962).



FRANK & ERNEST REPRINTED BY PERMISSION OF NEWSPAPER ENTERPRISE ASSOCIATION, INC.

To get elected (or reelected), politicians have a strong incentive to provide transfers to important interest groups to secure their support.

The fact that citizens realize their individual votes will not sway the outcome of an election also explains why so many of them don't vote. Even in a presidential election, only about half of all voting-age Americans take the time to register and vote. The turnout for state and local elections is generally still lower. Given the low probability that one's vote will be decisive, low voter turn out is an expected result.

Incentives Confronted by the Politician

What motivates political candidates and officeholders? Economics indicates that the pursuit of votes will primarily shape politicians' actions and political positions. No doubt, many of them genuinely care about the "public interest" and the quality of government, but they need to get elected to achieve their objectives, whatever they might be. To be successful, a candidate's positive attributes must be brought to the attention of rationally ignorant voters focused on their families, jobs, various civic activities, and local sports teams (which are probably more entertaining). The successful candidate needs an expert staff, sophisticated polling techniques to uncover popular issues and positions, and high-quality advertising to favorably shape his or her image. This, of course, will be costly. It is not unusual for an incumbent candidate to the U.S. Senate to spend more than $15 million or more to get reelected. In other words, votes are the ultimate objective of politicians, but money helps them get those votes. Predictably, the pursuit of campaign contributions therefore shapes the actions of politicians, too.

Are we implying that politicians are selfish, caring only for their pocketbooks and reelection chances? The answer is "No." Factors other than personal political gain, narrowly defined, may well influence their actions. Sometimes an elected official may feel so strongly about an issue that he or she will knowingly take a position that is politically unpopular and damaging to his or her future electoral prospects. None of this is inconsistent with the economic view of the political process we just described. Over time, however, the politicians most likely to remain in office are the ones who focus on how their actions will influence their reelection prospects. Just as profits are the lifeblood of the market entrepreneur, votes are the lifeblood of the politician.

Politicians face competition for elected office from other candidates. Just like market suppliers, political suppliers have an incentive to find ways to gain an advantage over their competitors. Catering to the views of voters and contributors is one way of doing that. Enacting rules that put potential challengers at a disadvantage is another. When geographic political districts are redrawn, for example, politicians frequently manipulate the process to increase their chances of reelection—a process known as "gerrymandering." Incumbents can also attempt to use government resources for their reelection campaigns, an advantage challengers do not have. Campaign finance "reforms" that make it more difficult for a challenger to raise funds may also provide incumbents with an additional advantage.

Incentives Confronted by the Government Bureaucrat

Like other people, bureaucrats who staff government agencies have narrowly focused interests.[4] They usually want to see their own agency's goals furthered. Many bureaucrats believe strongly in what they are trying to do. Furthering these goals, however, usually requires larger budgets. In turn, larger budgets lead to more prestige and career opportunities for the bureaucrats. ***Economic analysis suggests there is a strong tendency for government bureaucrats and employees to want to expand their budgets to sizes well beyond what is economically efficient.***

Legislative bodies are in charge of overseeing these bureaus, but the individual legislators themselves are generally not very knowledgeable about the true costs of running these agencies. This makes it even more likely that bureaucrats will be able to get funding beyond what's economically efficient.

The political process, which begins with voter-driven elections and proceeds to legislative decisions and bureaucratic actions, brings about results that please some voters and displease others. The goals of the three major categories of participants — voters, politicians, and bureaucrats — frequently conflict with one another. Each group wants more of the government's limited supply of resources. Coalitions form and the members of each coalition hope to enhance their ability to get the government to do what they want. Sometimes this results in productive activities on the part of the government, and sometimes it does not.

WHEN THE POLITICAL PROCESS WORKS WELL

Under what conditions are voting and representative government most likely to result in productive actions? People have a tendency to believe that support by a majority makes a political action productive. However, if a government project is truly productive, it will always be possible to find a way to allocate the cost so that *all* voters gain. This would mean that, even if voting rules required unanimity or near-unanimity, all truly productive government projects would pass if the costs were allocated in the right manner. **Exhibit 4** helps illustrate this point. Column 1 presents hypothetical data on the distribution of benefits from a government road construction project. These benefits sum to $40, which exceeds the $25 cost of the road, so the project is productive. But if the project's $25 cost were allocated equally among the voters (plan A), Adams and Chan gain substantially, but Green, Lee, and Diaz lose. If the fate of the project is decided by majority vote, the project will be defeated by the "no" votes of Green, Lee, and Diaz. The reason why this productive government project fails to obtain a majority vote, however, is because of the way that the costs have been allocated.

Because the project is indeed productive, there is an alternative way to allocate its costs so that Adams, Chan, Green, Lee, and Diaz all benefit. This can be accomplished by

Just as the general does not want his Camp Swampy budget cut, most heads of agencies want expanded budgets to help them do more and do it more comfortably.

BEETLE BAILEY BY MORT WALKER REPRINTED BY SPECIAL PERMISSION OF KING FEATURES SYNDICATE

[4]The economic analysis of bureaucracy was pioneered by William Niskanen. Reprints of some of his classic articles along with recent updated material can be found in William A. Niskanen, Jr., ***Bureaucracy and Public Economics*** (Aldershot, U.K.: Edward Elgar Publishing, 1994).

EXHIBIT 4
The Benefits Derived by Voters from a Hypothetical Road Construction Project

Voter	Benefits Received (1)	Tax Payment: Plan A (2)	Tax Payment: Plan B (3)
Adams	$20	$5	$12.50
Chan	12	5	7.50
Green	4	5	2.50
Lee	2	5	1.25
Diaz	2	5	1.25
Total	**$40**	**$25**	**$25.00**

When taxes are levied in proportion to benefits received (tax plan B), any efficient project can pass unanimously (and any inefficient project will fail unanimously). When taxes are not levied in accordance with benefits received (tax plan A), efficient projects can fail to win a majority vote (or inefficient projects can pass in a majority vote).

allocating the cost of the project among voters in proportion to the benefits that they receive (plan B). Under this arrangement, Adams would pay half ($12.50) of the $25 cost. since he receives half ($20) of the total benefits ($40). The other voters would all pay in proportion to the benefits they receive. Under this plan, all voters would gain from the proposal. Even though the proposal could not secure a majority when the costs were allocated equally among voters, it will be favored by all five voters when they are taxed in proportion to the benefits they receive (plan B).

This simple illustration highlights an extremely important point about voting and the efficiency of government action. ***When voters pay in proportion to benefits received, all voters will gain if the government action is productive, and all will lose if it is unproductive.***[5] ***When the benefits and costs derived by individual voters are closely related, the voting process will enact efficient projects while rejecting inefficient ones. When voters pay in proportion to the benefits they receive, there will tend to be harmony between good politics and sound economics.***

How might the cost of a government service be linked to the benefits received? **User charges**, which require people who use a service more to pay a larger share of the cost. provide one way. User charges are most likely to be levied at the local level. Local services such as electricity, water, and garbage collection are generally financed with user charges. Sometimes the intensity of the use of a service and the amount paid for it can be linked by specifying that the revenue from a specific tax be used for a designated purpose. For example, most states finance road construction and maintenance with the revenue collected from taxes on gasoline and other motor fuels. The more an individual drives, the more he or she benefits from the roads—and the more he or she pays.

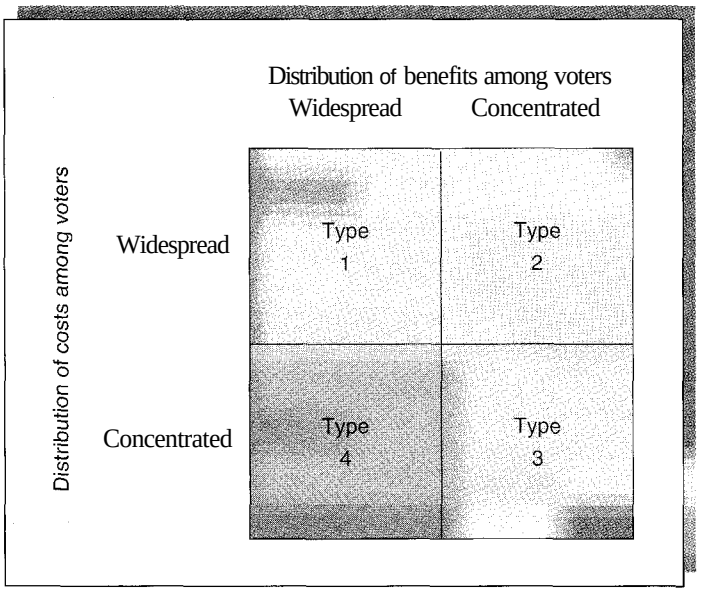
User charges Payments that users (consumers) are required to make if they want to receive certain services provided by the government.

Exhibit 5 provides a useful way to look at the possible linkage between the benefits and costs of government programs. The benefits from a government action may be either widespread among the general public or concentrated among a small subgroup (for example, farmers, students, business interests, senior citizens, or members of a labor union). Similarly, the costs may be either widespread or highly concentrated among voters. Thus, as the exhibit shows, there are four possible patterns of voter benefits and costs: (1) widespread benefits and widespread costs, (2) concentrated benefits and widespread costs. (3) concentrated benefits and concentrated costs, and (4) widespread benefits and concentrated costs.

When both the benefits and costs are widespread among voters (type 1 issue), essentially everyone benefits and everyone pays. Although the costs of type 1 measures may not be precisely proportional to the benefits individuals receive, there will be a rough

[5]The principle that productive projects generate the potential for political unanimity was initially articulated by Swedish economist Knut Wicksell in 1896. See Wicksell, "A New Principle of Just Taxation," in *Public Choice and Constitutional Economics*, James Gwartney and Richard Wagner (Greenwich, Conn.: JAI Press, Inc., 1988). Nobel laureate James Buchanan has stated that Wicksell's work provided him with the insights that led to his large role in the development of modem public-choice theory.

EXHIBIT 5
Distribution of Benefits and Costs among Voters



It is useful to visualize four possible combinations for the distribution of benefits and costs among voters to consider how the alternative distributions affect the operation of representative governments. When the distribution of benefits and costs is both widespread among voters (1) or both concentrated among voters (3), representative government will tend to undertake projects that are productive and reject those that are unproductive. In contrast, when the benefits are concentrated and the costs are widespread (2), representative government is biased toward the adoption of inefficient projects. Finally, when benefits are widespread but the costs concentrated (4), the political process may reject projects that are productive.

relationship. When type 1 measures are productive, almost everyone gains more than they pay. There will be little opposition, and political representatives have a strong incentive to support such proposals. In contrast, when type 1 proposals generate costs in excess of benefits, almost everyone loses, and representatives will face pressure to oppose such issues. Thus, for type 1 projects, the political process works pretty well. Productive projects will tend to be accepted and unproductive ones rejected.

Similarly, there is reason to believe that the political process will work fairly well for type 3 measures — those for which both benefits and costs are concentrated on one or more small subgroups. In some cases, the concentrated beneficiaries may be the same group of people paying for the government to provide them a service. In other cases, the subgroup of beneficiaries may differ from the subgroup footing the bill. Even in this case, however, when the benefits exceed the costs, the concentrated group of beneficiaries will have an incentive to expend more resources lobbying for the measure than those harmed by it will expend opposing it. Thus, when the benefits and costs are both concentrated, there will be a tendency for productive projects to be adopted and unproductive ones to be rejected.

WHEN THE POLITICAL PROCESS WORKS POORLY

Although the political process yields reasonable results when there is a close relationship between the receipt of benefits and the payment of costs (type 1 and type 3 projects), the harmony between good politics and sound economics breaks down when there is not (type 2 and type 4 projects). Inefficiency may also arise from other sources when governments undertake economic activities. In this section, we consider four major reasons why the political allocation of resources will often result in inefficiency.

Special-Interest Effect

Trade restrictions that limit the import of steel and lumber from abroad; subsidies for sports stadiums, the arts, and various agricultural products; federal spending on an indoor rain forest in Coralville, Iowa; a tattoo-removal program in San Luis Obispo County, California; the Rock and Roll Hall of Fame in Cleveland, Ohio; a golf awareness program in St. Augustine, Florida; and therapeutic horseback riding in Apple Valley, California. These seemingly diverse programs funded by the federal government have one thing in common: They reflect the attractiveness of special interests to vote-seeking politicians. A **special-interest** issue is one that generates substantial personal benefits for a small number of constituents while spreading the costs widely across the bulk of citizens (type 2 projects). Individually, a few people gain a great deal, but many others lose a small amount. In aggregate, the losses may exceed the benefits.

Special-interest issue
An issue that generates substantial individual benefits to a small minority while imposing a small individual cost on many other citizens. In total, the net cost to the majority might either exceed or fall short of the net benefits to the special-interest group.

How will a vote-seeking politician respond to special-interest issues? Since their personal stake is large, members of the interest group (and lobbyists representing their interests) will feel strongly about such issues. Many of the special-interest voters will vote for or against candidates strictly on the basis of whether they are supportive of their positions. In addition, interest groups are generally an attractive source of campaign resources—including financial contributions. In contrast, most other rationally ignorant voters will

APPLICATIONS IN ECONOMICS

Sweet Subsidies to Sugar Growers: A Case Study of the Special-Interest Effect

For many years, the price of sugar in the United States has been two or three times as high as the world price. For example, in February 2004, the domestic price of sugar was 20 cents per pound while the world price was less than 6 cents a pound. Why? Because the U.S. government severely restricts the quantity of sugar imported. This keeps the domestic price of sugar high. As a result, the roughly 60,000 sugar growers in the United States gain about $1.9 billion. That's more than $30,000 per grower! Most of these benefits are reaped by large growers with incomes far above the national average. On the other hand, these subsidies cost the average American household about $20 in the form of higher prices for products containing sugar. Even more important, the resources of Americans are wasted producing a good we are ill-suited to produce and one that could be obtained at a substantially lower cost through trade. As a result, Americans are worse off.

Why does Congress support this program year after year? Given the sizable impact the restrictions have on the personal wealth of sugar growers, it is perfectly sensible for them, particularly the large ones, to use their wealth and political clout to help politicians who support their interests. This is precisely what they have done. During the 2000 election cycle, the sugar lobby contributed almost $13 million to candidates and political action committees. In contrast, it makes no sense for the average voter to investigate this issue or give it any significant weight when deciding for whom to vote. In fact, most voters are unaware that this program is costing them money. Here, as in several other areas, politicians have a strong incentive to support policies favored by special interests, solicit those parties for political contributions, and use the funds to attract the support of other voters, most of whom know nothing about the sugar program. Even though the sugar program is counterproductive, it is still a political winner.

The sugar growers are not the only ones benefiting from government programs that are economically inefficient. Taxpayers and consumers spend approximately $20 billion annually to support grain, cotton, tobacco, peanut, wool, and dairy programs, all of which have structural characteristics similar to those of the sugar program. The political power of special interests also explains the presence of tariffs and quotas on steel, textiles, lumber, and many other products. Federally funded irrigation projects, subsidized agricultural grazing rights, subsidized business loans, numerous pork-barrel spending projects (the list goes on and on) are all policies rooted in the special-interest effect rather than economic efficiency and net benefits to Americans. Although each such program individually imposes only a small drag on the economy, together they exert a sizeable negative impact on our income levels and living standards.

either not know or will care little about special-interest issues. Even if voters know about some of these programs, it will be difficult for them to punish their legislators because each politician represents a bundle of positions on many different issues. While there is little to be gained from the support of the disorganized majority, organized interest groups provide politicians with vocal supporters, campaign workers, and, most important, financial contributions.

As a result, politicians have a strong incentive to support legislation giving concentrated benefits to special-interest groups at the expense of disorganized groups (like the bulk of taxpayers and consumers). Even if supporting such legislation is counterproductive, politicians will often still be able to gain by supporting programs favored by special interests. For a real-world illustration of how the special-interest effect works, see Applications in Economics, "Sweet Subsidies to Sugar Growers: A Case Study of the Special-Interest Effect."

Logrolling
The exchange between politicians of political support on one issue for political support on another.

Pork-barrel legislation
A package of spending projects benefiting local areas financed through the federal government. The costs of the projects typically exceed the benefits in total, but the projects are intensely desired by the residents of a particular district, who get the benefits without having to pay much of the costs.

The power of special interests is further strengthened by logrolling and pork-barrel legislation. **Logrolling** involves the practice of trading votes by a politician to get the necessary support to pass desired legislation. **Pork-barrel legislation** is the term used to describe the bundling of unrelated projects benefiting many interests into a single bill. Both logrolling and pork-barrel legislation will often make it possible for special-interest projects to gain legislative approval, even though these projects themselves are counterproductive and individually would be unable to muster the needed votes.

Exhibit 6 provides a numeric illustration of the forces underlying logrolling and pork-barrel legislation. Here we consider the operation of a five-member legislature considering three projects: construction of a post office in district A, dredging of a harbor in district B, and spending on a military base in district C. For each district, the net benefit or cost is shown—that is, the benefit to the district minus the tax cost imposed on it. The total cost of each of the three projects exceeds the benefits (as shown by the negative number in the total row at the bottom of the table), and therefore each is counterproductive. If the projects were voted on separately, each would lose by a 4-to-1 vote because only one district would gain, and the other four would lose. However, when the projects are bundled together through either logrolling (representatives A, B, and C could agree to trade votes) or pork-barrel legislation (all three programs put on the same bill), they can all pass, despite the fact that all are inefficient.[6] This can be seen by noting that the total combined net benefit is positive for representatives A, B, and C. Given the weak incentive for voters to acquire information,

EXHIBIT 6
Trading Votes and Passing Counterproductive Legislation

All three projects are inefficient, and would not pass majority vote individually. However, representatives from districts A, B, and C could trade votes (logrolling) or put together pork-barrel legislation that would result in all three projects passing.

	Net Benefits (+) or Costs (−) to Voters in District			
Voters of District[a]	Construction of Post Office in A	Dredging Harbor in B	Construction of Military Base in C	Total
A	+$10	−$ 3	−$ 3	**+$4**
B	−$ 3	+$10	−$ 3	**+$4**
C	−$ 3	−$ 3	+$10	**+$4**
D	−$ 3	−$ 3	−$ 3	**−$9**
E	−$ 3	−$ 3	−$ 3	**−$9**
Total	**−$ 2**	**−$ 2**	**−$ 2**	**−$6**

[a]*We assume the districts are of equal size.*

[6]*Logrolling and pork-barrel policies can sometimes lead to the adoption of productive measures. However, if a project is productive, there would always be a pattern of finance that would lead to its adoption even if logrolling and pork-barrel policies were absent. Thus, the tendency for logrolling and pork-barrel policies to result in the adoption of inefficient projects is the more significant point.*

those harmed by pork-barrel and other special-interest policies are unlikely to even be aware of them. Thus, the incentive to support projects like these is even stronger than is implied by the simple numeric example in Exhibit 6.

Why don't representatives oppose measures that force their constituents to pay for projects that benefit others? There is some incentive to do so, but the constituents of any one elected representative can capture only a small portion of the benefits of tax savings from improved efficiency, since the savings would be spread nationwide among all taxpayers. We would not, for example, expect the president of a corporation to devote any of the firm's resources to projects not primarily benefiting its stockholders. Neither should we expect an elected representative to devote political resources to projects like defeating pork-barrel programs when the benefits of spending reductions and tax savings will be derived mostly by constituents in other districts. Instead, each representative has a strong incentive to work for programs that concentrate benefits among his or her own constituents—especially organized interest groups that can help the representative be reelected. Heeding such incentives is a survival (reelection) tactic.

On the other hand, when the benefits of a governmental action are widespread and the costs are highly concentrated (type 4 of Exhibit 5), special-interest groups—those who stand to bear the cost—will strongly oppose and lobby against it. Most other voters will be largely uninformed and uninterested. Once again, politicians will have an incentive to respond to the views of the concentrated interests. A proposal to reduce or eliminate a tariff (tax) on an imported good would be an example of this type of legislation. Although many thousands of consumers would benefit from the lower prices that result, the domestic firms that compete with the imported good would devote substantial resources toward lobbying to keep the tariff in place. Projects of this type will tend to be rejected even when they are productive, that is, when they would generate larger benefits than costs.

The bottom line is clear: Public-choice analysis indicates that majority voting and representative democracy work poorly when concentrated interests benefit at the expense of the general public. In the case of special-interest issues, there is a conflict between good politics—getting elected—and the efficient use of resources. The special-interest effect helps explain the presence of numerous government programs that increase the size of government and reduce the overall size of the economic pie. As we discuss diverse topics throughout this text, counterproductive political action that has its foundation in the special-interest effect will arise again and again.

Shortsightedness Effect

Because voters have a weak incentive to acquire information, current economic conditions will have a major impact on their choices at election time. Complex issues, like reforming Social Security or restructuring health care programs that involve future benefits and costs, will be difficult for voters to assess. Thus, incumbent politicians will want to make sure economic conditions look good on election day. To accomplish this, they will favor policies that provide current benefits voters can easily identify at the expense of future costs that are complicated and difficult to identify. Similarly, they will tend to oppose legislation that involves immediate and easily identifiable costs (and higher taxes) but yield future benefits that are complex and difficult to identify. Economists refer to this bias inherent in the political process as the **shortsightedness effect**.

As a result of the shortsightedness effect, politicians will tend to favor programs that generate highly visible current benefits, even when the true cost of these programs outweighs the benefits. In contrast, their incentive is weak to support efficient programs that generate future benefits but involve current costs.

The shortsightedness effect sheds light on why legislators find debt financing so attractive. Debt financing makes it possible for officeholders to provide visible benefits to their constituents without having to levy an equivalent amount of taxes. During the last forty-five years, the federal budget has been in deficit forty times; there have been only five surpluses (1969 and 1998–2001). The bias toward budget deficits is a predictable result; it reflects the shortsighted nature of the political process. Similarly, the shortsighted-

Shortsightedness effect The misallocation of resources that results because public-sector action is biased (1) in favor of proposals yielding clearly defined current benefits in exchange for difficult-to-identify future costs and (2) against proposals with clearly identifiable current costs that yield less concrete and less obvious future benefits.

ness effect indicates that vote-seeking politicians will find it attractive to promise future benefits without levying a sufficient amount of taxes to finance them. This has been the case with both the Social Security and Medicare programs. The unfunded liabilities of these two programs are nearly *three times* the size of the official outstanding federal debt. By the time the higher taxes (or benefit cuts) for these programs are confronted, the politicians who gained votes from the promised benefits will be long gone.

It is worth taking a moment to consider the differences between the public and private sectors in terms of how future benefits and costs are considered in current decisions. As we explained in Chapter 2, private-property rights provide a means by which the value of future benefits can be immediately captured (or costs borne) by a property owner. Owners who do not invest now to properly maintain their homes or cars, for example, will bear the consequences of the reduced value of those assets. Correspondingly, the value of a firm's stock will immediately rise (or fall), depending on the shareholders' perception of the expected future benefits and costs of an action taken by the company's executives today. In contrast, the public sector tends to place more weight on current benefits and costs and less weight on the future. In areas where the primary benefits are in the future, and property rights can be well defined and enforced, there is good reason to believe that the private sector will do a better job than the government sector.

Rent Seeking

When buying and selling are controlled by legislation, the first things bought and sold are legislators.

—P. J. O'Rourke[7]

There are two ways individuals can acquire wealth: production and plunder. When individuals produce goods or services and exchange them for income, they not only enrich themselves but they also enhance the wealth of the society. Sometimes the rules—or lack of rule enforcement—also allow people to get ahead by taking, or plundering, what others have produced. This method not only fails to generate additional income—the gain of one is a loss to another—but it also consumes resources and thereby reduces the wealth of the society.

Rent seeking
Actions by individuals and groups designed to restructure public policy in a manner that will either directly or indirectly redistribute more income to themselves or the projects they promote.

Rent seeking is the term economists use when they refer to actions taken by individuals and groups seeking to use the political process to take the wealth of others.[8] Perhaps "favor seeking" would be a more descriptive term for this type of activity, which generally involves "investing" resources in lobbying and other activities designed to gain favors from the government. The incentive for individual5 to spend time and effort in rent seeking will be determined by how rewarding it is. Rent seeking will be unattractive when constitutional constraints prevent politicians from taking the property of some and transferring it to others (or forcing some to pay for things desired by others).

When a government fails to allocate the costs of public-sector projects to the primary beneficiaries (through user fees, for example), or when it becomes heavily involved in transfer activities, people will spend more time organizing and lobbying politicians and less time producing goods and services. Resources that would otherwise be used to create wealth and generate income are wasted as people fight over slices of the economic pie—a pie that is smaller than it could be if they were engaged in productive activities instead. When the government grants favors to some people at the expense of others (instead of simply acting as a neutral force protecting property rights and enforcing contracts), counterproductive activities will expand while productive activities will shrink. As a result, the overall income level will fall short of its potential.

There is ample evidence that rent-seeking consumes a substantial amount of resources. Washington, D.C., is full of organizations seeking subsidies and other favors from the federal government. More than 3,000 trade associations have offices in Washington, and they employ nearly 100,000people seeking to alter the actions of Congress. Of course, business and labor organizations are well represented, but so, too, are agricultural interests, health care providers, trial lawyers, senior citizens, export industries, and many others.

[7]Quoted in P. J. O'Rourke, *Insight Magazine*, Jan. *15–25: 35.*

[8]See the classic work of Charles K. Rowley, Robert D. Tollison. and Gordon Tullock, *The Political Economy of Rent-Seeking* (Boston: Kluwer Academic Publishers, 1988), for additional details on rent seeking.

As we noted earlier, income transfers have grown substantially during the last several decades. The government now taxes approximately one out of every seven dollars citizens earn, and transfers it to someone else. Rent seeking is the political "fuel" for most of these transfer activities. Interestingly, *means-tested transfers,* those directed toward the poor, constitute only about one-sixth of all transfers. No income test is applied to the other five-sixths of income transfers. These transfers are generally directed toward groups that are either well organized (like businesses and labor union interests) or easily identifiable (like the elderly and farmers). The people receiving these transfers often have incomes well above the average person.

Within the framework of public-choice analysis, the relatively small portion of income transfers directed toward the poor is not surprising. There is little reason to believe that transfers to the poor will be particularly attractive to vote-seeking politicians. After all, in the United States, the poor are less likely to vote than middle- and upper-income recipients. They are also less likely to be well informed on political issues and candidates. They are not an attractive source of political contributions. Politicians often argue that their proposed policies will help the poor, but there is little reason to believe that this will be a high priority for most of them.

There are three major reasons why government transfer activity will reduce the size of the economic pie. First, income redistribution weakens the link between productive activity and reward. When taxes take a larger share of a person's income, the reward from hard work and productive activity is reduced. Second, as public policy redistributes a larger share of income, more resources will flow into wasteful rent-seeking activities. Resources used for lobbying and other rent-seeking activities will not be available to increase the size of the economic pie. Third, higher taxes to finance income redistribution and an expansion in rent-seeking will induce taxpayers to focus less on income-producing activities, and more on actions to protect their income. More accountants, lawyers, and tax-shelter experts will be retained as people seek to limit the amount of their income redistributed to others. Like the resources allocated to rent seeking, resources allocated to protecting one's wealth from the reach of government will also be unavailable for productive activity. Predictably, the incentives created by government redistribution policies will exert a negative impact on the level of economic activity.

Inefficiency of Government Operations

Will government goods and services be produced efficiently? The pride of a job well done is likely to motivate both public- and private-sector suppliers. However, the incentive to reduce costs and operate efficiently differs substantially between the two. In the private sector, there is a strong incentive to produce efficiently because lower costs mean higher profits, and high costs mean losses and going out of business. This index of performance (profit) is unavailable in the public sector. Missing also are signals from the capital market. When a corporation announces a strategy or a plan that vigilant, personally committed investors believe to be faulty, the price of the corporation's stock will drop. There is no mechanism similar to the stock market in the public sector. Furthermore, direct competition in the form of other firms trying to woo the customers of a government agency or enterprise is largely absent in the public sector. As a result, bureaucrats have more freedom to pursue their narrow goals and interests without a strong regard for the control of costs relative to the benefits the public derives.

Bankruptcy weeds out inefficiency in the private sector, but there is no parallel mechanism to eliminate inefficiency in the public sector. In fact, failure to achieve a targeted objective (for example, a lower crime rate or improvement in student achievement scores) is often used as an argument for *increased* public-sector funding. Furthermore, public-sector managers are seldom in a position to gain personally from measures that reduce costs. The opposite is often true, in fact. If an agency fails to spend its entire budget for a given year, not only does it have to return the extra money, but its budget for the next year is likely to be cut. Because of this, government agencies typically go on a spending spree near the end of a budget period if they discover they have failed to spend all the current year's funds appropriated to them.

Just as the boy considers the quarter (his quarter) more important than the far greater cost (to the father) of the metal detector, so, too, does the leader of a bureau often consider the bureau's goals more important than the costs, even if the latter are far greater.

"Daddy, could you buy a new metal detector? I dropped my quarter in the snow."

FAMILY CIRCUS© REPRINTED BY PERMISSION OF KING FEATURES SYNDICATE

It is important to note that the argument of internal inefficiency is not based on the assumption that employees of a bureaucratic government are lazy or less capable. Rather, the emphasis is on the incentives and opportunities that government managers and workers confront. Government firms do not have owners that have risked their wealth on the future success of the firm. There is no entity that will be able to reap substantial economic gain if the firm produces more efficiently or incorporates a new product or service highly valued relative to its costs. The operation of the firm and the appointment of high-level managers might be influenced by political rather than economic considerations. Because the profitability criteria are absent, performance is difficult to evaluate. There are no tests to define economic inefficiency or measure it accurately—much less eliminate it. These perverse incentives are bound to affect efficiency.

The empirical evidence is consistent with this view. Economies dominated by government control, like those of the former Soviet bloc, India, Syria, and Nigeria (and many other African countries), have performed poorly. The level of output per unit of resource input in countries with numerous government enterprises is low. Similarly, when private firms are compared with government agencies providing the same goods or services (like garbage collection, hospitals, electric and water utilities, weather forecasting, and public transportation), studies indicate that private firms generally provide the services more economically.

ECONOMIC ORGANIZATION: WHO PRODUCES, WHO PAYS, AND WHY IT MATTERS

The structure of production and consumption will influence economic outcomes. Goods and services can be either produced by private enterprises or supplied by the government. They can be paid for either by the consumer directly or by the taxpayer or some other third party. As **Exhibit 7** shows, there are four possible combinations of production and consumption. Let's take a closer look at each and consider its impact on the allocation of resources and the incentive to economize.

In quadrant 1, goods are produced by private firms and purchased by consumers with their own money. Clearly, consumers will have a strong incentive to economize in this case. They will compare value with cost, and will make purchases only when they value items more than their purchase price. Correspondingly, the owners of private enterprises have a strong incentive to both cater to the views of consumers and supply goods efficiently. Net revenues can be increased if the output can be produced at a lower cost. Producers will continue supplying goods only if consumers are willing to pay an amount sufficient to cover their production costs. Essentially, the supply and demand analysis of Chapter 3 focused on quadrant 1 cases.

Quadrant 2 represents the case in which goods are produced privately but are paid for by the taxpayer or some other third party. Providing health care to citizens financed primarily by government (Medicare and Medicaid) or insurance is an example. If someone

Good is produced by	Good is paid for by: Consumer-Purchaser	Good is paid for by: Taxpayer or other Third Party
Private Enterprises	(1) Examples: apples, oranges, television sets, food, housing, most other goods	(2) Examples: health-care, food purchased with food stamps
Government Enterprises or Contracting	(3) Examples: Post Office, water and electricity in many cities, toll roads, many hospitals	(4) Examples: public schools, streets and roads, national defense, law enforcement

EXHIBIT 7
The Private- and Government-Sector Matrix of Production and Payment

The incentive to economize is influenced by who produces a good and who pays for it. Economizing behavior will be strongest when consumers purchase goods produced by private firms (quadrant 1). The incentive to economize is reduced when payment is made by a third party and when production is handled by the government.

else is paying the bill, consumers have little incentive to care much about the price of their health-care services. Instead of economizing, many consumers will simply purchase from suppliers they believe offer the highest quality, regardless of the price. The behavior of producers will also be affected. If consumers are largely insensitive to prices, producers have little reason to control costs and offer services at attractive prices. This can dramatically affect economic efficiency.

Quadrant 3 represents the situation in which consumers pay for a good or service, but production is handled by the government. First-class mail delivery via the U.S. Postal Service, water and electricity by municipal governments, and the operation of toll roads are examples that fall into this category. When consumers pay for a good or service directly, they will economize and seek the most value per dollar they spend. This will be true whether their purchases are from private or government enterprises. As we just discussed, however, there is reason to believe that government-operated firms will generally be less efficient than private enterprises. Cost consciousness is also likely to be reduced if the government firm is a monopolist—if it is protected from competition with potential private rivals. Competition, however, is difficult to maintain in some markets. When this is the case, government enterprises may offer a reasonable alternative. As we proceed, we will investigate this issue in more detail.

Quadrant 4 represents the case in which the government both provides the service and covers its costs through taxation. In this case, the political process determines what will be produced, how it will be produced, and how it will be allocated among the general public. Under these circumstances, consumers are in a very weak position to either discipline the suppliers or alter their production. The incentive to produce efficiently is weak, and there is likely to be a disconnect between the goods produced and the preferences of consumers. As we discussed in the previous chapter, the nature of public goods—items such as national defense—makes it difficult, if not impossible, to supply them through markets. In these cases, there may be little alternative to having the government provide them. In other instances, however, there are feasible alternatives. This is true for education.

Most goods and services in the United States are allocated under conditions approximating those of quadrant 1. Thus, most of our analysis focuses on this case. However, a sizable portion of economic activity takes place under conditions present in quadrants 2, 3, and 4, where the incentive structure often creates problems. As a result, our analysis also considers modifications that might improve the efficiency of activities currently undertaken in these quadrants.

THE ECONOMIC WAY OF THINKING ABOUT GOVERNMENT

Given its monopoly power over the legitimate use of force, people have a tendency to believe that the government, particularly a democratic representative government, can solve all types of problems. Further, if things do not go well, people tend to think that it is because the "wrong" people won the last election. Public-choice analysis suggests that the problem is more fundamental: there is sometimes a conflict between winning elections and following sound policies. For some types of activities, there is reason to believe that the political action that will help get one elected will, at the same time, reduce income levels and living standards.

Both the market and the political process have shortcomings. In Chapter 5, we focused on the shortcomings of the market and explained why markets sometimes result in the inefficient use of resources. This chapter provides a parallel analysis for the political process. The accompanying Thumbnail Sketch lists the major deficiencies of both sectors.

Understanding the strengths and weaknesses of both sectors is important if we are going to improve our current economic institutions. As we have stressed throughout this textbook, when the government protects property rights, enforces contracts, and provides a stable monetary environment, economic prosperity is more likely to ensue. The basic problem, however, is how a society can obtain the benefits of the protective functions of government and at the same time constrain it to those activities where it is a productive force. As the analysis of this chapter illustrates, this is not an easy task.

Could Constitutional Changes Help Promote Prosperity?

When we think about how to get the most out of our government, it is important to distinguish between ordinary politics and constitutional rules. Constitutions establish the procedures utilized to make political decisions. Constitutions can also limit the activities of government.

The framers of the U.S. Constitution were aware that even a democratic government might undertake counterproductive actions. Thus, they incorporated restraints on the economic role of government. They enumerated the permissible tax and spending powers of the central government (Article I, Section 8) and allocated all other powers to the states and the people (Tenth Amendment). They also prohibited states from adopting legislation "impairing the obligation of contracts" (Article I, Section 10). Furthermore, the Fifth Amendment specifies that private property shall not be "taken for public use without just compensation." Over time, however, these restraints have been significantly eroded, due in part to Supreme Court decisions that have effectively reinterpreted the Constitution. Today, it is difficult to think of an economic activity that is beyond the reach of majority rule or normal legislative procedure.

Public-choice analysis highlights the importance of constitutional rules and procedures capable of restraining government activities to those areas in which it will promote

THUMBNAIL SKETCH

What Weakens the Case for Market-Sector Allocation Versus Public-Sector Intervention, and Vice Versa?

These factors weaken the case for market-sector allocation:

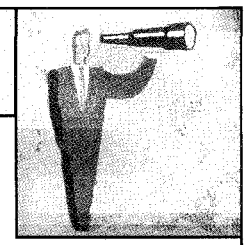
1. Lack of competition
2. Externalities
3. Public goods
4. Poor information

These factors weaken the case for public-sector intervention:

1. The special-interest effect
2. The shortsightedness effect
3. Rent seeking
4. Weak incentives for operational efficiency

prosperity. If left alone, even democratic governments will tend to cater to special-interest groups and draw significant resources into rent seeking. If we can figure out how to constrain the activities of government to those areas in which it is most likely to be productive, higher income levels can be achieved. The challenge before us is to develop constitutional rules and political institutions more consistent with economic efficiency and prosperity. The theory of public choice and its applications can help us do that. Needless to say, this topic is one of the most exciting and potentially fruitful areas of research in economics.

LOOKING AHEAD



Cases involving potential government intervention will be discussed repeatedly throughout this book. The tools presented in this chapter and the previous one will help us better understand both the potential and limitations of public policy as a source for economic progress.

KEY POINTS

- In recent years, government spending has been about one-third the size of the U.S. economy.
- There are both similarities and differences between markets and governments. Competition is present in both sectors. The government can use its taxing power to break the link between payment and receipt of a good for an individual, but not for the economy as a whole. In the public sector, voters face a "bundle" purchase problem; they are unable to vote for some policies favored by one candidate and other policies favored by the candidate's opponent. Power and income are distributed differently in the public sector than in the private sector.
- In a representative democracy, government is controlled by voters who elect politicians to set policy and hire bureaucrats to run government agencies. The incentives faced by all three classes of participants influence political outcomes.
- Voters have a strong incentive to support the candidate who offers them the greatest gain relative to their personal costs. Because collective decisions break the link between the choice of the individual and the outcome of the issue, voters are likely to be poorly informed on political matters.
- Politicians have a strong incentive to follow a strategy that will enhance their chances of getting elected (and reelected). Political competition more or less forces them to focus on how their actions influence their support among voters and potential contributors.
- The distribution of the benefits and costs among voters influences how the political process works. When voters pay in proportion to the benefits they receive from a public-sector project, productive projects tend to be approved and counterproductive ones rejected. When the costs of a policy are distributed among voters differently than are the benefits, democratic decision making will tend to be less efficient.
- Government actions will often lead to economic inefficiency as the result of (1) the special-interest effect, (2) the shortsightedness effect, (3) rent seeking, and (4) weak incentives to keep cost low within government enterprises and agencies. Thus, just as the market sometimes fails to allocate goods efficiently, so, too, will the government.
- Economic organization influences efficiency. The incentive to economize is strong when consumers use their own money to purchase goods and services from private firms. Both the payment by a third party and production by the government weaken the incentive to economize.

CRITICAL ANALYSIS QUESTIONS

1. Are voters likely to be well informed on issues and the positions of candidates? Why or why not?

*2. "The government can afford to take a long view when it needs to, while a private firm has a short-term outlook. Corporate officers, for example, typically care about the next 3 to 6 months, not the next 50 to 100 years. Government, not private firms, should own things like forests, which take decades to develop." Evaluate this view.

3. "If there are problems with markets, government will generally be able to intervene and correct the situation." Is this statement true or false? Explain your response.

"4. "The political process sometimes leads to economic inefficiency because we elect the wrong people to political office. If the right people were elected, a democracy governed by majority rule would allocate resources efficiently." Evaluate this statement.

5. What is rent seeking? When is it likely to be widespread? How does it influence economic efficiency? Explain.

*6. "The average person is more likely to make an informed choice when he or she purchases a personal computer than when he or she votes for a congressional candidate." Evaluate this statement.

7. "Government action is based on majority rule, whereas market action is based on mutual consent. The market allows for proportional representation of minorities, but minorities must yield to the views of the majority when activities are undertaken through government." In your own words, explain the meaning of this statement. Is the statement true? Why or why not?

"8. "Voters should simply ignore political candidates who play ball with special-interest groups and vote instead for candidates who will represent all the people when they are elected. Government will work far better when this happens." Evaluate this view.

9. If a project is efficient (if its total benefits exceed its total costs), would it be possible to allocate the cost of the project in a manner that would provide net benefits to each voter? Why or why not? Explain. Will efficient projects necessarily be favored by a majority of voters? Explain.

"10. "When an economic function is turned over to the government, social cooperation replaces personal self-interest." Is this statement true? Why *or* why not?

11. What is the shortsightedness effect? How does the shortsightedness effect influence the efficiency of public-sector action?

*12. What's wrong with this way of thinking? "Public policy is necessary to protect the average citizen from the power of vested interest groups. In the absence of government intervention, regulated industries such as airlines, railroads, and trucking, will charge excessive prices, products will be unsafe, and the rich would oppress the poor. Government curbs the power of special-interest groups."

13. "Since governnient-operated firms do not have to make a profit, they can usually produce at a lower cost and charge a lower price than privately owned enterprises." Evaluate this view.

14. What percentage of government income transfer payments go to the poor? Do you think that the political process in general works to the advantage of the poor? Why or why not?

15. Why does representative democracy often tax some people in order to provide benefits to others? When governments become heavily involved in tax-transfer activities, how will this involvement affect economic efficiency?

*16. The United States imposes highly restrictive sugar import quotas that result in a domestic price that is generally about three times as high as the world price. The quotas benefit sugar growers at the expense of consumers. Given that there are far more sugar consumers than growers, why aren't the quotas abolished? Has government action in this area improved the living standards of Americans? Why or why not?

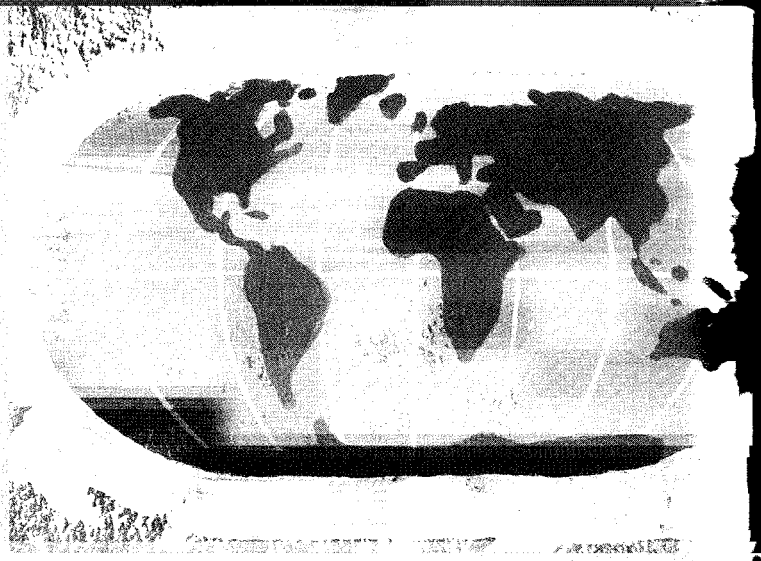
17. "The United States is rich because it is a political democracy where the people decide what policies will be followed." Is this statement true or false? Discuss.

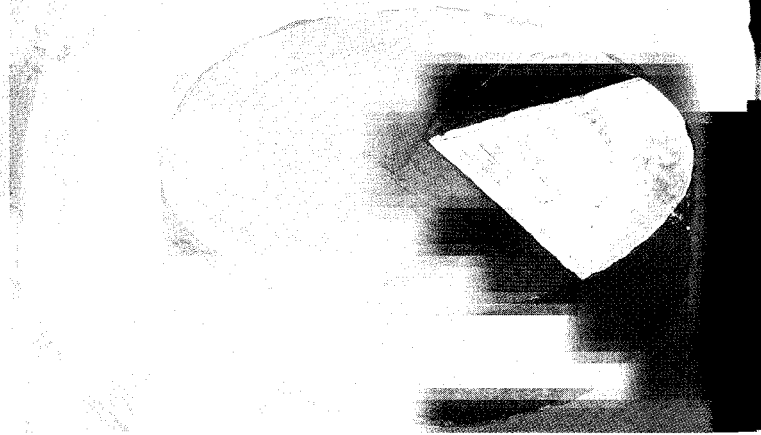
18. If the power of special interests were reduced, for example, through the adoption of a supra-majority voting rule, would economic efficiency improve? How would contributions to political campaigns be affected? Do you think politicians are very interested in curtailing the power of special interests? Why or why not?

*Asterisk denotes questions for which answers are given in Appendix B

PART 3

Microeconomics focuses on the choices of consumers, the operation of firms, the structure of markets, and the choices of resource suppliers and employers.





Core Microeconomics

PART 3

Consumers are the ultimate judge of both products and business firms. Their choices will determine which will survive and which will fail. The competitive process is highly dynamic. Numerous businesses come and go. Each year, newly incorporated businesses account for approximately 10 percent of the total. Approximately 1 percent of businesses file for bankruptcy during a typical year. In addition, many others close their doors or sell their assets to other, more successful (or more optimistic) operators.

CHAPTER 7

Consumer Choice and Elasticity

The most famous law in economics, and the one economists are most sure of, is the law of demand. On this law is built almost the whole edifice of economics.

—*David R. Henderson*[1]

A thing is worth whatever a buyer will pay for it.

—*Publilius Syrus, first century B.C.*[2]

Chapter Focus

- **What are the fundamental postulates underlying consumer choice?**
- **How does the law of diminishing marginal utility help explain the law of demand?**
- **How do the demand curves of individuals translate into a market demand curve?**
- **What is demand elasticity? What does it measure? Why is it important?**

[1]David R. Henderson, "Demand," in *The Concise Encyclopedia of Economics*, ed. David R. Henderson, http://www.econlib.org/library/CEE.html.

[2]Quoted in Michael Jackman, ed., *Macmillan Book of Business and Economic Quotations* (New York: Macmillan, 1984), 150.

The statement of David Henderson highlights the central position of the law of demand in economics. As Publilius Syrus noted more than 2,000 years ago, demand reflects the willingness of individuals *to* pay for what is offered in the market. In this section, we begin our examination of microeconomic markets for specific products with an analysis of the demand side of markets. In essence we will be going "behind" the market demand curve to see how it is made up of individual consumer demands and what factors determine the choices of individual consumers.[3]

FUNDAMENTALS OF CONSUMER CHOICE

Each of us must decide how to allocate our limited income among the many possible things we could purchase. The prices of goods, *relative to each other,* are very important determining factors. If your favorite cereal doubled in price, would you switch to a different brand? Would your decision be different if all cereals, not just yours, doubled in price? Your choice ***between*** brands of cereal will be affected only by the change in relative prices. If the prices of all cereals rose by a proportional amount, you might quit purchasing cereal, but this would not give you a strong reason to switch to a different brand. Relative prices measure opportunity cost. If cereal is *$5* per box when movie tickets are $10, you must give up two boxes of cereal to purchase a movie ticket.

Several fundamental principles underlie the choices of consumers. Let's take a closer look at the key factors influencing consumer behavior.

1. Limited income necessitates choice. Because of scarcity, we all have limited incomes. The limited nature of our income requires us to make choices about which goods we will and will not buy. When more of one good or service is bought, we must buy less of some other goods if we are to stay within our budget.

2. Consumers make decisions purposefully. The goals that underpin consumer choice can usually be met in alternative ways. If two products cost the same, a consumer will choose to buy the one expected to have the higher benefit. On the other hand, if two products yield equal benefits, the consumer will choose to buy the cheaper one. Fundamentally, we assume that consumers are rational—that they are able to weigh the costs and benefits of alternative choices.

3. One good can be substituted for another. Consumers can achieve utility—that is, satisfaction—from many different alternatives. Either a hamburger or a taco might satisfy your hunger, while going either to a movie or to a football game might satisfy your desire for entertainment. With $600, you might either buy a new TV set or take a short vacation. No single good is so precious that some of it will not be given up in exchange for a large enough quantity of other goods. Even seemingly unrelated goods are sometimes substituted one for another. For example, high water prices in Southern California have led residents there to substitute cactus gardens and flow constrictor shower heads for the relatively more expensive water. In Montana, where household electricity in recent years cost nearly twice as much as in nearby Washington, people substitute natural gas, fuel oil, insulation, and wool sweaters for the relatively more expensive electricity.

4. Consumers must make decisions without perfect information, but knowledge and past experience will help. In Chapter 1, we noted that information is costly to acquire. Asking family and friends, searching through magazines such as *Consumer Reports,* and contacting your local Better Business Bureau are all ways of gathering information about products and potential sellers. The time and effort consumers spend acquiring information will be directly related to the value derived from it. Predictably,

[3]You may want to review the section on demand in Chapter 3 before proceeding with this chapter.

consumers will spend more time and money to inform themselves when they are buying "big ticket" items such as automobiles or air-conditioning systems than when they are buying pencils or paper towels.

While no one has perfect foresight, experience — your own and that of others — will help you make better-informed choices. You have a pretty good idea of what to expect when you buy a cup of coffee at your favorite restaurant or ten gallons of gasoline at a service station you patronize regularly. Your expectations might not always be fulfilled precisely the same way every time (for example, the coffee may be stronger than expected or the gasoline may make your car's engine knock), but even then, you will gain valuable information that will help you project the outcome of future choices more accurately.

5. The law of diminishing marginal utility applies: as the rate of consumption increases, the marginal utility derived from consuming additional units of a good will decline. Utility is a term economists use to describe the subjective personal benefits that result from taking an action. The law **of diminishing marginal utility** states that the **marginal** (or additional) **utility** derived from consuming successive units of a product will eventually decline as the rate of consumption increases. For example, the law says that even though you might like ice cream, your marginal satisfaction from additional ice cream will eventually decline as you eat more and more of it. Ice cream at lunchtime might be great. An additional helping for dinner might also be good. However, after you have had it for lunch and dinner, another serving as a midnight snack will be less attractive. When the law of diminishing marginal utility sets in, the additional utility derived from still more units of ice cream declines.

Law of diminishing marginal utility
The basic economic principle that, as the consumption of a product increases, the marginal utility derived from consuming more of it (per unit of time) will eventually decline.

Marginal utility
The additional utility, or satisfaction, derived from consuming an additional unit of a good.

The law of diminishing marginal utility explains why, even if you really like a certain product, you will not spend your entire budget on it. As you increase your consumption of any good, including those that you like a lot, the utility you derive from each additional unit will become smaller and smaller and eventually it will be less than the cost of the unit. At that point, you will not want to purchase any more units of the good.

MARGINAL UTILITY, CONSUMER CHOICE, AND THE DEMAND CURVE OF AN INDIVIDUAL

The law of diminishing marginal utility helps us understand the law of demand and the shape of the demand curve. The height of an individual's demand curve at any specific unit is equal to the maximum price the consumer would be willing to pay for that unit — its **marginal benefit** to the consumer — given the number of units he or she has already purchased. Although marginal benefit is measured in dollars, the dollar amount reflects the opportunity cost of the unit in terms of other goods forgone. If a consumer is willing to pay, at most, $5 for an additional unit of the product, this indicates a willingness to give up, at most, $5 worth of other goods. ***Because a consumer's willingness to pay for a unit of a good is directly related to the utility derived from consuming the unit, the law of diminishing marginal utility implies that a consumer's marginal benefit, and thus the height of the demand curve, falls with the rate of consumption.***

Marginal benefit
The maximum price a consumer will be willing to pay for an additional unit of a product. It is the dollar value of the consumer's marginal utility from the additional unit, and therefore it falls as consumption increases.

Exhibit 1 shows this relationship for a hypothetical consumer Jones, relative to her weekly consumption of frozen pizza. Because of the law of diminishing marginal utility, each additional pizza consumed per week will generate less marginal utility for Jones than the previous pizza. For this reason, Jones's maximum willingness to pay — her marginal benefit — will fall as the quantity consumed increases. In addition, the steepness of Jones's demand curve, or its responsiveness to a change in price — its elasticity — is a reflection of how rapidly Jones's marginal utility diminishes with additional consumption. An individual's demand curve for a good whose marginal value declines more rapidly will be steeper.

Given what we now know about a consumer's maximum willingness to pay for additional units of a good, we are now in a position to discuss the choice of how many units the consumer will choose to purchase at various prices. ***At any given price, consumers***

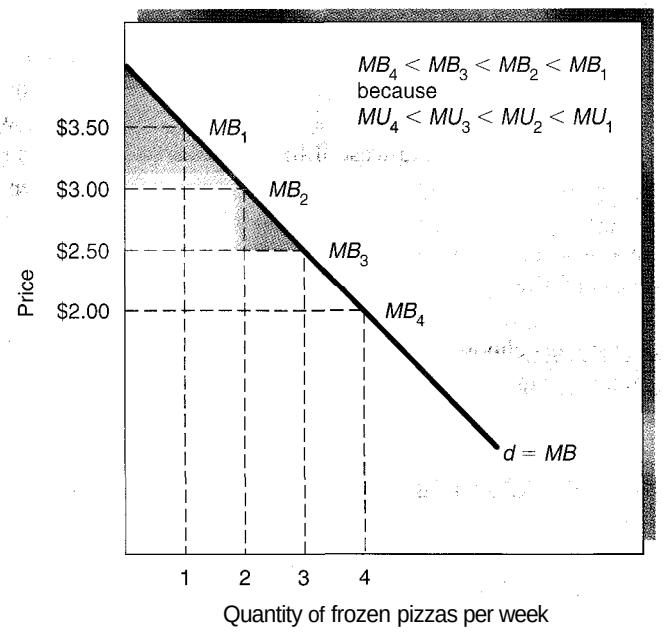


EXHIBIT 1
Diminishing Marginal Utility and the Individual's Demand Curve

An individual's demand curve, Jones's demand for frozen personal pizzas in this case, reflects the law of diminishing marginal utility. Because marginal utility *(MU)* falls with increased consumption, so does the consumer's maximum willingness to pay—marginal benefit (*MB*). A consumer will purchase until *MB = Price,* so at a price of $2.50 per pizza, Jones would purchase three pizzas and receive a consumer surplus shown by the shaded triangle.

will purchase all units of a good for which their maximum willingness to pay—their marginal benefit—is greater than the price. They will stop at the point where the next unit's marginal benefit would be less than the price. Although there are some problems related to dividing up certain kinds of goods (for example, it is hard to purchase half a car), we can generally say that a consumer will purchase all units of a good up to the point where the marginal benefit from it equals the price of the good ($MB = P$).

Returning to Exhibit 1, if the price of frozen pizza were $2.50, Jones would purchase three frozen pizzas per week.[4] Remember from Chapter 3 that consumer surplus is defined as the difference between the maximum price the consumer is willing to pay and the price actually paid. Jones's maximum willingness to pay for the first unit is $3.50, which, at a price of $2.50, generates $1.00 of consumer surplus for Jones. When a consumer has purchased all units to the point where $MB = P$, total consumer surplus is maximized. It is shown by the total triangular area under the demand curve that lies above the price. Total consumer surplus for Jones is shown as the shaded area in Exhibit 1.

Within this framework, how would a consumer respond to a decline in the price of a good? The consumer will increase purchases to the point where marginal benefit diminishes to the level of the new lower price. If marginal utility declines rapidly with consumption, the consumer will expand his or her purchases only slightly. If marginal utility declines less rapidly, it will take a larger expansion in purchases to reach this point. If the price were to rise, the consumer would cut back on purchases, eliminating those for which marginal benefit were now less than the price. This link between marginal benefit and maximum willingness to pay is the basis for the law of diminishing marginal utility, which underlies a person's demand curve for a product. The shape and steepness of the curve, for example, depends upon his or her marginal utility.

[4]Jones would certainly purchase the second unit because $MB > P$ For the third unit, $MB = P$ so Jones would be indifferent to buying the unit or not purchasing it For a good that is easily divisible, say, pounds of roast beef, the consumer would continue purchasing up to 2.9999 pounds. Thus, economist? are comfortable with simply concluding thdt the consumer will purchase this final unit, implying that Jones w.ll purchase three frozen pizzas

Consumer Equilibrium with Many Goods

The last time you were at the mall, you probably saw something, perhaps a nice billfold, that you liked. After all, there are many things we would like—many different alternatives that would give us utility. Next, you looked at the price tag: "Fifty dollars, wow! That's too much." What you were really saying was, "I like the billfold, but not as much as the $50 worth of other goods that I would have to give up for it." Consumer choice is a constant comparison of value relative to price. Consider another example: perhaps you prefer steak to less costly hamburger. Even if you do, your happiness might be better served if you were to buy the hamburger and then spend the extra money you saved on something else.

The idea that consumers choose among products by comparing their relative marginal utility (MU) to price (P) can be expressed more precisely. A consumer with a limited amount of income to spend on a group of products is not likely to do the following math, but will act as though he or she had, and will end up with

$$\frac{MU_A}{P_A} = \frac{MU_B}{P_B} = \ldots = \frac{MU_n}{P_n}$$

In this formula, MU represents the marginal utility derived from the last unit of a product, and P represents the price of the good. The subscripts $A, B, \ldots, n$ indicate the different products available to the consumer. ***This formula implies that the consumer will maximize his or her satisfaction (or total utility) by ensuring that the last dollar spent on each commodity yields an equal degree of marginal utility. Alternatively stated, the last unit of each commodity purchased should provide the same marginal utility per dollar spent on it.*** Thus, if the price of a gallon of ice cream is twice as high as the price of a liter of Coke, a consumer will purchase these items to the point where the marginal utility of the last gallon of ice cream is twice as high as the marginal utility of the last liter of Coke.

Perhaps the best way to grasp this point is to think about what happens when your ratios of marginal utility to price are not equal for two goods. Suppose that you are at a local restaurant eating buffalo chicken wings and drinking Coke. For simplicity, assume that a large Coke and an order of wings each costs $2. With your $10 budget, you decide to purchase four orders of wings and one large Coke. When you finish your Coke, there are still lots of wings left. You have already eaten so many wings, though, that those remaining do not look as attractive. You could get more utility with fewer wings and another Coke, but it is too late. You have not spent your $10 in a way that gets you the most for your money. Instead of satisfying the above condition, you find that the marginal utility of wings is lower than the marginal utility of a Coke, and because they both have the same price ($2), this implies that

$$\frac{MU_{wings}}{P_{wings}} < \frac{MU_{Coke}}{P_{Coke}}$$

If you had purchased fewer wings and more Coke, your total utility would have been higher. Spending more on Coke would have lowered its marginal utility, decreasing the value of the right side of the equation. Simultaneously, spending less on wings would have raised the marginal utility of wings, increasing the value of the left side of the equation. You will maximize your utility—and get the most "bang for the buck" from your budget—when you make these values (the ratios) equal.

The equation can also be used to illustrate the law of demand. Beginning with a situation in which the two sides were equal, suppose that the price of wings increased. It would lower the value of MU/P for wings below the MU/P for Coke. In response, you would reallocate your budget, purchasing fewer of the more costly wings and more Coke. Thus, we have the law of demand—as the price of wings rises, you will purchase less of them. When people try to spend their money in a way that gives them the greatest amount of satisfaction, the consumer decision-making theory outlined here is difficult to question. In the next section, we will take the theory a little further.

Price Changes and Consumer Choice

The demand curve or schedule shows the amount of a product that consumers are willing to buy at alternative prices during a specific time period. The law of demand states that the amount of a product bought is inversely related to its price. We have seen how the law of demand can be derived from fundamental principles of consumer behavior. Now, we go further and distinguish two different phenomena underlying a consumer's response to a price change. First, as the price of a product declines, the lower opportunity cost will induce consumers to buy more of it—even if they have to give up other products. This tendency to substitute a product that has become cheaper for goods that are now relatively more expensive is called the **substitution** effect of a price change.

Substitution effect
That part of an increase (decrease) in amount consumed that is the result of a good being cheaper (more expensive) in relation to other goods because of a reduction (increase) in price.

Second, if a consumer's money income is fixed, a reduction in the price of a product will increase his or her real income—the amount of goods and services he or she is able to purchase with that fixed amount of money income. If your rent were to decline by $100 per month, for example, that would allow you to buy more of a number of other goods. This increase in your real income has the same effect as if the rent had remained the same but your income had risen by $100 per month. As a result, this second way in which a price change affects consumption is called the **income effect.** Typically, consumers will respond to the income effect by buying more of the cheaper product and other products as well because they can better afford to do so. Substitution and income effects generally work in the same direction—in other words, in the same way; they both cause consumers to purchase more of a good as its price falls and less of a good as its price rises. (Both the income and substitution effects are derived graphically in the addendum to this chapter, titled "Consumer Choice and Indifference Curves.")[5]

Income effect
That part of an increase (decrease) in amount consumed that is the result of the consumer's real income being expanded (contracted) by a reduction (rise) in the price of a good.

Time Costs and Consumer Choice

You may have heard the saying that "time is money." It is certainly true that time has value and that this value can sometimes be measured in dollars. As we have learned, the monetary price of a good is not always a complete measure of its cost to the consumer. Consuming most goods requires not only money, but time as well; and time, like money, is scarce to the consumer. So a lower time cost, like a lower money price, will make a product more attractive. For example, patients in a dentist's office would prefer a shorter wait before receiving care. One study showed that dental patients are willing to pay more than $5 per minute saved to shorten their time spent in waiting rooms.[6] Similarly, commodities such as automatic dishwashers, prepared foods, air travel, and taxi service are demanded mainly for the time savings they offer. People are often willing to pay relatively high money prices for goods that help them save time.

Time costs, unlike money prices for goods, differ among individuals. They are higher for persons with higher wage rates, for example. Other things being equal, high-wage consumers choose fewer time-intensive (and more time-saving) commodities than people with lower time costs and wages. For example, high-wage consumers are overrepresented among airplane and taxicab passengers but underrepresented among television watchers, chess players, and long-distance bus travelers. Can you explain why? You can, if you understand how both money and time costs influence the choices of consumers.

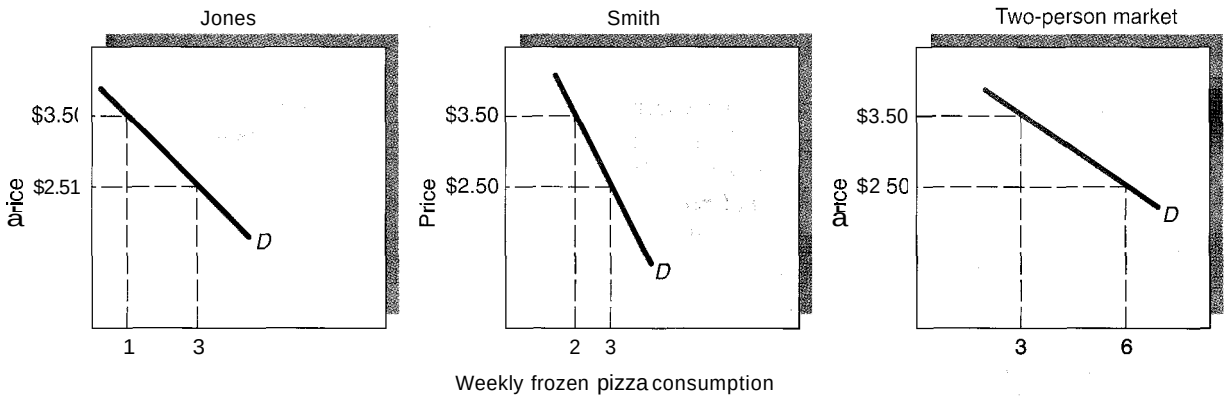
Failure to account for time costs can lead to bad decisions. For example, which is cheaper for consumers: (1) waiting in line three hours to purchase a $25 concert ticket or (2) buying the same ticket for $40 without standing in line? A consumer whose time is worth more than $5 per hour will find that $40 without the wait in line is less costly. As you can see, time costs matter. For example, when government-imposed price ceilings (discussed in Chapter 4) create shortages, rationing by waiting in line is frequently used. For many consumers, the benefit of the lower price due to the ceiling will be largely, if not entirely, offset by their increased time cost of having to wait in line.

[5]The substitution effect will always work in this direction. The income effect, however, may work in the reverse direction for some types of goods known as inferior goods. These will be addressed later in this chapter.

[6]Rexford E. Santerre and Stephen P. Neun, *Health Economics: Theories, insights and Industry Studies* (Orlando, Fla.: Harcourt, 2000), 113.

EXHIBIT 2
Individual and Market Demand Curves

The market demand curve is merely the horizontal sum of the individual demand curves. It will slope downward to the right just as individual demand curves do



MARKET DEMAND REFLECTS THE DEMAND OF INDIVIDUAL CONSUMERS

The market demand schedule is the relationship between the market price of a good and the amount demanded by all the individuals in the market area. Because individual consumers purchase less at higher prices, the amount demanded in a market area as a total is also inversely related to price.

Exhibit *2* shows the relationship between individual demand and market demand for a hypothetical two-person market. The individual demand curves for both Jones and Smith are shown. Jones and Smith each consume three frozen pizzas per week at a price of $2.50. The amount demanded in the two-person market is six pizzas. If the price rises to $3.50 per pizza, the amount demanded in the market will fall to three pizzas, one demanded by Jones and two by Smith. ***The market demand is simply the horizontal sum of the individual demand curves of consumers—in this case, Smith and Jones.***

In the real world, however, there can be millions of consumers in a market. But the relationship between the demand curves of individuals and the market demand curve will still be just like the one shown in Exhibit 2. At any given price, the amount purchased in the market will be the sum of the amounts purchased by each consumer in the market. Furthermore, the total amount demanded in the market will decline as price increases because individual consumers will purchase fewer units at the higher prices. The market demand curve reflects the collective choices of the individual consumers.

ELASTICITY OF DEMAND

Price elasticity of demand
The percentage change in the quantity of a product demanded divided by the percentage change in the price that caused the change in quantity. The price elasticity of demand indicates how responsive consumers are to a change in a product's price.

Although it is important to recognize that consumers will buy less of a product as its price increases, for many purposes, it is also important to know whether the increase will lead to a large or small reduction in the amount purchased. Economists have designed a tool called the price elasticity of demand to measure this sensitivity of amount purchased in response to a change in price. The equation for the **price elasticity of demand** is as follows:

$$\text{Price elacticity of demand} = \frac{\text{Percentage change in quantity demanded}}{\text{Percentage change in price}} = \frac{\%\Delta Q}{\%\Delta P}$$

This ratio is often called the *elasticity coefficient.* To express it more briefly, we use the notation $\%\Delta Q$ to represent percentage change in quantity and %AP to represent percentage change in price. (The Greek letter delta *[A]* means "change in.") The law of demand states that an increase in a product's price lowers the quantity of it purchased, whereas a decrease in price raises it. Because a change in price causes the quantity demanded to change in the opposite direction, the price elasticity coefficient is always negative, although economists often ignore the sign and use the absolute value of the coefficient.

To see how the concept of elasticity works, suppose that the price of the Ford Explorer rises 10 percent, while other prices remain the same. Ford could expect Explorer sales to fall substantially—perhaps 30 percent—as sport-utility vehicle (SUV) buyers respond by switching to other SUVs whose prices have not changed. This strong response by buyers means that the demand for the Explorer is elastic.

Now consider a different situation. Suppose that, because of a new tax, the price of not only the Explorer *but of all new SUVs* rises 10 percent. In this case, consumers' options are much more limited. They can't simply switch to a cheaper close substitute as they could when the price of the Explorer alone rose. They might either simply pay the extra money for a new SUV or settle for a used SUV instead. Because of this, the 10 percent rise in the price of all new SUVs will lead to a smaller consumer response, perhaps a 5 percent decline in sales of new SUVs.

To calculate the elasticity coefficient for the Explorer in our example above, we begin with the 30 percent decline in quantity demanded and divide by the 10 percent rise in the price that caused the decline. Thus, the elasticity of demand for the Explorer would be:

$$\frac{\%\Delta Quantity}{\%\Delta Price} = \frac{-30\%}{+10\%} = -3$$

(or 3.0 if we ignore the minus sign). This means that the percentage change in quantity demanded is three times the percentage change in price.

To calculate the demand elasticity for *all* SUVs (our second example), we see that the percentage change in quantity, 5 percent, divided by the percentage change in price, 10 percent, gives us $-1/2$, or -0.5. When it comes to the price elasticity of demand for all SUVs, the percentage change in quantity demanded (using our hypothetical numbers) is only half the percentage change in price, not three times the percentage change in price as it was with the Explorer.

Often, we will have to derive the percentage change in quantity and price. If you know the quantities that will be purchased at two different prices, you can then derive the percentage change in both the price and the quantity. For example, suppose that a price change from P_0 to P, causes a change in quantity demanded from Q, to Q,. The change in quantity demanded would therefore be Q, – Q,. To calculate the percentage change in quantity, we divide the actual change by the midpoint (or average) of the two quantities.[7] Although it is often easy to find the midpoint without a formula (halfway between \$4 and \$6 is \$5), it can also be found as $(Q, + Q_1)/2$. Finally, because 0.05 is simply 5 percent, we multiply by 100. Thus, we can express the percentage change in quantity demanded as:

$$\frac{Q_0 - Q_1}{(Q_0 + Q_1)/2} \times 100$$

Similarly, when the change in price is $P_0 - P_1$, the *percentage* change in price is

$$\frac{P_0 - P_1}{(P_0 + P_1)/2} \times 100$$

Dividing the resulting percentage change in quantity by the percentage change in price gives us the elasticity.

[7]This formula uses the average of the starting point and the ending point of the change so that it will give the same result whether we start from the lower or the higher price. This arc elasticity formula is not the only way to calculate elasticity, but it is the most frequently used.

There is a more direct way to compute elasticity using this same sort of approach. Dividing the percentage change in quantity by the percentage change in price and simplifying yields the following:

$$\frac{(Q_0 - Q_1)/(Q_0 + Q_1)}{(P_0 - P_1)/(P_0 + P_1)}$$

(Because each term is multiplied by 100 and the denominator of each term contains a 2, these factors cancel out of the final expression.)

A numerical example will help you understand this. Suppose that Trina's Cakes can sell fifty specialty cakes per week at $7 each, or it can sell seventy specialty cakes at $6 each per week. The percentage difference in quantity is the difference in the quantity demanded ($50 - 70 = -20$) divided by the midpoint (60) times 100. The result is a -33.33 percent change in quantity ($-20 \div 60 \times 100 = -33.33$).

Now that we've calculated the percentage change in quantity demanded of cakes, let's calculate the percentage change in the price. The percentage change in price is the difference in the two prices ($\$7 - \$6 = \$1$) divided by the midpoint price ($6.50) times 100, or a 15.38 percent change in price ($1 \div 6.5 \times 100 = 15.38$). Dividing the percentage change in quantity by the percentage change in price ($-33.33 \div 15.38$) gives an elasticity coefficient of -2.17. Alternatively, we could have expressed this directly as:

$$\frac{[(50 - 70)/(50 + 70)]}{[(7 - 6)/(7 + 6)]} = \frac{-20/120}{1/13} = \frac{-1/6}{1/13} = \frac{-13}{6} = -2.17$$

The same result is obtained either way. The elasticity of 2.17 (ignoring the sign) indicates that the percentage change in quantity is just over twice the percentage change in price.

The elasticity coefficient lets us make a precise distinction between elasticity and inelasticity. When the elasticity coefficient is greater than 1 (ignoring the sign), as it was for the demand for Trina's Cakes, demand is elastic. When it is less than 1, demand is inelastic. Demand is said to be of *unitary elasticity* if the price elasticity is exactly 1.

Graphic Representation of Price Elasticity of Demand

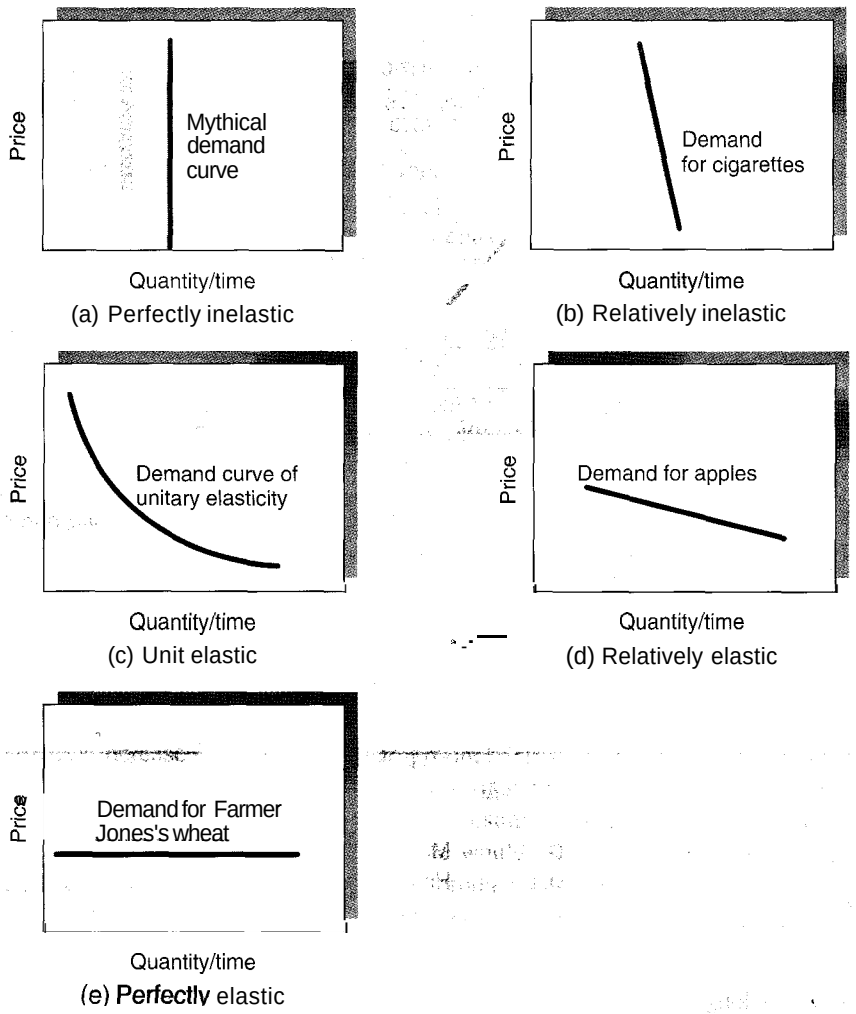
Exhibit 3 presents demand curves of varying elasticity. A demand curve that is completely vertical is said to be *perfectly inelastic*. In the real world, such demand is nonexistent because the substitutes for a good become more attractive as the price of that good rises. Moreover, because of the income effect, we should expect that a higher price will always reduce the quantity demanded, other things remaining the same. (Unlike our falling rent example earlier in the chapter, this example can be thought of as a "negative" income effect. When the price of a good rises, it leaves you with less overall buying power, not more.) Still, the (mythical) perfectly inelastic demand curve is shown in part (a) of Exhibit 3.

The more inelastic the demand, the steeper the demand curve *over any specific price range*. As you can see, the demand for cigarettes (shown in part b of Exhibit 3) is highly inelastic; a big change in price doesn't change quantity demanded much. People who crave nicotine will be very willing to pay the higher price. On the other hand, the demand for apples (shown in part d) is relatively elastic. People will find it easy to switch to oranges or bananas, for example, if the price of apples skyrockets.

When demand elasticity is unitary, as part (c) shows, a demand curve that is convex to the origin will result. When a demand curve is completely horizontal, an economist would say that it is *perfectly elastic*. Demand for the wheat marketed by a single wheat farmer, for example, would approximate perfect elasticity (part e).

Because elasticity is a relative concept, the elasticity of a straight line demand curve will differ ut each point along the demand curve. As **Exhibit 4** shows, the elasticity of a straight-line demand curve (one with a constant slope) will range from highly elastic to highly inelastic. In this exhibit, when the price rises from $10 to $11, sales decline from 20 to 10. According to the formula, the price elasticity of demand is -7.0. Demand is very elastic in this region. In contrast, demand is quite inelastic in the $1 to $2 price range. As

EXHIBIT 3
Price Elasticity of Demand



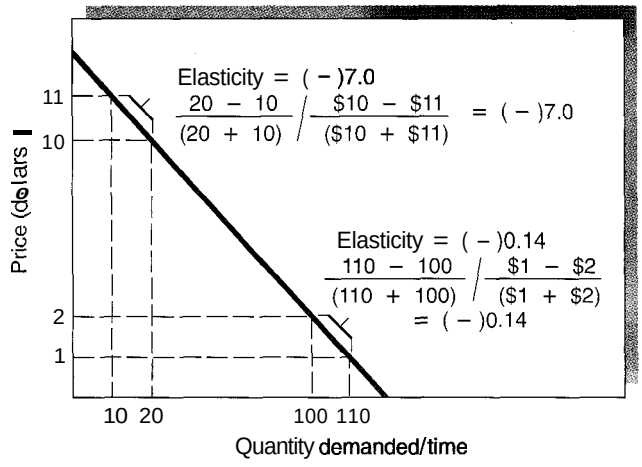
(a) Perfectly inelastic: Despite an increase in a product's price, consumers still purchase the same amount of it. Substitution and income effects prevent this from happening in the real world, though.
(b) Relatively inelastic: A percentage increase in a product's price results in a smaller percentage reduction in its sales. The demand for cigarettes has been estimated to be highly inelastic.
(c) Unit elastic: The percentage change in quantity demanded of a product is equal to the percentage change in its price. A curve with a decreasing slope results. Sales revenue (price times quantity sold) is constant.
(d) Relatively elastic: A percentage increase in a product's price leads to a larger percentage reduction in purchases of it. When good substitutes are available for a product (as in the case of apples), the amount of it purchased will be highly sensitive to price changes.
(e) Perfectly elastic: Consumers will buy all of Farmer Jones's wheat at the market price, but none will be sold above the market price.

the price increases from $1 to $2, the amount demanded declines from 110 to 100. The ten-unit change in quantity is the same, but it is a smaller *percentage* change. And the $1 change in price is the same, but it is now a larger *percentage* change. The elasticity of demand in this range is only −0.14; demand in this example is highly inelastic.

How Large Are the Demand Elasticities of Various Products?

Economists have estimated the price elasticity of demand for many products. As **Exhibit 5** shows, the elasticity of demand varies substantially among products. The demand is highly inelastic for several products — salt, toothpicks, matches, light bulbs, and newspapers, for example — in their normal price range. On the other hand, the demand curves for fresh tomatoes, Chevrolet automobiles, and fresh green peas are highly elastic. The demand for movies, housing, private education, radios, and television sets is near 1.0 (unitary).

EXHIBIT 4
Slope of the Demand Curve Versus Price Elasticity



With this straight line (constant slope) demand curve, demand is more elastic in the high price range. The formula for elasticity shows that, when price rises from $1 to $2 and quantity falls from 110 to 100, demand is inelastic. A price rise of the same magnitude (but of a smaller percentage), from $10 to $11 leads to a decline in quantity of the same size (but of a larger percentage), so that elasticity is much greater. (Price elasticities are negative, but economists often ignore the sign and look only at the absolute value.)

EXHIBIT 5
Estimated Price Elasticity of Demand for Selected Products

INELASTIC		APPROXIMATELY UNITARY ELASTICITY	
Salt	− 0.1	Movies	− 0.9
Matches	− 0.1	Housing, owner occupied, long run	− 1.2
Toothpicks	− 0.1	Shellfish, consumed at home	− 0.9
Airline travel, short run	− 0.1	Oysters, consumed at home	− 1.1
Gasoline, short run	− 0.2	Private education	− 1.1
Gasoline, long run	− 0.7	Tires, short run	− 0.9
Residential natural gas, short run	− 0.1	Tires, long run	− 1.2
Residential natural gas, long run	− 0.5	Radio and television receivers	− 1.2
Coffee	− 0.25	**ELASTIC**	
Fish (cod), consumed at home	− 0.5	Restaurant meals	− 2.3
Tobacco products, short run	− 0.45	Foreign travel, long run	− 4.0
Legal services, short run	− 0.4	Airline travel, long run	− 2.4
Physician services	− 0.6	Fresh green peas	− 2.8
Dental services	− 0.7	Automobiles, short run	− 1.2–1.5
Taxi, short run	− 0.6	Chevrolet automobiles	− 4.0
Automobiles, long run	− 0.2	Fresh tomatoes	− 4.6

Sources: Hendrick S Houthakker and Lester D. Taylor, Consumer Demand *in* the United States, *1929-1970* (Cambridge, Mass : Harvard University Press, 1966, 1970); Douglas R. Bohi, Analyzing Demand *Behavior* (Baltimore: Johns Hopkins University Press, 1981); Hsaing-tai Cheng and Oral Capps Jr , "Demand for Fish," *American* Journal *of Agricultural* Economics, August 1988; U.S. Department *of* Agriculture; and Rexford E. Santerre and Stephen P. Neun, Health *Economics: Theories,* Insights and *Industry* Studies (Orlando, Fla.. Harcourt, 2000).

Why Do the Price Elasticities of Demand Vary?

The primary determinants of a product's price elasticity of demand are the availability of good substitutes and to some extent the share of the typical consumer's total budget expended on a product. Let's consider each of these factors.

Availability of Substitutes ***The most important determinant of the price elasticity of demand is the availability of substitutes. When good substitutes for a product are available, a price increase induces many consumers to switch to other products. Demand is elastic.*** For example, if the price of felt tip pens increases, many consumers will switch to pencils, ballpoint pens, or (for children) crayons. If the price of apples increased, consumers might substitute oranges, bananas, peaches, or pears.

When good substitutes are unavailable, the demand for a product tends to be inelastic. Medical services are an example. When we are sick, most of us find witch doctors, faith healers, palm readers, and aspirin to be highly imperfect substitutes for the services of a physician. Not surprisingly, the demand for physician services is inelastic.

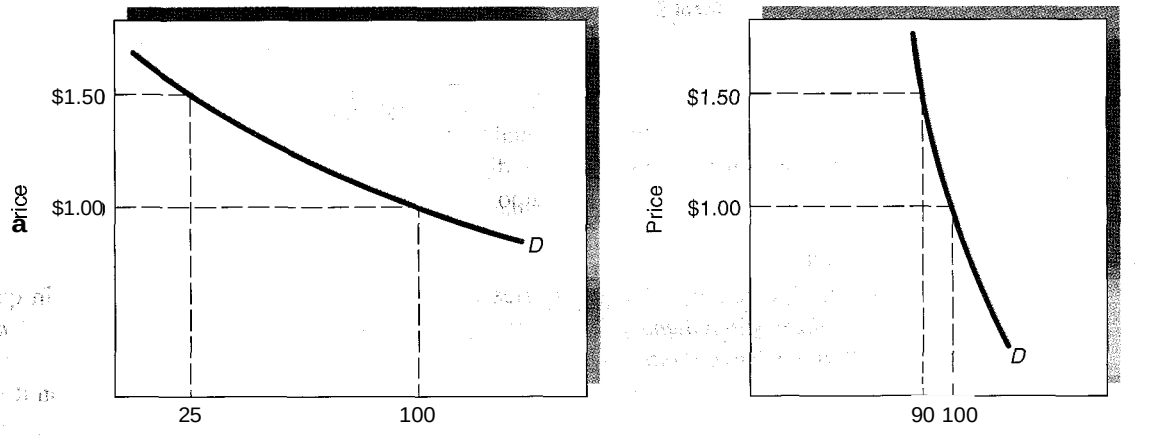
The availability of substitutes increases as the product class becomes more specific, thus increasing price elasticity. For example, as Exhibit 5 shows, the price elasticity of Chevrolets, a narrow product class, exceeds that of the broad class of automobiles in general. If the price of Chevrolets alone rises, many substitute cars are available. But if the prices of all automobiles rise together, consumers have fewer good substitutes.

Product's Share of the Consumer's Total Budget If the expenditures on a product are quite small relative to the consumer's budget, the income effect will be small even if there is a substantial increase in the price of the product. This will make demand less elastic. Compared to one's total budget, expenditures on some commodities are minuscle. Matches, toothpicks, and salt are good examples. Most consumers spend only $1 or $2 per year on each of these items. A doubling of their price would exert little influence on a family's budget. Therefore, even if the price of such a product were to rise sharply, consumers will still not find it worthwhile to spend much time and effort looking for substitutes.

Exhibit 6 provides a graphic illustration of both elastic and inelastic demand curves. In part (a), the demand curve for ballpoint pens is elastic because there are good

EXHIBIT 6
Inelastic and Elastic Demand

As the price of ballpoint pens (a) rose from $1.00 to $1.50, the quantity purchased plunged from 100,000 to 25,000. The percentage reduction in quantity is larger than the percentage increase in price. Thus, the demand for the pens is elastic. In contrast, an increase in the price of cigarettes from $1.00 to $1.50 results in only a small reduction in the number purchased (b). Because the demand for cigarettes is inelastic, the percentage reduction in quantity is smaller than the percentage increase in price.



substitutes — for example, pencils and felt tip pens. Therefore, when the price of the pens increases from $1.00 to $1.50, the quantity purchased of them declines sharply from 100,000 to only 25,000. The calculated price elasticity equals −3.0. The fact that the absolute value of the coefficient is greater than 1 confirms that the demand for ballpoint pens is elastic over the price range shown.

Part (b) of Exhibit 6 shows the demand curve for cigarettes. Because most smokers do not find other products to be a good substitute, the demand for cigarettes is highly inelastic. If a unit of six cigarettes is worth a dollar, a substantial increase in price (from $1.00 to $1.50) leads to only a small reduction in the quantity demanded. The price elasticity coefficient is −0.26, substantially less in absolute value than 1, confirming that the demand for cigarettes is inelastic. *(Exercise:* Use the price elasticity formula to verify the values of these elasticity coefficients.)

Time and Demand Elasticity

As changing market conditions raise or lower the price of a product, both consumers and producers will respond. However, the response will not be instantaneous, and it is likely to become larger over time. ***In general, when the price of a product increases, consumers will reduce their consumption by a larger amount in the long run than in the short run. Thus, the demand for most products will be more elastic in the long run than in the short run. This relationship between elasticity and the length of the adjustment period is sometimes referred to as the second law of demand.***

The first law of demand says that buyers will respond predictably to a price change, purchasing more when the price is lower than when the price is higher, if other things remain the same. The second law of demand says that the response of buyers will be greater after they have had time to adjust more fully to a price change.

HOW DEMAND ELASTICITY AND PRICE CHANGES AFFECT TOTAL EXPENDITURES (OR REVENUES) ON A PRODUCT

By looking at demand elasticity, we can determine changes in total consumer spending on a product when its price changes. We can do this in three different ways: by looking at (1) changes in an individual's total spending, using the demand elasticity from his or her demand curve for the product, (2) changes in the total combined spending of all consumers, using the elasticity from the total market demand curve, or (3) changes in total consumer spending on the product, using the demand curve facing that firm that produces it. In other words, this third method allows us to look at elasticity based not on what consumers spend, but on what the producer receives from selling the product.

Total expenditures (or revenues) simply amount to the price of the product times the number of units of it purchased (or sold).

Total Expenditures	=	Price	x	Quantity
?	−	↑	x	↓
?	−	↓	x	↑

Because total expenditures are equal to the price times the quantity, and because the price and the quantity move in opposite directions, the net effect of a price change on the total spending on a product depends upon whether the (percentage) price change or the (percentage) quantity change is greater.

When demand is inelastic, the price elasticity coefficient is less than 1. This means that the percentage change in price is greater than the percentage change in quantity. ***Therefore, when demand is inelastic, the change in the price will dominate and, as a result, the price and total expenditures will change in the sume direction.*** In other words, when the price of an inelastic product (say, cigarettes) increases, spending on it will increase, too — and vice versa. On the other hand, when demand is elastic, the change in quantity will be greater than the change in the price. ***As a result, the impact of the change in quantity will dominate, and therefore the price and expenditures will move in oppo-***

EXHIBIT 7
Demand Elasticity and How Changes in Price Affect Total Consumer Expenditures or a Firm's Total Revenue

Price Elasticity of Demand	Numerical Elasticity Coefficient (in absolute value)	Impact of raising Price on Total Consumer Expenditures or a Firm's Total Revenue	Impact of Lowering price on Total Consumer Expenditures or a Firm's Total Revenue
Elastic	1 to ∞	decrease	increase
Unit Elastic	1	unchanged	unchanged
Inelastic	0 to 1	increase	decrease

site directions. In other words, when the price of an elastic product (say, a ballpoint pen) increases, spending on it will decrease—and vice versa.

When demand elasticity is unitary, the change in quantity demanded will be equal in magnitude to the change in price. With regard to their impact on total expenditures, these two effects will exactly offset each other. ***Thus, when price elasticity of demand is equal to 1, total expenditures will remain unchanged as price changes.***

Exhibit 7 summarizes the relationship between changes in the price of a product and changes in total spending on it when demand is elastic, inelastic, and unit elastic. The demand curves shown in Exhibit 6 can also be used to show the link between elasticity and changes in total spending. In the case of cigarettes (part b), the price elasticity of demand for the price increase from $1.00 to $1.50 is 0.26, indicating that demand is inelastic. This increase in cigarette prices leads to an increase in spending on the product from $100 million ($1.00 × 100 million units) to $135 million ($1.50 × 90 million units). If the change had occurred in the opposite direction, with the price falling from $1.50 to $1.00, total expenditures would have declined.

The price elasticity of demand for a ballpoint pen when its price increases from $1.00 to $1.50 (part a of Exhibit 6) is 3.0, indicating that demand is elastic. This increase in the price of ballpoint pens lowers total consumer spending on the product from $100,000 ($1.00 × 100,000 pens) to $37,500 ($1.50 × 25,000 pens). If the change had occurred in the opposite direction, with the price falling from $1.50 to $1.00, total expenditures would have risen.

You can see how important it is for business decision makers to understand the concept of elasticity. When a firm increases the price of its product, its revenues may rise, fall, or remain the same. If the demand for the firm's product is inelastic, the higher price will expand the firm's total revenue. However, if the demand for the firm's product is elastic, a price increase will lead to substantially lower sales and a decline in total revenue. In the case of unitary elasticity, the price increase will leave total revenue unchanged.

Beyond the price elasticity of demand, two other elasticity relationships are important in any given market. We therefore end this chapter with a brief discussion of income elasticity of demand and price elasticity of supply.

INCOME ELASTICITY

Increases in consumer income will increase the demand (the quantity demanded at each price) for most goods. Income elasticity tells us how responsive the demand for a product is to income changes. **Income elasticity** is defined as:

$$\text{Income elasticity} = \frac{\text{Percentage change in quantity demanded}}{\text{Percentage change in income}}$$

As **Exhibit 8** shows, although the income elasticity coefficients for products vary from one good to another, they are normally positive. In fact, the term **normal good** refers to any good with a positive income elasticity of demand. Some normal goods have lower

Income elasticity
The percentage change in the quantity of a product demanded divided by the percentage change in consumer income that caused the change in quantity demanded. It measures the responsiveness of the demand for a good to a consumer's change in income.

Normal good
A good that has a positive income elasticity. so that, as consumer income rises, demand for the good rises, too.

EXHIBIT 8
Estimated Income Elasticity of Demand for Selected Products

LOW-INCOME ELASTICITY		HIGH-INCOME ELASTICITY	
Margarine	−0.20	Private education	2.46
Fuel	0.38	New cars	2.45
Electricity	0.20	Recreation and amusements	1.57
Fish (haddock)	0.46	Alcohol	1.54
Food	0.51		
Tobacco	0.64		
Hospital care	0.69		

Sources: Hendrick S. Houthakker and Lester D. Taylor, Consumer Demand in the *United* States, *1929–1970* (Cambridge, Mass : Harvard University Press, 1966); L. Taylor, "The Demand for Electricity' A Survey," Bell Journal *of Economics* (Spring 1975); F. W. Bell, "The Pope and the Price of Fish," American *Economic* Review 58 (December 1968); and Rexford E. Santerre and Stephen P. Neun, Health Economics: Theories, *Insights* and Industry *Studies* (Orlando, Fla.: Harcourt, 2000).

income elasticities than others, however. In general, goods that people regard as "necessities" will have low income elasticities (between 0 and 1). Significant quantities are purchased even at low incomes, and, as income increases, spending on these items will increase by less than a proportional amount. It is understandable that items such as fuel, electricity, bread, tobacco, economy clothing, and potatoes have a low income elasticity.

Goods that consumers regard as "luxuries" generally have a high (greater than 1) income elasticity. For example, private education, new automobiles, recreational activities, donations to environmental groups, swimming pools, and vacation air travel are all highly income elastic. As the consumer's income increases, the demand for these goods expands even more rapidly, and therefore spending on these items increases as a proportion of income.

A few commodities, such as margarine, low-quality meat cuts, and bus travel, actually have a negative income elasticity. Economists refer to goods with a negative income elasticity as **inferior goods. As** income expands, the demand for inferior goods will decline. Conversely, as income declines, the demand for inferior goods will increase.

Inferior good
A good that has a negative income elasticity, so that, as consumer income rises, the demand for the good falls.

PRICE ELASTICITY OF SUPPLY

Price elasticity of supply
The percentage change in quantity supplied, divided by the percentage change in the price that caused the change in quantity supplied.

The **price elasticity of supply** is the percentage change in quantity supplied, divided by the percentage change in the price causing the supply response. Because this measures the responsiveness of sellers to a change in price, it is analogous to the price elasticity of demand. However, the price elasticity of supply will be positive because the quantity producers are willing to supply is directly related to price. Like demand elasticity, time plays a role once again. Supply elasticities will be greater when suppliers have a longer time to respond to a price change. In the next two chapters, we will discuss more fully the factors that determine supply elasticity. For now, it is important simply to recognize the concept of supply elasticity and the fact that suppliers (like buyers) will be more responsive to a price change when they have had more time to adjust to it.

LOOKING AHEAD

Market demand indicates how strongly consumers desire a good or service. In the following chapter, we will turn to a firm's costs of production—costs that arise because resources are demanded for alternative uses. These two topics—consumer demand and the cost of production—are central to understanding how markets work and the conditions necessary for the efficient allocation of resources.

KEY POINTS

- Consumers will try to allocate their limited incomes among a multitude of goods in a way that maximizes their utility. The role of relative prices, information, and preferences, as well as the law of diminishing marginal utility help explain the choices consumers make and the downward slope of a person's demand curve for products.
- The market demand curve for a product is the horizontal sum of people's individual demand curves.
- The price elasticity of demand measures the responsiveness of the quantity of a product purchased to a change in its price.
- The availability of substitutes is the primary determinant of the price elasticity of demand for a product. When there are good substitutes available, and the item is a sizable component of the consumer's budget, its demand will tend to be elastic. When only poor substitutes are available, demand will tend to be inelastic.
- Typically, the price elasticity of a product will increase as consumers have had more time to adjust to a change in its price. This direct relationship between the size of the elasticity coefficient and the length of the adjustment period is often referred to as the *second law of demand.*
- The concept of elasticity helps us determine how a change in price will affect total consumer expenditures on a product or a firm's total revenues derived from it. When the demand for a product is elastic, a price change will cause total spending on it to change in the opposite direction. When demand for a product is inelastic, a change in price will cause total spending on it to change in the same direction.
- The concept of elasticity can also be applied to consumer income (which is called the income elasticity of demand) and supply (which is called the price elasticity of supply).

CRITICAL ANALYSIS QUESTIONS

***1.** Suppose that, in an attempt to raise more revenue, Nowhere State University (NSU) increases its tuition. Will this necessarily result in more revenue? Under what conditions will revenue (a) rise, (b) fall, or (c) remain the same? Explain this, focusing on the relationship between the increased revenue from students who enroll at NSU despite the higher tuition and the lost revenue from lower enrollment. If the true price elasticity were -1.2, what would you suggest the university do to expand revenue?

***2.** A bus ticket between two cities costs $50 and the trip will take twenty-eight hours, whereas an airplane ticket costs $300 and takes three hours. Mary values her time at $12 per hour, and Michele values her time at $8 per hour. Will Mary take the bus or the plane? Which will Michele take? Explain.

***3.** Recent research confirms that the demand for cigarettes is inelastic, but it also indicates that smokers with incomes in the lower half of all incomes respond to a given price increase by reducing their purchases by amounts that are more than four times as large as the purchase reductions made by smokers in the upper half of all incomes. How can the income and substitution effects of a price change help explain this finding?

4. A consumer is currently purchasing three pairs of jeans and five T-shirts per year. The price of jeans is $30, and T-shirts cost $10. At the current rate of consumption, the marginal utility of jeans is 60, and the marginal utility of T-shirts is 30. Is this consumer maximizing his utility? Would you suggest that he buy more jeans and fewer T-shirts, or more T-shirts and fewer jeans?

5. A few years ago, when residential electricity in the state of Washington cost about half as much as in nearby Montana, the average household in Washington used about 1,200 kilowatt-hours per month, whereas Montanans used about half that much per household. Do these data provide us with two points on the average household's demand curve for residential electricity in this region? Why or why not?

***6.** The wealthy are widely believed to have more leisure time than the poor. However, even though we are a good deal wealthier today than our great-grandparents were 100 years ago, we appear to live more hectic lives and have less free time. Can you explain why?

7. What are the major determinants of a product's price elasticity of demand? Studies indicate that the demand for Florida oranges, Bayer aspirin, watermelons, and airfares to Europe are elastic. Why?

8. Most syctems of medical insurance substantially lower the out-of-pocket costs consumers have to pay for additional units of physician services and hospitalization. Some reduce these costs to zero. How does this method of payment affect the consumption levels of medical services? Might this method of organization result in "too much" consumption of medical services? Discuss.

*9. Are the following statements true or false? Explain your answers.
   a. A 10 percent reduction in price that leads to a 15 percent increase in the amount purchased indicates a price elasticity of more than 1.
   b. A 10 percent reduction in price that leads to a 2 percent increase in total expenditures indicates a price elasticity of more than 1.
   c. If the percentage change in price is less than the resultant percentage change in quantity demanded, demand is elastic.

*10. Respond to the following questions: If you really like pizza, should you try to consume as much pizza as possible? If you want to succeed, should you try to make the highest possible grade in your economics class?

*11. Sue loves ice cream but cannot stand frozen-yogurt desserts. In contrast, Carole likes both foods and can hardly tell the difference between the two. Who will have the more elastic demand for yogurt?

*12. Patsy's Specialty Bakery projects the following demand for Patsy's pies:

Price ($)	Quantity Purchased
9	130
10	110
11	95

   a. Calculate the price elasticity of demand between $9 and $10. Is demand in this range elastic or inelastic?
   b. Calculate the price elasticity of demand between $10 and $11. Is demand in this range elastic or inelastic?

13. Suppose Bobby, the owner-manager of Bobby's Red Hot BBQ restaurant, projects the following demand for his Baby Back Rib platter:

Price ($)	Quantity Purchased (per night)
9	110
11	100
13	80

   a. Calculate the price elasticity of demand between $9 and $11.
   b. Is the price elasticity of demand between $9 and $11 elastic, unit elastic, or inelastic?
   c. Will Bobby's total revenue rise if he increases the price from $9 to $11?
   d. Calculate the price elasticity of demand between $11 and $13.
   e. Is the price elasticity of demand between $11 and $13 elastic, unit elastic, or inelastic?
   f. Will Bobby's total revenue rise if he increases the price from $11 to $13?

*Asterisk denotes questions for which answer? are given in Appendix B

ADDENDUM: CONSUMER CHOICE AND INDIFFERENCE CURVES

Advanced Material

In many classes, this material will be optional.

In the text of this chapter, we used marginal utility analysis to develop the demand curve of an individual. In developing the theory of consumer choice, economists usually rely on a more formal technique—*indifference curve* analysis. Because this technique is widely used at a more advanced level, many instructors like to include it in their introductory course. In this addendum, we use indifference curve analysis to develop the theory of demand in a more formal—some would say more elegant—manner.

What Are Indifference Curves?

There are two elements in every choice: (1) preferences (the desirability of various goods) and (2) opportunities (the attainability of various goods). The **indifference curve** relates to the former: the preferences of an individual. It separates better (more preferred by this individual) bundles of goods from inferior (less preferred) bundles, providing a diagrammatic picture of how an individual ranks alternative consumption bundles.

To illustrate indifference curves, we begin with the title character from the classic *Robinson Crusoe,* by Daniel Defoe, published in 1719 and thought to be the first English novel. Crusoe was shipwrecked on a desert island. In **Exhibit A1,** we assume that he is initially consuming eight fish and eight breadfruit per week (point **A**). This initial bundle provides him with a certain level of satisfaction (utility). He would, however, be willing to trade this initial bundle for certain other consumption alternatives if the

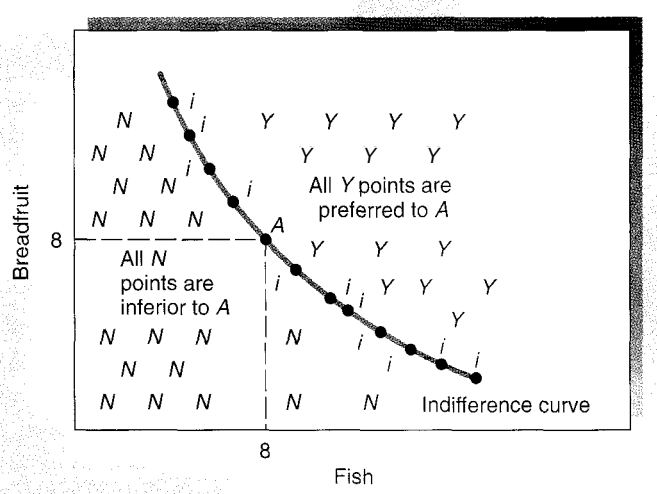


EXHIBIT A1
Indifference Curve of Robinson Crusoe

The curve generated by connecting Crusoe's "I do not care" answers separates the combinations of fish and breadfruit that he prefers to the bundle *A* from those that he judges to be inferior to *A*. The *i* points map out an indifference curve.

opportunity presented itself. Since he likes both fish and breadfruit, he would especially like to obtain bundles to the northeast of point ***A*** in the diagram, since they represent more of both goods. However, he would also be willing to give up some breadfruit if in return he received a compensatory amount of fish. Similarly, if the terms of trade were right, he would be willing to exchange fish for breadfruit. The trade-offs he is just willing to make—those that would make him no better and no worse off—lie *along* the indifference curve. Of course, he is happy to move to any bundle on a higher indifference curve.

Starting from point **A** (eight fish and eight breadfruit), we ask Crusoe if he is willing to trade that bundle for various other bundles. He answers "Yes" (*Y*), "No" (*N*), or "I do not care" (*i*). Exhibit A1 shows the pattern of his responses. Crusoe's "I do not care" answers indicate that the original bundle (point ***A***) and each alternative indicated by an *i* are valued equally by Crusoe. These *i* points, when connected, form the indifference curve. This line separates the preferred bundles of fish and breadfruit from the less-valued combinations. Note that such a curve is likely to be entirely different for any two people. The preferences of individuals vary widely.

We can establish a new indifference curve for the individual by starting from any point not on the original curve and following the same procedure. If we start with a point (a consumption bundle) to the northeast of the original indifference curve, all points on the new curve will have a higher level of satisfaction for Crusoe than any on the old curve. The new curve will probably have about the same shape as the original.

Characteristics of Indifference Curves

In developing consumer theory, economists assume that the preferences of consumers exhibit certain properties. These properties enable us to make statements about the general pattern of indifference curves. What are these properties, and what do they imply about the characteristics of indifference curves?

1. More goods are preferable to fewer goods—thus, bundles on indifference curves lying farthest to the northeast of a diagram are always preferred. Assuming the consumption of only two commodities that are both desired, the individual will always prefer a bundle with more of one good (without loss of the other) to the original bundle. This means that combinations to the northeast of a point on the diagram will always be preferred to points lying to the southwest.

2. Goods are substitutable—therefore, indifference curves slope downward to the right. As we discussed in this chapter, individuals are willing to substitute one good for another. Crusoe will be willing to give up some breadfruit if he is compensated with enough fish. Stated another way, there will be some amount of additional fish such that Crusoe will stay on the same indifference curve, even though his consumption of breadfruit has declined. However, in order to remain on the same indifference curve, Crusoe must always acquire more of one good to compensate for the loss of the other. The indifference curve for goods thus will always slope downward to the right (run northwest to southeast).

3. The value a consumer puts on a *good* declines as he or she consumes more of it—therefore, indifference curves are always convex when viewed from below. The slope of the indifference curve represents the willingness of the individual to substitute one good for the other. Economists refer to the amount of one good that is just sufficient to compensate the consumer for the loss of a unit of the other good as the **marginal rate of substitution**. It is equal to the slope of the indifference curve. Reflecting the principle of diminishing marginal utility, the marginal rate of substitution of a good will

decline as the good is consumed more intensively relative to other goods. Suppose Crusoe remains on the same indifference curve while continuing to expand his consumption of fish relative to breadfruit. As his consumption of fish increases (and his consumption of breadfruit declines), his valuation of fish relative to breadfruit will decline. It will take more and more units of fish to compensate for the loss of still another unit of breadfruit. The indifference curve will become flatter, reflecting the decline in the marginal rate of substitution of fish for breadfruit as Crusoe consumes more fish relative to breadfruit.

Of course, just the opposite will happen if Crusoe's consumption of breadfruit increases relative to that of fish—if he moves northwest along the same indifference curve. In this case. as breadfruit is consumed more intensively, Crusoe's valuation of it will decline relative to that of fish, and the marginal rate of substitution of fish for breadfruit will rise (the indifference curve will become steeper and steeper). Therefore, since the valuation of a good declines as it is consumed more intensively, indifference curves must be convex when viewed from the origin.

4. Indifference curves are everywhere. We can construct an indifference curve starting from any point on the diagram. This simply means that any two bundles of goods can be compared by the individual.

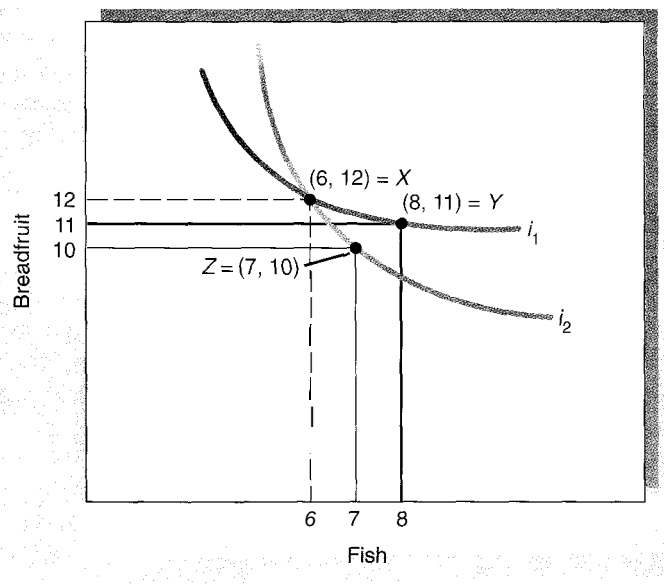
5. Indifference curves cannot cross—if they did, rational ordering would be violated. If indifference curves crossed, our postulate that more goods are better than fewer goods would be violated. **Exhibit A2** shows this point. The crossing of the indifference curves implies that points *Y* and Z are equally preferred because they both are on the same indifference curve as X. Consumption bundle *Y,* though, represents more of both fish and breadfruit than bundle *Z,* so *Y* must be preferred to Z. Whenever indifference curves cross, this type of internal inconsistency (irrational ranking) will arise. So, the indifference curves of an individual must not cross.

The Consumer's Preferred Bundle

Used together with the opportunity constraint of the individual, indifference curves can be used to indicate the most preferred consumption alternatives available to an individual. The **mption opportunity constraint** separates consumption bundles that are attainable from those that are unattainable.

Assuming that Crusoe could produce only for himself, his consumption opportunity constraint would look like the production possibilities curves discussed in Chapter 2. What would happen if natives from another island visited Crusoe and offered to make exchanges with him? If a barter market existed that permitted Cmsoe to exchange fish for breadfruit at a specified exchange rate. his options would resemble those of the market constraint shown by **Exhibit A3.** First, let us consider the case where Crusoe inhabits a barter economy in which the current market exchange rate is two fish to one breadfruit. Suppose that, as a result of his expertise as a fisherman, Crusoe specializes in this activity and is able to bring sixteen fish to the market per week. What consumption alternatives will be open to him? Because two fish can

EXHIBIT A2
Indifference Curves Cannot Cross



If the indifference curves of an individual crossed, they would show the inconsistency pictured here. Points X and *Y* must be equally valued, since they are both on the same indifference curve (i_1). Similarly, points X and *Z* must be equally preferred, since they are both on the indifference curve (i_2). If this is true, *Y* and *Z* must also be equally preferred, since they are both equally preferred to X. However, point *Y* represents more of both goods than *Z,* so *Y* has to be preferred to *Z.* When indifference curves cross, this type of internal inconsistency always arises.

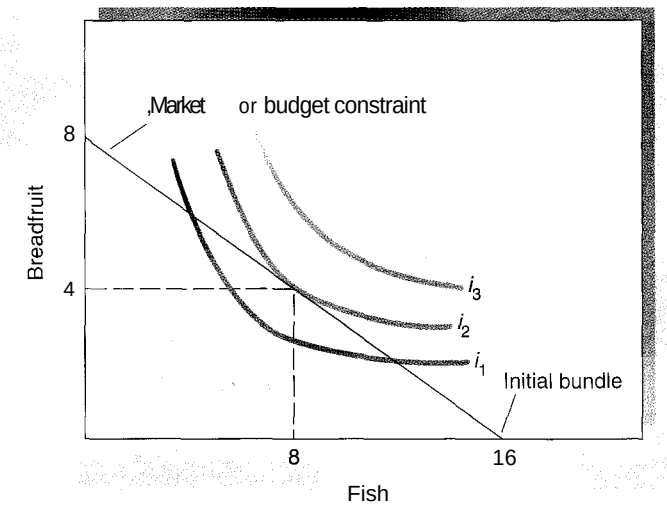


EXHIBIT A3
Consumer Maximization in a Barter Economy

Suppose that the set of indifference curves shown here outlines Crusoe's preferences. The slope of the market (or budget) constraint indicates that two fish trade for one breadfruit in this barter economy. If Crusoe produces sixteen fish per week, he will trade eight fish for four breadfruit in order to move to the consumption bundle (eight fish and four breadfruit) that maximizes his level of satisfaction.

be bartered in the market for one breadfruit, Crusoe will be able to consume sixteen fish, or eight breadfruit, or any combination on the market constraint indicated by the line between these two points. For example, if he trades two of his sixteen fish for one breadfruit, he will be able to consume a bundle consisting of fourteen fish and one breadfruit. Assuming that the set of indifference curves of Exhibit A3 outlines Crusoe's preferences, he will choose to consume eight fish and four breadfruit. Of course, it will be possible for Crusoe to choose many other combinations of breadfruit and fish, but none of the other attainable combinations would enable him to reach as high a level of satisfaction. Because he is able to bring only sixteen fish to the market, it would be impossible for him to attain an indifference curve higher than i_3

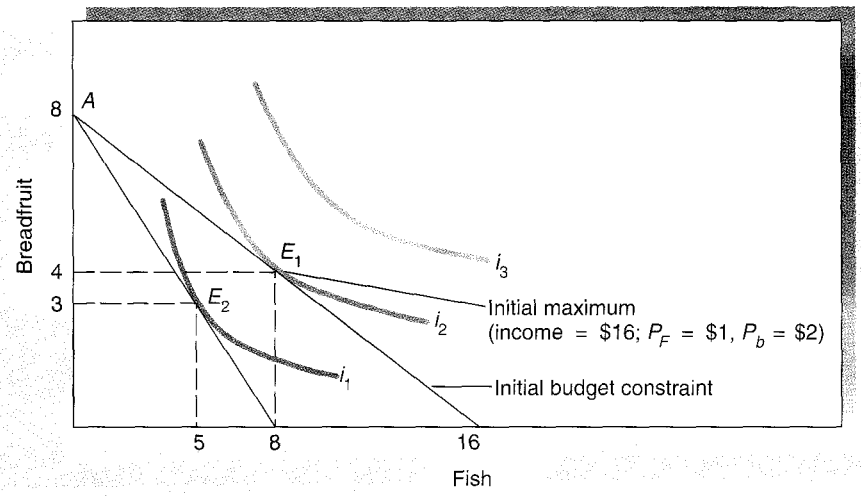
Crusoe's indifference curve and the market constraint curve will coincide (they will be tangent) at the point at which his attainable level of satisfaction is maximized. At that point (eight fish and four breadfruit), the rate at which Crusoe is willing to exchange fish for breadfruit (as indicated by the slope of the indifference curve) will be just equal to the rate at which the market will ***permit*** him to exchange the two (the slope of the market constraint). If the two slopes differ at a point, Crusoe will always be able to find an attainable combination that will permit him to reach a ***higher*** indifference curve. He will always move down the market constraint when it is flatter than his indifference curve, and up if the market constraint is steeper.

Crusoe in a Money Economy

As far as the condition for maximization of consumer satisfaction is concerned, moving from a barter economy to a money income economy changes little. **Exhibit A4** shows this point. Initially, the price of fish is $1, and the price of breadfruit is $2. The market therefore permits an exchange of two fish for one breadfruit, just as was the case in Exhibit A3. In Exhibit A4, we assume that Crusoe has a fixed money income of $16. At this level of income, he confronts the same market constraint (usually called a **budget constraint** in an economy with money) as in Exhibit A3. Given the product prices and his income, Crusoe can choose to consume sixteen fish, or eight breadfruit, or any combination indicated by a line (the budget constraint) connecting these two points. Given his preferences, Crusoe will again choose the combination of eight fish and four breadfruit if he wishes to maximize his level of satisfaction. As was true for the barter economy, when Crusoe maximizes his satisfaction (moves to the highest attainable indifference curve), the rate at which he is willing to exchange fish for breadfruit will just equal the rate at which the market will permit him to exchange the two goods. Stated in more technical terms, when his level of satisfaction is at a maximum, Crusoe's marginal rate of substitution of fish for breadfruit, as indicated by the slope of the indifference curve at E_1, will just equal the price ratio (P_F/P_b, which is also the slope of the budget constraint).

What will happen if the price of fish increases? Exhibit A4 also answers this question. Since the price of breadfruit and Crusoe's *money* income are constant, a higher fish price will have two effects. First. it will make Crusoe poorer, even though his money income will be unchanged. His budget constraint will turn clockwise around point A, illustrating that his consumption options are now more limited—that is, his real income has declined. Second, the budget line will be steeper, indicating that a larger number of breadfruit must now be sacrificed to obtain an additional unit of fish. It will no longer be possible for Crusoe to attain indifference curve i_3 The best he can do is indifference curve i_1, which he can attain by choosing the bundle of five fish and three breadfruit.

EXHIBIT A4
Consumer Maximization in a Money Economy



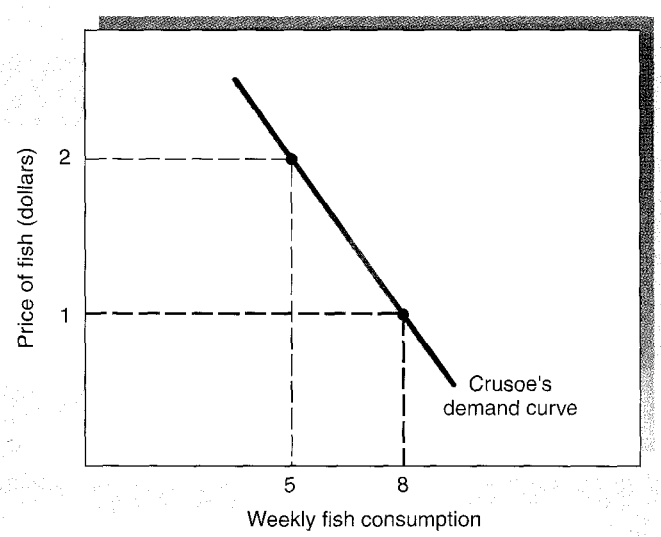
Suppose that Crusoe's income is $16 per day, the price of fish (P_F) is $1, and the price of breadfruit (P_b) is $2. Thus, Crusoe confronts exactly the same price ratio and budget constraint as in Exhibit A3. Assuming that his preferences are unchanged, he will again maximize his satisfaction by choosing to consume eight fish and four breadfruit. What will happen if the price of fish rises to $2? Crusoe's consumption opportunities will be reduced. His budget Constraint will turn clockwise around point *A*, reflecting the higher price of fish. His fish consumption will decline to five units. (Note: Because Crusoe's real income has been reduced, his consumption of breadfruit will also decline.)

Using the information supplied by Exhibit A4, we can now locate two points on Crusoe's demand curve for fish. When the price of fish was $1, Crusoe chose eight fish; when the price rose to $2, Crusoe reduced his consumption to five (see **Exhibit A5**). Of course, other points on Crusoe's demand curve could also be located if we considered other prices for fish.

The demand curve of Exhibit A5 is constructed on the assumption that the price of breadfruit remains $2 and that Crusoe's money income remains constant at $16. If either of these factors were to change, the entire demand curve for fish, shown by Exhibit A5. would shift.

The indifference curve is a useful way to show how a person with a fixed budget chooses between two goods. In the real world, of course, people have hundreds, or even thousands, of goods to choose from, and the doubling of only one price usually has a small impact on a person's overall consumption and satisfaction possibilities. In our

EXHIBIT A5
Crusoe's Demand for Fish



As Exhibit A4 shows, when the price of fish is $1, Crusoe chooses eight units. When the price of fish increases to $2, he reduces his consumption to five units. This gives us two points on Crusoe's demand curve for fish. Other points on the demand curve could be derived by confronting Crusoe with still other prices of fish. (Note: Crusoe's money income [$16] and the price of breadfruit [$2] are unchanged in this analysis.)

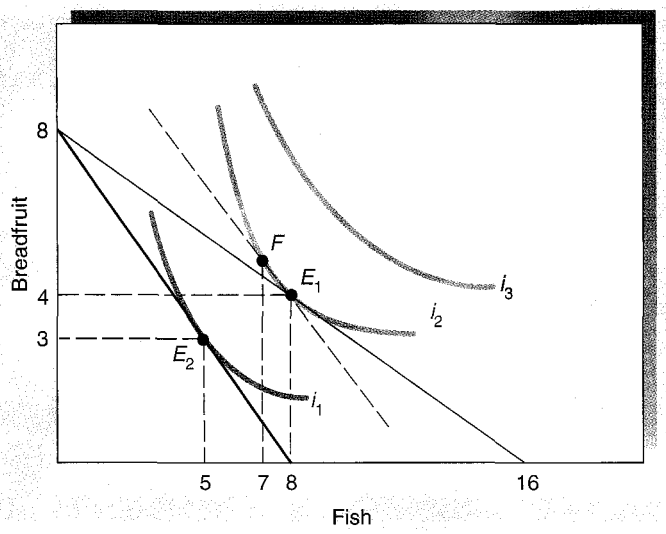


EXHIBIT A6
Income and Substitution Effects of a Price Change

Here we break down Crusoe's response to the rise in the price of fish from $1 to $2 (see Exhibit A4) into the substitution and income effects. The move from E_1 to F shows the substitution effect, whereas the move from F to E_2 reflects the income effect.

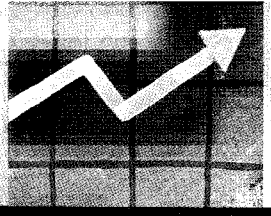
simple example, the twofold increase in the price of fish makes Crusoe much worse off, because he spends a large portion of his budget on the item.

Income and Substitution Effects of a Price Change

In the text, we indicated that, when the price of a product rises, the amount consumed will change as a result of both an *income effect* and a *substitution effect.* Indifference curve analysis can be used to separate these two effects. **Exhibit A6** is similar to Exhibit A4. Both exhibits show Crusoe's response to an increase in the price of fish from $1 to $2 when money income ($16) and the price of breadfruit ($2) are held constant. Exhibit A6, however, breaks down his total response into the substitution effect and the income effect. The reduction in the consumption of fish solely because of the substitution (price) effect, holding Crusoe's real income (level of utility) constant, can be found by constructing a line tangent to Crusoe's original indifference curve (i_2), and having a slope indicating the higher price of fish. This line (the broken line in Exhibit A6), which is parallel to Crusoe's actual budget constraint (the line containing point E_2), reflects the higher price of fish. It is tangent to the original indifference curve i_2 so Crusoe's real income is held constant. As this line indicates, Crusoe's consumption of fish would fall from eight to seven, due strictly to the fact that fish are now more expensive. This move from E_1 to F is a pure substitution effect.

Real income, though, has actually been reduced. As a result, Crusoe will be unable to attain point F on indifference curve i_2. The best he can attain is point E_2, which decreases his consumption of fish by another two units to five. Since the broken line containing F and the budget constraint containing E_1 are parallel, the relative price of fish and breadfruit is held constant as Crusoe moves from F to E_2. This move from F to E_2 is thus a pure income effect. *(Note:* Because the consumption of both goods drops in this move, when income falls but the prices do not change, both goods must be normal goods.) This reduction in the consumption of fish (and breadfruit) in the move from F to E_2 is due entirely to the decline in Crusoe's real income.

Indifference curve analysis highlights the assumptions and considerations that enter into consumer decisions. The logic of the proof that there is an inverse relationship between the price and the amount demanded is both elegant and reassuring. It is elegant because of the internal consistency of the logic and the precision of the analysis. It is reassuring because it conforms with our expectations, which are based on the central postulate of economics—that incentives matter in a predictable way.



CHAPTER 8

Costs and the Supply of Goods

From the standpoint of society as a whole, the "cost" of anything is the value that it has in alternative uses.

—*Thomas Sowell*[1]

Chapter Focus

- **Why are business firms used by societies everywhere to organize production?**
- **How are firms organized in market economies?**
- **What are explicit and implicit costs, and how do they guide the behavior of the firm?**
- **How does economic profit differ from accounting profit? Why is this difference important?**
- **How will increases in output influence the firm's costs in the short run? How will costs vary with output in the long run?**
- **What are the major factors that would cause the firm's cost curves to shift?**

[1]Thomas Sowell, *Basic Economics* (New York: Basic Books, 2000), 10.

Demand and supply interact to determine the market price of a product. In the preceding chapter, we showed that the demand for a product reflects the strength of consumer desire for that product. In this chapter, we will focus on the cost of production. The resources needed to produce one good could be used to produce other goods instead. As Thomas Sowell asserts in the quotation that begins this chapter, the cost to society of anything is the value that it has in alternative uses. The market for resources makes that cost clear to producers. The maker of soccer balls, for example, must compete against producers of other goods when purchasing the machines, materials, and labor needed to produce the balls. In turn, the firm incurs costs as these resources are purchased.

Costs carry an important message: they tell producers the value of the resources in the production of other things. Further, the message comes with an important incentive attached. If the per-unit cost of producing a good exceeds its price, producers will suffer losses. Under these conditions, they are unlikely to continue supplying the good because doing so will reduce their income and wealth. Thus, supply and the cost of production are closely linked. For example, a producer who faces a cost of $1,500 to produce a high-quality TV set is unlikely to continue supplying the sets for very long if their market price is $1,000. In the long run, firms will continue to supply a good only if they can sell it for a price that covers their per-unit cost.

In this chapter, we will lay the foundation for a detailed investigation of the links between costs, business output, and market supply. What do economists count as costs, and why? How do costs guide the owners and managers of firms in a market economy? Why are costs important to managers, even when they personally do not pay those costs? We will discuss these and related questions in this chapter. ■

THE ORGANIZATION OF THE BUSINESS FIRM

The business firm is an entity designed to organize raw materials, labor, and machines with the goal of producing goods and/or services. Firms (1) purchase productive resources from households and other firms, (2) transform them into a different commodity, and (3) sell the transformed product or service to consumers.

Every society, no matter what type of economy it has, relies on business firms to organize resources and transform them into products. In market economies, most business firms choose their own price, output level, and methods of production. They reap the benefits of sales revenues, but they also must pay the costs of the resources they use. In socialist countries, governments often set the selling prices of goods and services and constrain the actions of business firms in various other ways. Firms typically do not pay all their bills from their revenues, and they are often not allowed to keep revenues that exceed costs. In any case, the business firm is the entity used to organize production in capitalist and socialist economies alike. In this chapter, we will focus on the organization and behavior of firms in a market economy.

Incentives, Cooperation, and the Nature of the Firm

In capitalist countries, most firms are privately owned. Owners risk their wealth on the success of the business. If the firm is successful and earns profits, these financial gains go to the owners. Conversely, if the firm suffers losses, the owners must bear the consequences. Because the owners receive what remains after the revenue of the firm is used to pay the contractual costs, they are called **residual claimants.**

In a market economy, the property right of owners to the residual income of the firm plays a very important role: it provides owners with a strong incentive to organize and structure their business in a manner that will keep the cost of producing output low (relative to its value). The wealth of these residual claimants is directly influenced by the success or failure of the firm. Thus, they have both the authority and a strong incentive to see that resources under their direction are used efficiently and directed toward production of goods that are valued more highly than their costs.

Residual claimants
Individuals who personally receive the excess, if any, of revenues over costs. Residual claimants gain if the firm's costs are reduced or revenues increase.

Team production
A production process in which employees work together under the supervision of the owner or the owner's representative.

There are two ways of organizing productive activity: contracting and **team production,** in which workers are hired by a firm to work together under the supervision of the owner, or the owner's representative — a manager. Most business firms use both contracting and team production.

In principle, all production could be accomplished solely through contracting. For example, a builder might have a house built by contracting with one person to pour the concrete, another to construct the wooden part of the house, a third to install the roofing, a fourth to do the electrical wiring, and so on. No employees would have to be involved in such a project. More commonly though, goods and services are produced with some combination of contracting and the use of team production by employees of a firm.

Why do firms use team production? If contracting alone is used to produce something, the producer must, for each project, (1) determine what needs to be produced and how, given the circumstances, current technology, and prices, (2) search out reliable suppliers, and (3) negotiate and enforce the contracts. The entrepreneur who wants to produce by this method must have specialized knowledge in a variety of areas and must devote a great deal of time and effort to the planning and contracting processes. Not many people have the expertise or the time to perform all these tasks by themselves except on a small scale. Team production for certain tasks can be more practical and less costly.

Accordingly, a builder with multiple projects is likely to hire knowledgeable, experienced workers to plan the construction process, purchase materials, and build the structures. The builder will then contract with others to obtain materials and more specialized labor services.

The firm can reduce many of the transaction costs associated with contracting by using team production. Team production, however, comes with another set of problems. Team members — the employees working for the firm — must be monitored and given incentives to avoid shirking, or working at less than the expected rate of productivity. Taking long work breaks, paying more attention to their own convenience than to work results, and wasting time when diligence is called for are examples of **shirking**. A worker will shirk more when the costs of doing so are shifted to other team members, including the owners of the firm. Hired managers, even including those at the top, must be monitored and given incentives to avoid shirking.

Shirking
Working at less than the expected rate of productivity, which reduces output. Shirking is more likely when workers are not monitored, so that the cost of lower output falls on others.

Imperfect monitoring and imperfect incentives are always a problem with team production. It is part of a larger class of what economists call **principal-agent problems** A person taking a car to an auto mechanic confronts this problem. The mechanic wants to get the job done quickly and make as much money on it as possible. The car owner wants to get the job done quickly also, but wants the problem fixed in a lasting way, at the lowest possible cost. Because the mechanic typically knows far more about the job than the customer, it is hard for the customer to monitor the mechanic's work. There is a possibility, therefore, that the mechanic may charge a large amount for a "quick fix" that will not last.

Principal-agent problem
The incentive problem that occurs when the purchaser of services (the principal) lacks full information about the circumstances faced by the seller (the agent) and cannot know how well the agent performs the purchased services. The agent may to some extent work toward objectives other than those sought by the principal paying for the service.

The owner of a firm is in a similar situation. It is often difficult to monitor the performance of individual employees and motivate them in a way that will encourage high productivity. Nonetheless, the ability of the firm to use resources effectively and succeed in a competitive market depends crucially upon resolving these problems. To keep costs low and the value of output high, a firm must discover and use an incentive structure that motivates managers and workers, and discourage5 shirking. The problem extends all the way to the top.

Even top-level executives hired to manage a firm do not have the same objectives as owners — who care mainly about profit maximization — unless, of course, the managers are the owners. So the judgments of executives, too, are influenced by what is in their personal best interests. They want perks, personal job security, and other benefits that may not be consistent with profit maximization for the firm. The problem becomes more serious as firms grow larger and acquire more managers and employees. Ultimately, it is the job of the owners, as residual claimants, to develop an incentive structure to minimize the principal-agent problem. For the owner, the saying "the buck stops here" always applies.

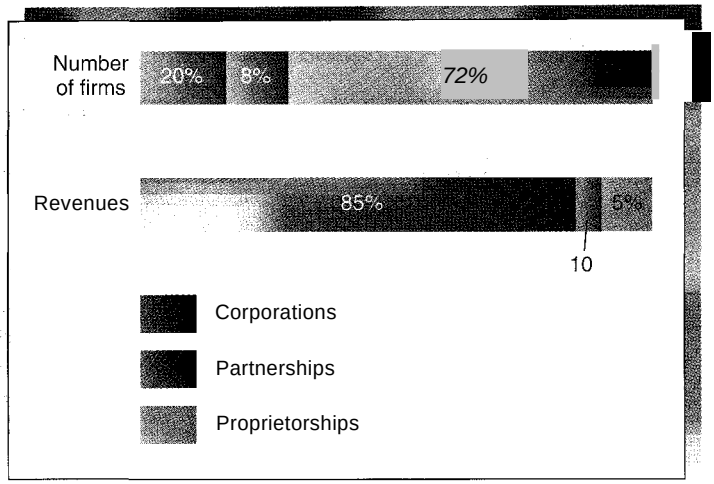


EXHIBIT 1
How Business Firms Are Organized

Nearly three out of every four firms are proprietorships, but only 5 percent of all business revenue is generated by proprietorships. Corporations account for only one out of every five firms, but generate 85 percent of all revenues.

Source: Statistical Abstract *of* the United States, *2003,* Table *732.* (Data are for *2000.)*

Three Types of Business Firms

Business firms can be organized in one of three ways: as a proprietorship, a partnership, or a corporation. The structure chosen determines how the owners share the risks and liabilities of the firm and how they participate in making decisions.

A **proprietorship** is a business firm that is owned by a single individual who is fully liable for the debts of the firm. In addition to assuming the responsibilities of ownership, the proprietor often works directly for the firm, providing managerial and other labor services. Many small businesses, including neighborhood grocery stores, barbershops, and farms, are business proprietorships. As **Exhibit 1** shows. proprietorships account for 72 percent of the business firms in the United States. Because most proprietorships are small, however, they account for less than 5 percent of all business revenues.

Proprietorship
A business firm owned by an individual who possesses the ownership right to the firm's profits and is personally liable for the firm's debts.

A **partnership** consists of two or more persons who are co-owners of a business firm. The partners share risks and responsibilities in an agreed-upon manner. There is no difference between a proprietorship and a partnership in terms of owner liability. In both cases, the owners are fully liable for all business debts incurred by the firm. Many law, medical, and accounting firms are organized along partnership lines. However, this form of business structure accounts for only 8 percent of the total number of firms and 10 percent of all business revenues.

Partnership
A business firm owned by two or more individuals who possess ownership rights to the firm's profits and are personally liable for the debts of the firm.

The business firms that are **corporations** account for more than 8.5 percent of total business revenue, even though they constitute only 20 percent of all firms. What accounts for the attractiveness of this business structure? From its start. by an Act of the British Parliament in 1862, the corporation, or "joint stock company," as it is also called, grew in importance for two main reasons. First, although the stockholders of the corporation are the legal owners, their liability is limited to the value of their shares of the corporation. If a Corporation owes you money, you cannot directly sue the stockholders. Of course, you can sue the corporation. However, if a corporation goes bankrupt, you and others to whom the firm owes money may simply be out of luck. This limited liability makes it possible for corporations to attract investment funds from a large number of "owners" who do not participate in the day-to-day management of the firm.

Corporation
A business firm owned by shareholders who possess ownership rights to the firm's profits, but whose liability is limited to the amount of their investment in the firm.

Second, ownership can easily be transferred under the corporate structure. The shares, or ownership rights, of an owner who dies can be sold by the heirs to another owner without disrupting the business firm. Because of this, the corporation is an ongoing concern. Similarly, stockholders who become unhappy with the way a corporation is run can bail out merely by selling their stock.

The stockholders of large corporations, through their elected board of directors, usually hire managers — trained and experienced experts — to operate the firm. Will the managers operate the firm efficiently and satisfy customers? Offering consumers value at

a low cost is the ticket to profitability. In an owner-managed firm, the owner's property **right to** the residual income provides a strong incentive both to reduce costs and to please consumers. This is not necessarily so for the managers of a corporation.

For a large corporation with many stockholders—millions in some cases—the situation is complex. Stockholders own the residual income (profits), but professional managers operate the firm. Stockholders want managers to cut costs while increasing output and revenues, but managers might want high salaries for themselves, large offices, first-class travel, and other expensive perks. They might also prefer the power and prestige associated with expanding the business (by taking over another firm, perhaps), even if it might reduce profitability. Can the stockholders control the actions of managers and direct them toward the pursuit of profitability? Direct control by many stockholders is unlikely. Few of them own enough shares to give them the incentive or information they would need to exercise direct control. Most find it too expensive even to attend the annual shareholders' meeting, much less to monitor managers closely. Instead, stockholders elect a board of directors, which in turn appoints the company's high-level managers. As the following section describes, however, not all is lost when stockholders don't directly run the company themselves. Internal corporate policies and competition for control of the firm by outsiders help mitigate principal-agent problems when managers run the company instead.

COSTS, COMPETITION, AND THE CORPORATION

Three major factors in a market economy promote cost efficiency and customer service within the corporation, limiting the power of corporate managers to shirk on their duties to shareholders:

1. Competition among firms for investment funds and customers. Even without direct control of their corporation, stockholders (and the investment advisers, pension fund managers, and others hired to help them) have an incentive to monitor the corporation's management in order to anticipate problems and search for constructive changes. Investors who are the first to spot a profitable new management strategy can buy stock early, before others realize the opportunity, and bid the price up. **A** rising stock price is both a signal of approval to good managers and an incentive to manage the corporation well. Conversely, when managers' decisions are contrary to the interests of stockholders, the opposite occurs: the stock price will fall. Knowing that stock prices can plummet motivates stockholders (along with professional managers of stock holdings) to be the first to spot a problem with the company. This way they can dump their stock before everyone else does, and thus prior to the share price falling. Some investment firms even specialize in selling shares "short" (writing contracts to deliver shares of the stock later, when they expect the stock price to be lower than its current level) if they believe a company is overvalued on the market. So managers get constant feedback via stock price changes, which can be just as important as current profits to stockholders and boards of directors.

Similarly, consumers have an incentive to monitor the quality and price of the firm's output. No one forces them to buy the corporation's product, so if other firms supply superior products or offer lower prices, consumers can take their business elsewhere. Because investors are free to buy and sell the company's shares and customers to buy its products or those of other firms, the ability of managers to benefit personally at the expense of either customers or stockholders is limited. While some managers are still able to enrich themselves at stockholders' expense, at least temporarily, the recent cases of corporations like Tyco and Enron show that their actions tend to catch up with them.

2. Compensation and management incentives. The compensation of managers can be structured to bring the interests of managers more into harmony with those of shareholders. Corporations usually tie the compensation of managers to the market success of the business. The salary increases and bonuses of most high-level managers are directly related to the firm's profitability and the price of its shares.

How important are these incentives? In recent years, salaries have constituted only about 10 percent of the compensation of chief executive officers (CEOs). The other 90 percent has been in the form of bonuses, often stock awards and stock options (the right to buy shares at a certain price). Both forms of payment have a "vesting period," which means they will be granted only if the CEO stays with the firm for a certain amount of time and meets specific goals. Policies like these encourage corporate managers to maximize the the firm's profits—and, not coincidentally, the value of its shares—both strongly in the interest of shareholders. Managers who develop a good track record of adding value to the firm gain not only more pay and greater job security, but also better job offers should they later decide to switch firms.

© BONNIE KAMIN/PHOTOEDIT

Corporate managers are more likely to serve the interests of customers and stockholder because of the following: competition for investment funds and consumer sales, the threat of a takeover, and managerial compensation packages based on stock and stock options.

Incentive pay brings with it another unwanted incentive for managers, though: the incentive to gain personally by manipulating the firm's accounting records to make its financial performance look better than it really is. (See the accompanying Applications in Economics feature: "Cooking the Books.") However, stockholders and portfolio managers, and especially investors who specialize in selling stocks short when they think the firm is overvalued, will be scrutinizing the records to detect phony accounting designed to mislead investors.

3. The threat of corporate takeover. Managers who do not serve the interests of their shareholders leave the firm vulnerable to a takeover. This is a move by an outside person or group, noticing the bad management, to gain control of the firm. As we previously noted, shareholders who lose confidence in the firm's management can exit the arrangement by selling their shares. When a significant number of shareholders follow this course of action, the market value of the firm's stock will decline. This will make it an attractive prospect for takeover specialists shopping for a poorly run business, the value of which could be substantially increased by a new and better management team.

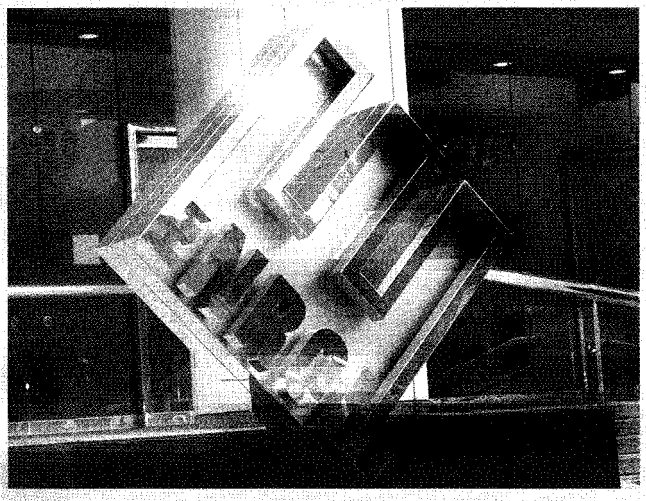
Consider a firm currently earning $1.50 per share. At present, the market value of the firm's stock is $15 per share. If the firm's earnings are low because the current management team is pursuing its own objectives at the expense of profitability. then a corporate takeover could lead to substantial gain for someone. Suppose outsiders believe the? can restructure the firm, improve the management. and double the firm's earnings. They then "tender" a takeover bid—make an offer to buy the shareholders' stock or persuade them to pressure the board of directors to sell out. Suppose they offer S20 per share. This is more than the current market value of the firm, so shareholders and the board will be tempted to accept the offer and gain wealth from the sale. If the takeover team gains control of the firm. improves its performance, and increases its earnings to $3 per share, then the stock value of the firm will rise accordingly (to $30 per share).

Of course, the current managers have an incentive to resist the takeover. After all, they are likely to lose their jobs if the potential new owners are successful. Unfortunately for them. though, the shareholders or their board of directors will ultimately decide whether or not to accept the offer. The takeover threat helps keep current managers from straying too far from a strategy to maximize profit. Some corporate managers have instituted policies to help defend their firms against takeovers. However, limiting the power of shareholders this way can significantly lower the stock price of a badly managed

APPLICATIONS IN ECONOMICS

Cooking the Books: How the Market Responds to Criminal Behavior

"Cooking the books" refers to the practice of using accounting procedures in an intentionally misleading way. By cooking the books, a corporation's profits—or at least what appear to be its profits—can be increased. Stockholders are always looking for shares of high-profit firms, so a recent record of high profits will increase the demand for the corporation's stock shares and drive up their price. Cooking the books can be artificially encouraging to shareholders in the short term, but it can conceal problems and even lead to disaster in the long term.



© TIM JOHNSON/BLOOMBERG NEWS/LANDOV

Enron, the energy-trading company, is a dramatic example. High-level executives in the firm used misleading accounting procedures to make Enron look highly profitable when, in fact, it was not. Moreover, Arthur Andersen, a highly respected accounting firm, audited Enron's books and certified their accuracy, knowing there were irregularities. As a result, Enron's stock price climbed, along with the value of its managers' stock options—even though the company was really on its way to bankruptcy. These executives were then able to cash in on their stock options to enrich themselves while the company lay on the path to ruin.

When Enron made some of its accounting procedures public in November 2001, market participants took quick action, selling their shares. The stock price then plummeted. In light of its true financial position, Enron filed for bankruptcy less than a month later. At the beginning of the year, its stock had been worth $60 billion. By the end of the year, it was nearly worthless. By July 2004, Enron was employing about 9,000 workers, down from 32,000 at its peak, and its bankruptcy plan had been approved by a federal judge. The Justice Department had filed criminal charges against thirty-one members of Enron's management, including its chief executive.

Arthur Andersen, Enron's accounting firm, faced similar consequences. Many of the auditing firm's clients quickly stopped doing business with it. One of its partners was subsequently convicted of obstructing justice, and the firm announced it would no longer audit public companies in the U.S. Its reputation as a trusted accounting company was ruined, and criminal charges were brought against several Arthur Andersen executives.

company.[2] Moreover, managers thought to be doing a bad job will have to make it extraordinarily difficult for a takeover to occur in order to avoid being ousted.

How Well Does the Corporate Structure Work?

Perhaps history provides the best answer to this question. If the corporate structure were not an effective form of business organization, it would not have continued to survive, nor would it be so prevalent today. Rival forms of business organization, including proprietorships, partnerships, consumer cooperatives, employee ownership, and mutually owned companies can and do compete in the marketplace for investment funds and customers. In certain industries, some of these alternative forms of business organization are dominant. Nonetheless, in most industries, the corporate structure is the dominant form of business organization (see Exhibit 1). This is strong evidence that, despite its defects, the corporation is generally a cost-efficient, consumer-sensitive form of organization.

*Economists Paul Gompers, Joy Ishii, and Andrew Metrick, in "Corporate Governance and Equity Prices" (NBER working paper 8449, released Augurt 2001), find that "firms with weaker shareholder rights earned significantly lower returns, were valued lower, had poorer operating performance, and engaged in greater capital expenditure and takeover activity." However, some of the same actions that favor managers—large "golden parachute" retirement plans, for example—can also be guarantees that increase the ability of a company's directors to hire or retain good managers, as shown by Dino Falaschetti, "Golden Parachutes: Credible Commitments or Evidence of Shirking?," *Journal of Corporate Finance* 8 (2002): 150–78.

THE ECONOMIC ROLE OF COSTS

Consumers would like to have more economic goods, but resources to produce them are scarce. How much of each desired good should be produced? Every economic system must balance consumers' competing desires. When decisions are made in the political arena, the budget process performs this balancing function. Legislators, a central planning committee, or a monarch decide which goods will be produced and which will be forgone. Taxes and budgets are set accordingly.

In a market economy, though, consumer demand and production costs are central to performing this balancing function. ***The demand for a product represents the voice of consumers instructing firms to produce the good. On the other hand, a firm's costs represent the desire of consumers not to sacrifice goods that could be produced if the same resources were employed elsewhere.*** A profit-seeking firm will try to produce only those units of output buyers are willing to pay full cost for. Proper measurement and interpretation of costs by the firm are critical to both the firm's profitability and the efficient use of resources.

Calculating Economic Costs and Profits

We can reasonably assume that business firms in a market setting, regardless of their size, are concerned first with profit. Profit is simply the firm's total revenue minus its total costs. But to state profit correctly, costs must be measured properly. Most people, including some who are in business, think of costs as amounts paid for raw materials, labor, machines, and similar inputs. However, this concept of cost, which stems from accounting procedures, excludes some important components of the firm's costs. When cost is miscalculated, so, too, is profit because it is merely revenue minus cost. Bad economic decisions can result from miscalculating cost and profit.

The key to understanding the economist's concept of profit is to remember the idea of opportunity cost—the highest valued alternative forgone by the resource owner when the resource is used. These costs may either be explicit or implicit. **Explicit** costs result when the firm makes a monetary payment to resource owners. Money wages, interest, and rental payments are a measure of what the firm gives up to employ the services of labor and capital resources. These are relatively easy to track. But firms also incur **implicit** costs—those associated with the use of resources owned by the firm. For example, the owners of small proprietorships often work for their own businesses, for little or no pay. These businesses incur an implicit cost—an opportunity cost—associated with the use of this resource (the owners' labor services). The highest valued alternative forgone in this case is the maximum amount of money the owners could have made doing something else. The total **cost** of production is the sum of these explicit and implicit costs incurred by the employment of all resources involved in the production process.

Explicit costs
Payments by a firm to purchase the services of productive resources.

Implicit costs
The opportunity costs associated with a firm's use of resources that it owns. These costs do not involve a direct money payment. Examples include wage income and interest forgone by the owner of a firm who also provides labor services and equity capital to the firm.

Total cost
The costs, both explicit and implicit, of all the resources used by the firm. Total cost includes a normal rate of return for the firm's equity capital.

Accounting statements generally omit the implicit cost of equity capital—the cost of funds supplied by the firm's owners. If a firm borrows financial capital from a bank or other private source, it will have to pay interest. Accountants properly record this interest sxpense as a cost. In contrast, when the firm acquires financial capital by issuing **stock.** accountants don't record this as an expense. Essentially. this is because the stockholders *are* the firm's owners. Either way, acquiring capital has an opportunity cost. People who supply capital to a firm expect to earn at least a normal rate of return—a return comparable to what they could have earned if they had chosen other investment opportunities. Banks will demand interest payments, for example. Stockholders will demand a share of rhe company's profits, or dividends. Or they might expect the price of the stock they purchase to rise measurably, thereby enhancing their wealth.

When calculating the normal rate of return, economists use the normal return on financial capital as a basis for determining the implicit opportunity cost **of equity** capital. If the normal rate of return on financial capital is 10 percent, equity investors will refuse funds to firms that persistently fail to earn a 10 percent rate of return on capital assets.

Opportunity cost of equity capital
The rate of return that must be earned by investors to induce them to supply financial capital to the firm.

How Do Economic and Accounting Profit Differ?

Economic profit
The difference between the firm's total revenues and its total costs. including both the explicit and implicit cost components.

Normal profit rate
Zero economic profit, providing just the competitive rate of return on the capital (and labor) of owners. An above-normal profit will draw more entry into the market, whereas a below-normal profit will lead to an exit of investors and capital.

Economists include both explicit and implicit costs when they measure total cost. **Economic profit** is total revenues minus total costs, including both the explicit and implicit cost components. Economic profit will be positive only if the earnings of the business exceed the opportunity cost of all the resources used by the firm, *including the opportunity cost of assets owned by the firm and any unpaid labor services supplied by the owner.* In contrast, economic losses result when the earnings of the firm are insufficient to cover explicit and implicit costs. That is why the **normal profit rate** is zero economic profit, yielding just the competitive rate of return on the capital (and labor) of owners. A higher rate would draw more competitors and their investors into the market; a lower rate would cause competitors and their investors to exit the market.

Remember, zero economic profits do not imply that the firm is about to go out of business. On the contrary, they indicate that the owners are receiving exactly the normal profit rate, or the competitive market rate of return on their investment. They are earning no more and no less than they could earn elsewhere on what they use in the firm.

Accounting profits
The sales revenues minus the expenses of a firm over a designated time period, usually one year. Accounting profits typically make allowances for changes in the firm's inventories and depreciation of its assets. No allowance is made, however, for the opportunity cost of the equity capital of the firm's owners, or other implicit costs.

Whenever accounting procedures omit implicit costs, like those associated with owner-provided labor services or capital, the firm's opportunity costs of production will be understated. This understatement of cost leads to an overstatement of profits. Therefore, the **accounting profits** of a firm are generally greater than the firm's economic profits (see the Applications in Economics feature on accounting costs). For most large corporations, though, omitting the implicit costs of services provided by an owner isn't an issue. In this case, the accounting profits approximate the returns to the firm's equity capital. High accounting profits (measured as a rate of return on a firm's assets), relative to those of other firms, suggest that a firm is earning an economic profit. Correspondingly, a low rate of accounting profit implies economic losses. Either positive or negative economic profits, of course, call for a change in output. Such a change, however, will take time.

SHORT-RUN AND LONG-RUN TIME PERIODS

Short run (in production)
A time period so short that a firm is unable to vary some of its factors of production. The firm's plant size typically cannot be altered in the short run.

A firm cannot instantly adjust its output. Time plays an important role in the production process. All of a firm's resources can be expanded (or contracted) over time, but for specialized or heavy equipment, expanding (and contracting) availability quickly is likely to be very expensive or even impossible. Economists often speak of the **short run** as a time period so short that the firm is unable to alter its present plant size. In the short run, the firm is typically stuck with its existing plant and heavy equipment. These assets are "fixed" for a given time period, in other words. The firm can alter output, however, by applying larger or smaller amounts of "variable" resources, like labor and raw materials. In this way, the existing plant capacity can be used more or less intensively in the short run.

How long is the short run? ***The short run is that period of time during which at least one factor of production, usually the size of the firm's plant, cannot be changed.*** The length varies from industry to industry. A trucking firm might be able to hire more drivers and buy or rent more trucks and double its hauling capacity in a few months. In other industries, particularly those that use assembly lines and mass-production techniques (for example, the automotive factory supplying trucks), increasing production capacity might take a year or even several years.

Long run (in production)
A time period long enough to allow the firm to vary all of its factors of production.

The **long run** is a time period long enough for existing firms to alter the size of their plants and for new firms to enter (or exit) the market. All of the firm's resources are variable in the long run. In the long run, firms can expand their output by increasing the sizes of their plants—perhaps by adding on to them or by constructing entirely new facilities.

An example might help you understand the distinction between the short- and long-run time periods: If a battery manufacturer hired 200 additional workers and ordered more raw materials in order to squeeze more production out of its existing plant, it would be making a short-run adjustment. In contrast, if the manufacturer built an additional plant (or expanded the size of its current facility) and installed additional heavy equipment, it would be making a long-run adjustment.

CATEGORIES OF COSTS

To describe various aspects of the firm's costs and the relationship of those costs to the level of output, even without detailed knowledge of the production process, we need some definitions. We have emphasized that in the short run some of a firm's factors of production, such as the size of its plant, will be fixed. Other productive resources will be variable. In the short run, then, we can break the firm's costs into these two categories—fixed and variable. Each category of costs behaves differently. Seeing that behavior graphically will help us understand how the profit-maximizing le\el of the firm's output is determined. It will be important to distinguish between a firm's total costs and its per-unit costs, called "average" costs.

Each of the firm's fixed costs, and their sum. called total **fixed cost** (*TFC*). will remain unchanged when output rises or falls in the short run. For example. a firm's insurance premiums, its property taxes, and, most significantly. the **opportunity** cost **of** using its fixed assets will be present whether the firm produces a large or small amount of output. These costs will not vary with output. The) are "fixed"as long as the firm remains in business. Fixed costs will be present at all levels of output. including zero output. They can be avoided only if the firm goes out of business.

Total fixed cost
The sum of the costs that do not vary with output. They will be incurred as long as a firm continues in business and the assets have alternative uses.

What will happen to average fixed **cost** ***(AFC),*** fixed costs per unit, as output expands? Remember that the firm's fixed cost **will** be the same whether output is 1, 100, or 1,000. The *AFC* is simply fixed cost divided by output. It's the firm's fixed cost spread across all of the units produced. As output increases, *AFC* declines because the fixed cost will be spread over more and more units (see part a of **Exhibit 2).**

Average fixed cost
Total fixed cost divided by the number of units produced It always declines as output increases.

APPLICATIONS IN ECONOMICS

Economic and Accounting Costs: A Hypothetical Example

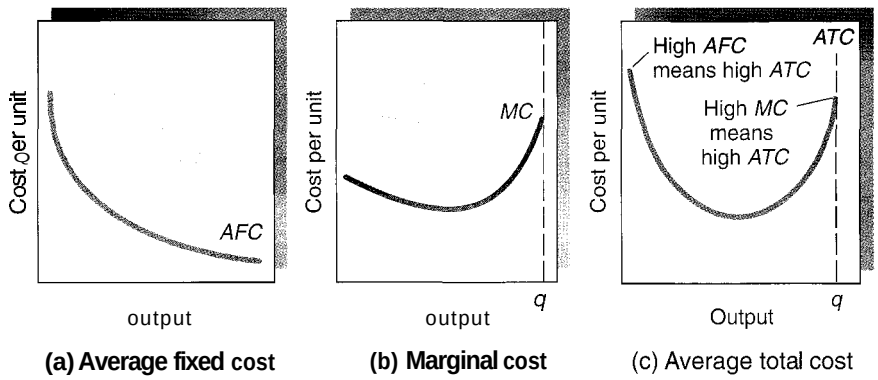
The revenue-cost statement for a corner grocery store owned and operated by Terry Smith is presented here.

TOTAL REVENUE	
Sales (groceries)	$170,000
Costs (explicit)	
Groceries, wholesale	$76,000
Utilities	4,000
Taxes	6,000
Advertising	2,000
Labor services (employees)	12,000
Total (explicit)costs	$100,000
Net (accounting)profit	$70,000
Additional (implicit)costs	
Interest (personal investment)	$ 7,000
Rent (Terry's building)	18,000
Salary (Terry's labor)	50,000
Total (implicit)costs	$75,000
TOTAL EXPLICIT AND IMPLICIT COSTS	$175,000
ECONOMIC PROFIT (TOTAL REVENUE MINUS EXPLICIT AND IMPLICIT COSTS)	−$5,000

Terry works full-time as the manager, chief cashier, and janitor. Terry has $140,000 worth of refrigeration and other equipment invested in the store. Last year, his total sales were $170,000, and suppliers and employees were paid $100,000. His revenues therefore exceeded explicit costs by $70,000. This is what was recorded on the books as profit.

But did Terry really make a profit last year? Let's look at his opportunity costs and see: If Terry didn't have $140,000 of his own money invested in equipment, he could be earning 5 percent interest on the money, which would add up to $7,000 each year. Similarly, if the building he owns weren't being used as a grocery store, it could be rented to someone else for $1,500 per month. Rental income forgone is therefore $18,000 per year. In addition. because Terry is tied up working in the grocery store, a $50.000 managerial position with the local Safeway is forgone. Considering the interest, rental, and salary income that Terry had to forgo to operate the grocery store last year. his implicit costs were $75,000. This makes his total costs—both explicit and implicit—$175,000. (Recall that explicit costs were $100,000.) That's $5,000 *less* than his actual revenues of $170,000. As a result, Terry incurred an economic loss of $5,000, despite the accounting profit of $70,000 recorded on the store's books.

EXHIBIT 2
The General Characteristics of Short-Run Cost Curves



Average fixed costs (a) will be high for small rates of output, but they will always decline as output expands. Marginal cost (b) will rise sharply as the plant approaches its production capacity, *q*. As graph (c) shows, *ATC* will be a U-shaped curve because *AFC* will be high for small rates of output, and *MC* will be high as the plant's production capacity is approached.

Total variable cost
The sum of those costs that rise as output increases. Examples of variable costs are wages paid to workers and payments for raw materials.

Average variable cost
The total variable cost divided by the number of units produced.

Average total cost
Total cost divided by the number of units produced. It is sometimes called per-unit cost.

Marginal cost
The change in total cost required to produce an additional unit of output.

Some costs vary with output. For example, additional output can usually be produced by hiring more workers and buying more raw materials. The sum of those and other costs that rise as output increases are the firm's **total variable cost** (*TVC*). At any given level of output, the firm's **average variable cost** (*AVC*) is the total variable cost divided by output. It's the firm's variable cost spread across all of the units produced.

We have noted that total cost (*TC*) includes explicit and implicit costs. At the same time, the total cost of producing a good is also the sum of the fixed and variable costs at each output level. At zero output, total cost will equal total fixed cost. As output expands from zero, variable cost and fixed cost must be added to obtain total cost. **Average total cost** (*ATC*), sometimes referred to as unit cost, can be found by dividing total cost by the total number of units produced. *ATC* is also equal to the sum of the average fixed and average variable costs. It indicates the amount of revenue per unit of output needed to cover total cost.

The economic way of thinking focuses on what happens "at the margin." How much does it cost to produce an additional unit? **Marginal cost** (*MC*) is the change in total cost that results from the production of one additional unit. The profit-conscious decision maker recognizes *MC* as the addition to cost that must be covered by additional revenue if producing the marginal unit is to be profitable. In the short run, as illustrated by part (b) of Exhibit 2, *MC* will generally decline at first if output is increased, reach a minimum, and then increase. The rising *MC* simply reflects the fact that it becomes increasingly difficult to squeeze additional output from a plant as the facility's maximum capacity (the dotted line of part b of Exhibit 2) is approached. The accompanying Thumbnail Sketch summarizes how the firm's various costs are related to one another.

OUTPUT AND COSTS IN THE SHORT RUN

As a firm changes its rate of output in the short run, how will its unit cost be affected? First, look at this question intuitively. In the short run, the firm can vary its output by using its fixed plant size more (or less) intensively. Exhibit 2 shows two ways that this can result in high unit costs. First, when the output rate of a plant is small relative to its capacity, the facility is being underutilized, causing *AFC* to be high and *ATC* to be high, too. It will be costly to operate a large plant, with its high fixed costs, substantially below its production capacity. Alternatively, overutilization can also cause high unit costs. An overutilized plant will mean congestion—time spent by workers waiting for machines and similar costly delays. Requiring output beyond the least-cost, or designed, output level of a plant will lead to high *MC* and therefore to high *ATC*.

Thus, the* ATC *curve will be U-shaped, as pictured in part (c) of Exhibit 2.* ATC *will be high for both an underutilized plant and an overutilized plant. It will be high for an underutilized plant because average fixed cost will be high. It will be high for an overutilized plant because marginal cost will be high.

THUMBNAIL SKETCH

Compact Glossary on Cost

Term	Symbol	Equation	Definition
Fixed cost			Cost that is independent of the output level
Variable cost			Cost that varies with the output level
Total fixed cost	TFC		Cost of the fixed inputs (equals sum of quantity times unit price for each fixed input)
Total variable cost	*TVC*		Cost of the variable inputs (equals sum of quantity times unit price for each variable input)
Total cost	*TC*	$TC = TFC + TVC$	Cost of all inputs (equals fixed costs plus variable costs)
Marginal cost	*MC*	$MC = ATC \div \Delta q$	Change in total cost resulting from a one-unit rise in output (q) [equals the change in total cost divided by the change in output]
Average fixed cost	*AFC*	$AFC = TFC \div q$	Total fixed cost per unit of output (equals total fixed cost divided by total output)
Average variable cost	*AVC*	$AVC = TVC \div q$	Total variable cost per unit of output (equals total variable cost divided by total output)
Average total cost	*ATC*	$ATC = AFC + AVC$	Total cost per unit of output (equals average fixed cost plus average variable cost)

Diminishing Returns and Production in the Short Run

Our analysis of the changes in unit cost as the output rate rises reflects a long-established economic law. This **law of diminishing returns** states that, as more and more units of a variable factor are applied to a fixed amount of other resources, output will eventually increase by smaller and smaller amounts. Therefore, the impact on output of additional units of the variable factor will diminish. The cost per unit of adding the variable factor may be the same, but the added output per dollar spent falls. The impact on cost per unit of output is clear: when the returns to the variable factor are rising, marginal costs (the additions to total variable cost needed to add a unit of output) are falling. Similarly, when the returns to the variable factor are falling, marginal cost is rising.

Law of diminishing returns
The postulate that, as more and more units of a variable resource are combined with a fixed amount of other resources, using additional units of the variable resource will eventually increase output only at a decreasing rate. Once diminishing returns are reached, it will take successively larger amounts of the variable facto: to expand output by one unit.

The law of diminishing returns is as famous in economics as the law of gravity is in physics. It is based on common sense and real-life observation. Have you ever noticed that, as you apply a single resource more intensively, the resource eventually tends to accomplish less and less? Consider a wheat farmer who applies fertilizer (a variable resource) more and more intensively to an acre of land (a fixed factor). At some point, the application of additional 100-pound units of fertilizer will expand the wheat yield by successively smaller amounts.

Essentially, the law of diminishing returns is a constraint imposed by nature. If it were not valid, it would be possible to raise all the world's food on an acre of the best land, or even in a flowerpot. Logically, then, there would be no point in cultivating any of the less-fertile land. We would be able to increase output simply by applying another unit of labor and fertilizer to the world's most fertile flowerpot! In the real world, of course, this is not the case; the law of diminishing returns is valid and it restricts our options.

Total product
The total output of a good that is associated with each alternative utilization rate of a variable input.

Exhibit 3 illustrates the law of diminishing returns numerically. Column 1 indicates the quantity of the variable resource, labor in this example, that is combined with a specified amount of the fixed resource. Column 2 shows the **total product** that will result as the utilization rate of labor increases. Column 3 provides data on the **marginal product**, the

Marginal product
The increase in the total product resulting from a unit increase in the employment of a variable input. Mathematically, it is the ratio of the change in total product to the change in the quantity of the variable input.

EXHIBIT 3
The Law of Diminishing Returns (Hypothetical Data)

(1) Units of the Variable Resource, Labor (per day)	(2) Total Product (output)	(3) Marginal Product	(4) Average Product
0	0	—	—
1	8	8	8.0
2	20	12	10.0
3	34	14	11.3
4	46	12	11.5
5	56	10	11.2
6	64	8	10.7
7	70	6	10.0
8	74	4	9.3
9	75	1	8.3
10	73	−2	7.3

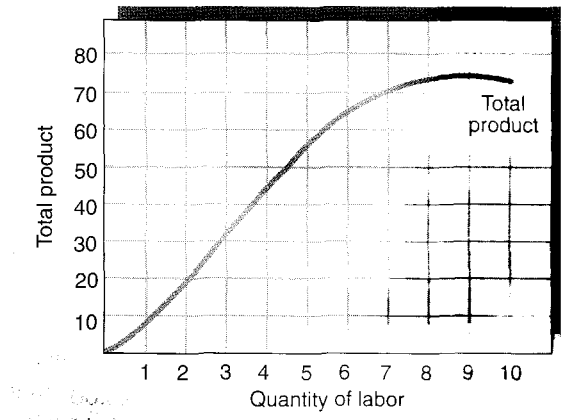
change in total output associated with each additional unit of labor. Without the application of labor, output will be zero. As additional units of labor are applied, total product (output) rises. As the first three units of labor are applied, total product increases by successively larger amounts (8, then 12, then 14). Beginning with the fourth unit, however, diminishing returns are confronted. When the fourth unit is added, marginal product—the change in the total product—declines to 12 (down from 14, when the third unit was applied). As additional units of labor are applied, marginal product continues to decline. It is increasingly difficult to squeeze a larger total product from the fixed resources (for example, plant size and equipment). Eventually, marginal product becomes negative (beginning with the tenth unit). Column 4 of Exhibit 3 provides data for the **average product** of labor, which is simply the total product divided by the units of labor applied. Note that the average product increases as long as the marginal product is greater than the average product. Whenever the marginal unit's contribution is greater than the average, it must cause the average to rise. (A good analogy would be your grade point average. If the grade you get in this course is higher than your overall grade point average, your grade point average has to go up.) Here, this is true through the first four units. The marginal product of the fifth unit of labor, though, is 10, less than the average product for the first four units of labor (11.5). Therefore, beginning with the fifth unit, the average product declines as additional labor is applied. When marginal productivity is below the average, it brings down the average product.

Average product
The total product (output) divided by the number of units of the variable input required to produce that output level.

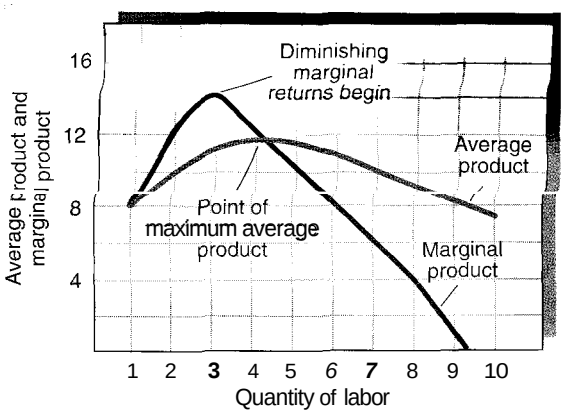
Using the data from Exhibit 3, **Exhibit 4** illustrates the law of diminishing returns graphically. Initially, the total product curve (part a) increases quite rapidly. As diminishing marginal returns are confronted (beginning with the fourth unit of labor), total product increases more slowly. Eventually, a maximum output (75) is reached with the application of the ninth unit of labor. The marginal product curve (part b) reflects the total product curve. Geometrically, marginal product is the slope—the rate of increase—of the total product curve. That slope, the marginal product, reaches its maximum here with the application of three units of labor. Beyond three units, diminishing returns are present. Eventually, at ten units of labor, the marginal product becomes negative. When marginal product becomes negative, total product is necessarily declining. The average product curve rises as long as the marginal product curve is above it, since each added unit of labor is raising the average. The average product reaches its maximum at four units of labor. Beyond that, each additional unit of labor brings down the average product, and the curve declines.

Diminishing Returns and the Shape of the Cost Curves

What impact will diminishing returns have on a firm's costs? Once a firm confronts diminishing returns, larger and larger additions of the variable factor are required to



(a) Total product curve



(b) Average and marginal product curve

EXHIBIT 4
The Law of Diminishing Returns

As units of variable input (labor) are added to a fixed input, total product will increase, first at an increasing rate and then at a declining rate (a). This will cause both marginal and average product curves (b) to rise at first and then decline. Note that the marginal product curve intersects the average product curve at its maximum (when four units of labor are used). The smooth curves indicate that labor can be increased by amounts of less than a single unit.

expand output by one unit. This will cause marginal costs (MC) to rise. As MC continues to increase, eventually it will exceed average total cost. Until that point, MC is below ATC, bringing ATC down. When MC is greater than ATC, the additional units cost more than the average, and ATC must increase. As suggested above, think about what happens when you get a grade on an exam above your current class average. Your class average goes up. What happens if a unit of above-average cost is added to output? Average total cost rises. The firm's *MC* curve therefore crosses the *ATC* curve at the *ATC*'s lowest point. For output rates beyond the minimum ***ATC,*** the rising *MC* causes *ATC* to increase.

Exhibit 5 numerically illustrates the effect of the law of diminishing returns on a firm's short-run cost curve. Here, we assume that Royal Roller Blades, Inc. combines units of a variable input with a fixed factor to produce units of output (pairs of inline skates). Columns 2, 3, and 4 indicate how the total cost schedules vary as output is expanded. Total fixed costs *(TFC)*, representing the opportunity cost of the fixed factors of production, are $50 per day at all levels of output. For the first four units of output, total variable costs (*TVC*) increase at a decreasing rate—by $15 with the production of the first unit, $10 with the production of the second unit, $9 with the third, and so on. Why? In this range, there are increasing returns to the variable input. Beginning with the fifth unit of output, however, diminishing marginal returns are present. From this point on, *TVC* and *TC* increase by successively larger amounts as output is expanded.

Columns 5 through 8 of Exhibit 5 are the average and marginal cost schedules. For small output rates, the *ATC* of producing roller blades is high, primarily because of the high *AFC*. Initially, *MC* is less than *ATC*, so *ATC* is falling. When diminishing returns set in for output rates beginning with five units, however, *MC* rises. Beginning with the sixth unit of output, *MC* exceeds AVC, causing AVC to rise. Beginning with the eighth unit of output, *MC* exceeds *ATC*, causing it also to rise. *ATC* thus reaches its minimum at seven

EXHIBIT 5
The Numerical Short-Run Cost Schedules of Royal Roller Blades, Inc.

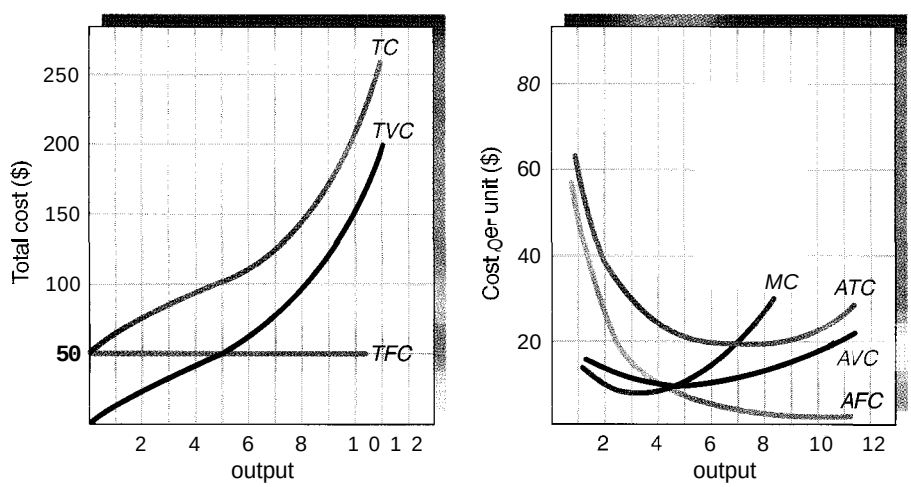
	TOTAL COST DATA (PER DAY)			AVERAGE/MARGINAL COST DATA (PER DAY)			
(1) OUTPUT PER DAY	(2) *TFC*	(3) *TVC*	(4) *TC* (2) + (3)	(5) *AFC* (2) ÷ (1)	(6) *AVC* (3) ÷ (1)	(7) *ATC* (4) ÷ (1)	(8) *MC* Δ(4) ÷ Δ(1)
0	$50	$ 0	$ 50	—	—	—	—
1	50	15	65	$50.00	$15.00	$65.00	$15
2	50	25	75	25.00	12.50	37.50	10
3	50	34	84	16.67	11.33	28.00	9
4	50	42	92	12.50	10.50	23.00	8
5	50	52	102	10.00	10.40	20.40	10
6	50	64	114	8.33	10.67	19.00	12
7	50	79	129	7.14	11.29	18.43	15
8	50	98	148	6.25	12.25	18.50	19
9	50	122	172	5.56	13.56	19.11	24
10	50	152	202	5.00	15.20	20.20	30
11	50	202	252	4.55	18.36	22.91	50

units of output. Look carefully at the data of Exhibit 5 to be sure that you fully understand the relationships among the various cost curves. Do you understand how columns 4 to 8 are derived from columns 1 to 3?

Using the numeric data of Exhibit 5, **Exhibit 6** graphically illustrates the total, the average, and the marginal cost curves. Note that the *MC* curve intersects both the AVC and *ATC curves* at their minimum points (part b). As *MC*, driven up by diminishing returns, continues to rise above *ATC*, unit costs rise higher and higher as output increases beyond seven units.

In sum, the firm's short-run cost curves reflect the law of diminishing marginal returns. Assuming that the price of the variable resource is constant, *MC* declines so long as the marginal product of the variable input is rising. This results because, in this range, smaller and smaller additions of the variable input are required to produce each extra unit of output. The situation is reversed, however, when diminishing returns are confronted. Once diminishing returns set in, more and more units of the variable factor are required to generate each additional unit of output. *MC* will rise, because the marginal product of the

EXHIBIT 6
Costs in the Short Run



(a) Total cost data (b) Average and marginal cost data

Using data from Exhibit 5, this exhibit shows the general shape of the firm's short-run total cost curves (a), and average and marginal cost curves (b). Note that when output is small (for example, two units), *ATC* will be high because the *AFC* is so high. Similarly, when output is large (for example, eleven units), per-unit cost (*ATC*) will be high because additional units will be extremely costly to produce at this point. Thus, the short-run ATC curve will be U-shaped.

variable resource is declining. Eventually, *MC* exceeds *AVC* and *ATC,* causing these costs also to rise. **A** U-shaped, short-run average total cost curve results.

OUTPUT AND COSTS IN THE LONG RUN

The short-run analysis relates costs to output *for a specific size of plant.* Firms. though, are not committed forever to their existing plants. In the long run. a **firm** can alter its plant size and all other factors of production. **All** resources used by the **firm** are variable in the long run. Thus, there are no fixed and variable cost categories in the long run.

How will the firm's choice of plant size affect per-unit production costs'? **Exhibit 7** illustrates the short-run *ATC* curves for three different plant sizes, ranging from small to large. If these three plant sizes were the only possible choices, which one should the **firm** choose as it plans for the future? The answer depends on the rate of output the firm expects to produce from the plant. The smallest plant would have the lowest cost if an output rate of less than q_1 were produced. The medium-size plant would provide the least-cost method of producing output rates between q_1 and q_2. For any output level greater than q_2, the largest plant would be the most cost-efficient.

The long-run* ATC *curve shows the minimum average cost of producing each output level when the firm is free to choose among all possible plant sizes. It can best be thought of as a planning curve, because it reflects the expected per-unit cost of producing alternative rates of output while plants are still in the blueprint stage.

Exhibit 7 illustrates the long-run *ATC* curve when only three plant sizes are possible, and the planning curve *ABCD* is thus mapped out. Of course, given sufficient time, firms can usually choose among many plants of various sizes. **Exhibit 8** presents the long-run

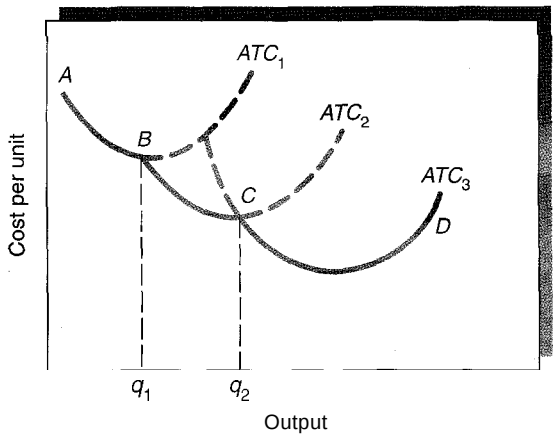


EXHIBIT 7
Long-Run Average Total cost

The short-run average total cost curves are shown for three alternative plant sizes. If these three were the only possible plant sizes, the long-run average total cost curve would be *ABCD.*

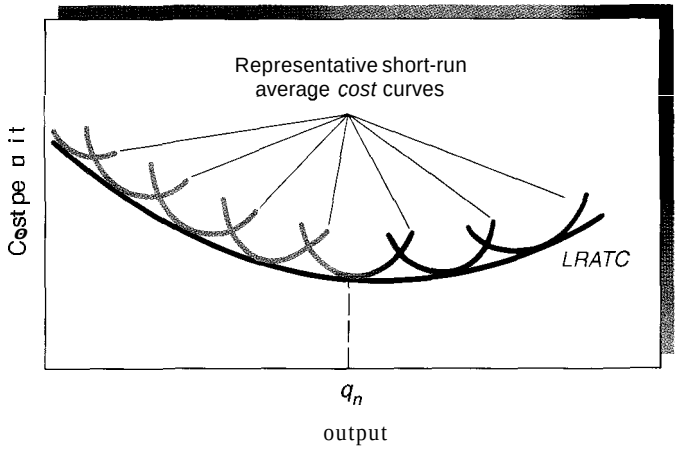
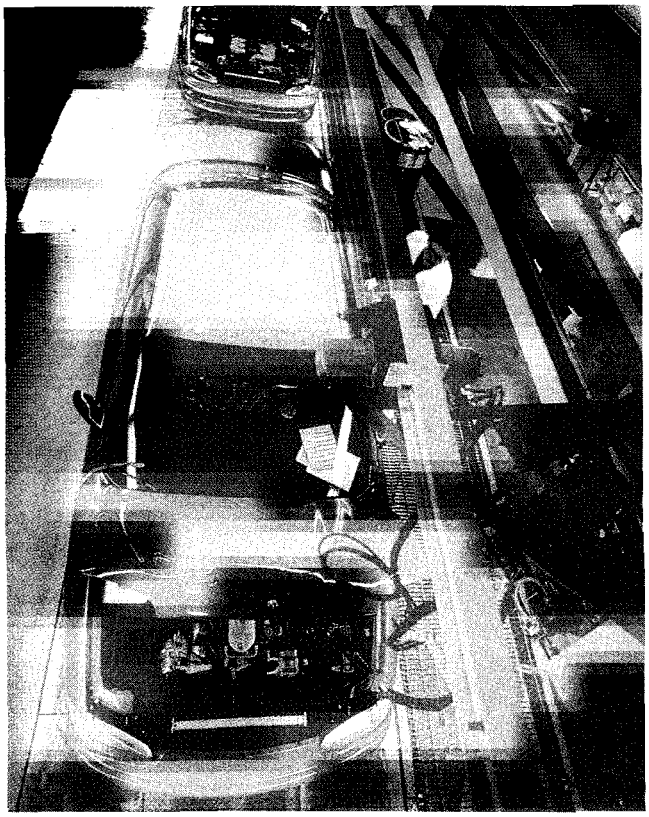


EXHIBIT 8
The Planning Curve (*LRATC*)

When many different plant sizes are possible, the long-run average total cost curve *(LRATC)* can be mapped out. When firms are able to plan large volumes of output, using mass-production methods will generally lead to lower per-unit costs. This helps explain why the *LRATC* has a downward-sloping portion.



© GREENLAR /THE IMAGE WORKS

When firms are able to plan large volumes of output, using mass-production methods will generally lead to lower per-unit costs. This helps explain why the *LRATC* has a downward-sloping portion.

planning curve under these circumstances. It is a smooth curve, with each short-run *ATC* curve tangent to it.

It is important to keep in mind that no single plant size could produce the alternative output rates at the costs indicated by the planning curve* LRATC *in Exhibit 8. Any of the planning curve options are, of course, available to the firm before a plant size is chosen and the plant is built. But it can *operate* in the short run only *after* a plant size has been chosen and put in place. The *LRATC* curve outlines the *possibilities* available in the planning stage. It shows the expected output and average total costs of production for the firm depending on the plant size it chooses.

Economies and Diseconomies of Scale

Do larger firms have lower minimum unit costs than smaller ones? The answer to this question depends on which industries are being considered. There is a sound basis, though, for expecting some initial reductions in per-unit cost from large-scale production methods. Why? Large firms typically produce a large total volume of output.[3] Volume of output denotes the total number of units of a product that the firm expects to produce.[4] There are three major reasons why planning a larger volume generally reduces, at least initially, unit costs: (1) economies accompanying the use of mass-production methods, (2) higher productivity as a result of specialization and "learning by doing," and (3) economies in promotion and purchasing. Let's consider each of these factors.

Mass-production techniques usually are economical only when large volumes of output are planned, since they tend to involve large development and setup costs. Once the production methods are established, though, marginal costs are low. For example, the use of molds, dies, and assembly line production methods reduce the per-unit cost of automobiles only when the planned volume is in the millions. High-volume methods, although cheaper to use for high rates of output and high volumes, will typically require high fixed costs and therefore cause unit costs to be far higher for low volumes of production.

Large-scale operation also allows the specialized use of labor and machines. In a giant auto plant, hundreds of different jobs must be done, and many of them require a training period for each worker. In a small plant, the same worker might do ten or twenty of these jobs, so each worker would have a much longer, more costly training period. Even then, the worker doing so many tasks might never fully develop the same level of proficiency of the more specialized worker. Baseball players improve by playing baseball, and pianists by playing the piano. Similarly, the employees of a firm improve their skills as they experience "learning by doing" in their jobs. Even better, concentration on a narrower range of tasks can help workers discover or develop cost-reducing techniques. The result of greater size and specialization is often more output per unit of labor.

Large firms are also able to achieve lower costs by spreading fixed costs (like the costs of advertising, developing specialized equipment, and searching out and negotiating better input prices, for example) over many more units. For example, both McDonald's and General Motors are able to spread these costs over a large number of stores and volume of sales. The cost advantages of scale come in many forms.

[3]Throughout this section, we assume that firms with larger plants necessarily plan a larger volume of output than do their smaller counterparts. Reality approximates these conditions. Firms choose large plants because they are planning to produce a large volume.

[4]Note the distinction between rate and volume of output. Rate of output is the number of units produced during a specific period (for example, the next six months). Volume is the total number of units produced during all time periods. For example, Boeing might produce two 777 airplanes per month (rate of output) while planning to produce a volume of two hundred 777s during the expected life of the model. Increasing the rate (reducing the time period during which a given output is produced) tends to raise costs, whereas increasing the volume (total amount produced) tends to lower costs.

© JANIE OSBORNE

The consistently high quality of gourmet restaurants like the Savory Olive in Bozeman, Montana, can seldom be duplicated by chain restaurants, in part because a gourmet chef must make decisions daily about which locally available fresh ingredients will be used and how they will be used to produce a constantly innovative menu, delivered by an attentive and dedicated staff Firms like the Savory Olive operate with diseconomies of scale

Economic theory explains why, at least initially, larger firms have lower unit costs than comparable smaller firms. Declining unit costs mean that **economies of scale** are present over the initial range of outputs. The long-run *ATC* curve is falling.

Economies of scale
Reductions in the firm's per-unit costs associated with the use of large plants to produce a large volume of output.

What about diseconomies of scale? **As** output continues to expand, is there reason to believe that larger firms will eventually have higher average total costs than smaller ones? The underlying causes of diseconomies of scale are less obvious, but they do occur. As a firm gets bigger and bigger, beyond some point bureaucratic inefficiencies *may* result. Inflexible procedures tend to replace managerial genius. Innovation requires clearance from more levels of management and becomes more difficult and costly. Motivating the workforce, carrying out managerial directives, and monitoring results of plans are also more complex when the firm is larger, and principal-agent problems grow as the number of employees increases and more levels of communication and monitoring are needed.

Circumstances vary, so diseconomies of scale set in at lower levels of firm size for some kinds of firms than for others. For example, firms in the fast-food industry can be very large and remain efficient; economies of scale apparently outweigh the diseconomies, even for giants like McDonald's. But in the fine-dining segment of the restaurant industry, the best restaurants seem to be small. Customers demand individual attention, and a constantly changing, innovative menu that takes advantage of the constantly changing array of locally available fresh ingredients, with consistently high quality as the only constant, is important. There are few truly gourmet restaurant chains because diseconomies seem *to* set in at a much smaller size at these firms. The bottom line for diseconomies of scale is this: for some firms, bureaucratic inefficiencies, principal-agent problems, difficulties with innovation, and similar problems that increase with firm size cause long-run average total costs to rise beyond some output level. However. there is considerable variation among industries and even among firms in the same industry concerning the precise output level at which diseconomies of scale begin to occur.

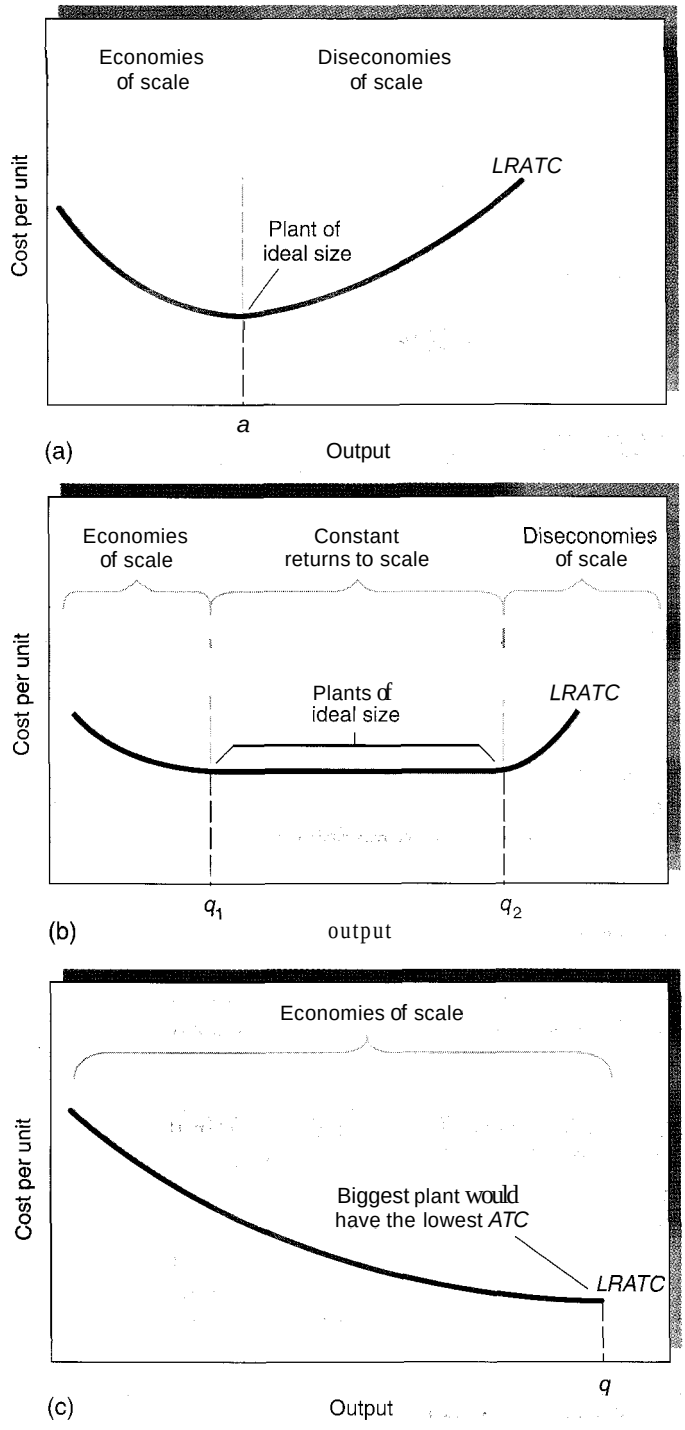
It is important to note that scale economies and diseconomies stem from sources different from those of increasing and diminishing returns. ***Economies and diseconomies of scale are long-run concepts. They relate to conditions of production when all factors are variable. In contrast, increasing and diminishing returns are short-run concepts, applicable only when the firm has at least one fixed factor of production.***

Alternative Shapes of the *LRATC*

Exhibit 9 outlines three different long-run average total cost *(LRATC)* curves, each describing real-world conditions in differing industries. For a firm described by the cost curve in part (a), both economies and diseconomies of scale are present. Higher per-unit

EXHIBIT 9
Three Different Types of Long-Run Average Total Cost Curves

For one type of *LRATC* curve, economies of scale are present for output levels less than q. but immediately beyond q. diseconomies of scale occur (a). In another instance, economies of scale are important until some minimum output level (q_1) is attained. Once the minimum level has been attained, there is a wide range of output levels (q_1 to q_2) consistent with the minimum *ATC* for the industry (b). In the third situation, economies of scale exist for all relevant output levels (c). As we will see later, this type of *LRATC* curve has important implications for how industries are structured.



costs will result if the firm chooses a plant size other than the one that minimizes the cost of producing output q. If each firm in an industry faces the same cost conditions, we can generalize and say that all plants larger or smaller than this ideal size will experience higher unit costs. A very narrow range of plant sizes would be expected in industries with the *LRATC* depicted by part (a). Some agricultural products and retail lines approximate these conditions.

Part (b) demonstrates the general shape of the *LRATC* that economists believe is present in most industries. Initially, economies of scale exist, but once a minimum efficient scale is reached, wide variation in firm size is possible. Firms smaller than the minimum efficient size would have higher per-unit costs, but firms larger than that would not gain a cost advantage. **Constant returns to scale** are present for a broad

Constant returns to scale
Unit costs that are constant as the scale of the firm is altered. Neither economies nor diseconomies of scale are present

range of output rates (between q_1 and q_2), in other words. This situation is consistent with real-world conditions in many industries. For example, small firms can be as efficient as larger ones in the apparel, lumber, and publishing industries, as well as in several retail industries.

In part (c) of Exhibit 9, economies of scale exist for all relevant output levels. The larger the firm size, the lower the per-unit cost. The *LRATC* for local telephone service can approximate the curve shown here.

WHAT FACTORS CAUSE COST CURVES TO SHIFT?

When we drew the general shapes of a firm's cost curves in both the long run and short run. we assumed that certain other factors — resource prices. taxes. regulations, and technology — remained constant as the firm altered its rate of output. Let's now consider how these other factors would affect production costs if they did not remain constant but changed either in response to increasing production or for some reason unrelated to the firm's output decisions.

Prices of Resources

If the price of resources used should rise, the firm's cost curves will shift upward, as **Exhibit 10** shows. Higher resource prices will increase the cost of producing each alternative output level. For example, what happens to the cost of producing automobiles when the price of steel rises? The cost of producing automobiles also rises. Conversely, lower resource prices will reduce costs and shift the cost curves downward at every plant size.

Taxes

Taxes are a component of a firm's cost. Suppose that an excise tax of 20 cents were levied on the seller for each gallon of gasoline sold. What would happen to the seller's costs? They would increase, just as they did in Exhibit 10. The firm's average total and marginal cost curves would shift upward by the amount of the tax. If the tax were an annual business license fee instead, it would raise the average cost, but not the variable cost.

Regulations

The government often imposes health, safety, environmental, and production regulations on business firms. Regulations might require businesses to provide a certain number of water fountains and restrooms for customers and workers. Federal regulations under the Americans with Disabilities Act compel many firms to make their facilities accessible to people in wheelchairs. Other regulations force firms to build certain features into their products. Strong bumpers and air bags in automobiles are examples. Although regulations

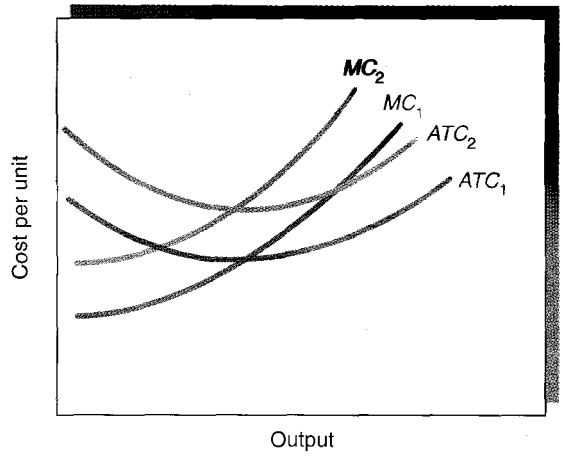


EXHIBIT 10
Higher Resource Prices and Cost

An increase in resource prices will cause the firm's cost curves to shift upward.

have their benefits, they are also costly. Just processing the paperwork that must be submitted to regulators is costly. and so are the compulsory changes themselves. Like tax increases, increases in regulatory compliance costs will shift cost curves upward. In some cases, only fixed costs will be affected; in other instances, variable costs will be altered as well. In both cases, the firm's *ATC* will be higher.

Technology

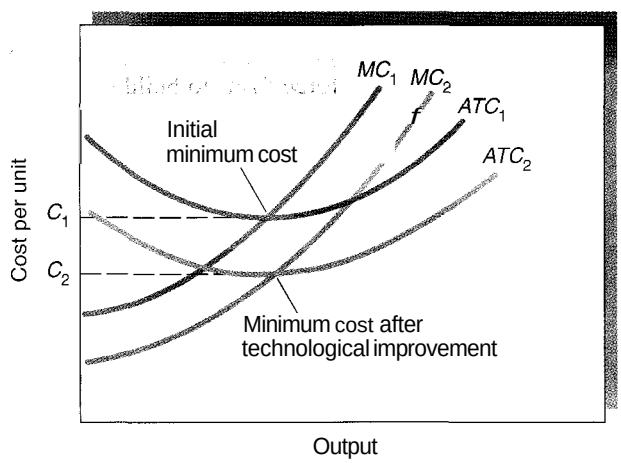
Improvements in technology often make it possible to produce a specific amount of output with fewer resources. Computers and robots have lowered costs in many industries. The Internet has made it easy to find and outsource many business-to-business services ranging from building maintenance and bookkeeping to the purchase of manufacturing components. Small firms can gain from specialization without being large themselves, and even large firms are participating. Federated Department Stores, which owns Bloomingdale's, Bon-Macy's, and Burdines-Macy's among others, now contracts with a specialty firm to manage all of its building facilities worldwide. Federated expects to save $150 million per year as the result of this change.[5] The *Wall Street Journal* reported that Micropub Systems of Rochester, New York, can install an entire brewery in ninety-six square feet, allowing the premium local brews to be made and delivered at lower cost than those currently available from the largest national brewers.[6] As **Exhibit 11** shows, a technological improvement will shift the firm's cost curves downward, reflecting the lower amount of resources needed to produce different levels of output.

THE ECONOMIC WAY OF THINKING ABOUT COSTS

When economists analyze the firm's costs, they often present a highly mechanical — some would say unrealistic — view that doesn't take into account the subjective nature of costs and the uncertainty about the future payoffs of many choices decision makers really face.

It is important to keep in mind that costs are incurred when choices are made. When business decision makers choose to purchase raw materials, hire new employees, or renew the lease on a plant, they incur costs. All these decisions, like other choices, must be made under conditions of uncertainty. Of course, past experience can help business decision makers anticipate the likely costs of various decisions. But the world is constantly changing; the future may differ substantially from the past.

EXHIBIT 11
Egg Production Costs and Technological Change



Suppose an egg producer discovers (or develops) a "super" mineral water that makes it possible to get more eggs from the same number of chickens. Because of this technological improvement, various output levels of eggs can now be produced with less feed, space, water, and labor. Costs will be reduced. The egg producer's *ATC* and *MC* curves will shift downward.

[5] "Economies of Scale: Spend to Save," *Chain Store Age 76*, no. 2 (February 2000): 2B.

[6] Thomas Petzinger Jr., "The Rise of the Small and Other Trends to Watch This Year," *Wall Street Journal*, January 9, 1998.

MYTHS OF ECONOMICS

"A good business decision maker will never sell a product for less than its production costs."

This statement contains a grain of truth. A profit-seeking entrepreneur will not undertake a project knowing the costs can't be covered. However, the statement fails to emphasize (1) the time dimension of the production process and (2) the uncertainty associated with business decisions. The production process takes time. Raw materials must be purchased, employees hired, and plants equipped. Retailers must contract with suppliers. As these decisions are made, costs result. Many of the firm's costs of production are incurred long before its product is ready for marketing.

Even a good business decision maker is not always able to predict the future because market conditions can change quickly and unexpectedly. At the time the product is ready for sale, buyers might be unwilling to pay a price that will cover the seller's past costs of production. These past costs, however, are now sunk costs and no longer relevant. Decisions must now be made on the basis of the firm's current cost and revenues.

Should a grocer refuse to sell oranges that are about to spoil because their wholesale cost cannot be covered? The grocer's current opportunity cost of selling the oranges at this point is nearly zero. The alternative would be to throw them in the garbage next week. Almost any price, even one far below past costs, will be better than letting the oranges spoil.

Consider another example. Suppose a couple who owns a house plans to relocate temporarily. Should they refuse to rent the house they're moving out of for $500 (if this is the best offer available) because their monthly house payment is $800? Of course not. The house payment will go on, regardless of whether or not they rent the house. If the homeowners can cover their opportunity costs (perhaps wear and tear plus a $60 monthly fee for a property management service), they will gain by renting rather than leaving the house vacant.

Past mistakes provide useful lessons for the future, but they cannot be reversed. Bygones are bygones, even if they resulted in business loss. There is no need to fret over spilt milk, burnt toast, or yesterday's business losses.

Opportunity costs are expected costs—they represent the highest valued option that the decision maker expects to give up as the result of a choice. Think for a moment of what the cost curves developed in this chapter really mean. The firm's short-run *MC* curve represents the opportunity cost of expanding output, *given the firm's current plant size.* The firm's long-run *ATC* curve represents the opportunity cost per unit of output associated with varying plant sizes and rates of output, *given that the alternative plants are still on the drawing board.* Opportunity costs look forward, reflecting expectations of what will be forgone as a result of current decisions. At the time decisions must be made, neither the short-run *MC* nor the long-run *ATC* can be determined from accounting records, since accounting costs look backward. Accounting figures yield valuable information about historical costs, but, as the following section illustrates, they must be interpreted carefully when they are used to forecast future costs.

What Are Sunk Costs?

Sunk costs are the historical costs of past decisions that cannot be reversed. Sunk costs give managers hindsight when it comes to making current decisions, but the specific costs themselves are no longer relevant. When past choices cannot be reversed—no refund is available, for example—money that has been spent is gone for good. Today's choices must be based on the costs and benefits expected under *current and future* market conditions, if mistakes are to be avoided (see the accompanying Myths of Economics feature).

Sunk costs
Costs that have already been incurred as a result of past decisions. They are sometimes referred to as historical costs.

To minimize costs, business decision makers need to realize that sunk costs are, indeed, *sunk.* A simple example will emphasize this point. Suppose that the firm in Exhibit 5 pays $100,000 to purchase and install a roller blade–producing machine. The machine is expected to last ten years. The company's books record the cost of the machine as $10,000 each year under the heading of depreciation. The machine can be used only to make roller blades, though. Because dismantling and reinstallation costs are high, it cannot be leased or sold to another firm. Also, it has no scrap value. In other words, there are no alternative uses for the machine. The machine'? annual production of roller blades will generate $50,000 of revenues for the firm when it is employed with raw

materials and other factors of production that cost $46,000. Thus, the net revenue generated by the machine is $4,000.

Should the firm continue to use the machine? Its depreciation figures suggest that the machine is costing the firm $10,000 annually, compared to the $4,000 net revenue it generates. Put another way, the machine is reducing the firm's profit by $6,000 annually. The machine's depreciation cost, however. is a sunk cost. It was incurred when the machine was purchased and installed, which is over and done with now. The *current* opportunity cost of the machine is therefore precisely zero. In other words, the firm is not giving up anything by continuing to use it today. Since using the machine generates $4,000 of additional net revenue, the firm can gain by continuing to use it. Of course, if market conditions are not expected to improve, the firm will not purchase a similar machine or replace the machine when it wears out, but this should not influence its decision whether to continue operating the one it already has. The irrelevance of sunk costs helps explain why it often makes sense to continue using older equipment (it has a low opportunity cost), even if it would not be wise to purchase similar equipment again.

How Will Cost Influence Supply?

Economists are interested in cost because they seek to explain firms' supply decisions. A strictly profit-maximizing firm will compare the expected revenues derived from a decision or a course of action with the expected costs. If the expected revenues exceed costs, then the course of action will be chosen because it will expand profits (or reduce losses).

For short-run supply decisions, the marginal cost of producing additional units is what should be considered. To maximize profits, the decision maker should compare the expected marginal costs with the expected additional revenue from larger sales. If the latter exceeds the former, output (the quantity supplied) should be expanded.

Marginal costs are central to the choice of short-run output, whereas expected average total cost is vital to a firm's long-run supply decision. Before entering an industry (or purchasing capital assets for expansion or replacement), a profit-maximizing decision maker will compare the expected market price with the expected long-run average total cost. Profit-seeking potential entrants will supply the product if, and only if, they expect the market price to exceed their long-run average total cost. Similarly, existing firms will continue to supply a product in the long run only if they expect that the market price will enable them at least to cover their long-run average total cost.

LOOKING AHEAD

In this chapter; we outlined several basic principles that affect costs for business firms. We will use these basic principles in the chapters that follow when we analyze the price and output decisions of firms under alternative market structures.

KEY POINTS

- The business firm is used to organize productive resources and transform them into goods and services. There are three major types of business structure—proprietorships, partnerships, and corporations.
- To solve the principal-agent problem, which tends to reduce worker efficiency in team production, every firm must provide work incentives and monitoring.

- The demand for a product indicates the intensity of consumers' desires for the item. The (opportunity) cost of producing the item indicates the intensity of consumers' desires for other goods that could have been produced instead, with the same resources.
- In economics, total cost includes not only explicit payments for resources employed by the firm, but also the implicit costs associated with the use of productive resources owned by the firm (like the opportunity cost of the firm's equity capital or owner-provided services) that could be used elsewhere.
- Because accounting methods omit the cost of equity capital (and sometimes other implicit costs), they tend to understate the opportunity cost of producing a good and overstate the firm's economic profit.
- Economic profit (or loss) results when a firm's sales revenues exceed (or are less than) its total costs, both explicit and implicit. Firms that are earning the market (or "normal") rate of return on their assets will therefore make zero economic profit.
- The firm's short-run average total cost *(ATC)* curve will tend to be U-shaped.
- The law of diminishing returns explains why a firm's short-run marginal and average total costs will eventually rise. When diminishing marginal returns are present, successively larger amounts of the variable input will be required to increase output by one more unit. As this happens, marginal cost will rise.
- The long-run *ATC (LRATC)* reflects the costs of production for plants of various sizes. When economies of scale are present, *LRATC* will decline. When constant returns to scale are experienced, *LRATC* will be constant. When diseconomies of scale are present, *LRATC* will rise.
- Changes in: (1) resource prices, (2) taxes, (3) regulations, and (4) technology will cause the cost curves of firms to shift.
- Sunk costs are costs that have already been incurred and cannot be recovered. Sunk costs give managers hindsight when it comes to making current decisions, but the specific costs themselves are no longer directly relevant for current and future decisions.

CRITICAL ANALYSIS QUESTIONS

***1.** What is economic profit? How might it differ from accounting profit? Explain why firms that are making zero economic profit are likely to continue in business.

***2.** Which of the following statements do you think reflect sound economic thinking? Explain your answer.

a. "I paid $400 for this economics course. Therefore, I'm going to attend the lectures even if they are useless and boring."

b. "Because we own rather than rent, and the house is paid for, housing doesn't cost us anything."

c. "I own 100 shares of stock that I can't afford to sell until the price goes up enough for me to get back at least my original investment."

d. "Private education is costly to produce, whereas public schooling is free."

3. Suppose a firm produces bicycles. Will the firm's accounting statement reflect the opportunity cost of the bicycles? Why or why not? What costs would an accounting statement reveal? Should current decisions be based on accounting costs? Explain.

4. What is the principal-agent problem? When will the principal-agent problem be most severe? Why might there be a principal-agent problem between the stockholder-owners and the managers of a large corporation?

***5.** "If a firm maximizes profit. it must minimize the cost of producing the profit-maximizing output." Is this statement true or false? Explain your answer.

6. What are some of the advantages of the corporate business structure of ownership for large business firms? What are some of the disadvantages'? Is the corporate form of business ownership cost-efficient? In a market economy, how would you tell whether the corporate structural form was efficient?

***7.** Explain the factors that cause a firm's short-run average total costs initially to decline. but eventually to increase, as the rate of output rises.

8. Which of the following are relevant to a firm's decision to increase output: (a) short-run average total cost, (b) short-run marginal cost, (c) long-run average total cost? Justify your answer.

9. Economics students often confuse (a) diminishing returns related to the variable factors of production and (b) diseconomies of scale. Explain the difference between the two, and give one example of each.

10. "Firms that make a profit have increased the value of the resources they used; their actions created wealth. In contrast, the actions of firms that make losses reduce wealth. The discovery and undertaking of profit-making opportunities are key ingredients of economic progress." Evaluate the statement.

***11.** Is profit maximization consistent with the self-interest of corporate owners? Is it consistent with the self-interest of corporate managers? Is there a conflict between the self-interests of owners and those of managers?

***12.** What is the opportunity cost of (a) borrowed funds and (b) equity capital? Under current tax law, firms can record as an expense the opportunity cost of borrowed funds, but not equity capital. How does this tax law affect the amount of debt the firm wants to incur, compared to the amount of money it raises by selling equity?

13. Why do economists consider normal returns to capital to be a cost? How does economic profit differ from normal profit?

***14.** Draw a U-shaped, short-run *ATC* curve for a firm. Construct the accompanying *MC* and AVC curves.

15. What is shirking? If the managers of a firm are attempting to maximize its profits, will they have an incentive to limit shirking? How might they go about doing so?

16. What are implicit costs? Do implicit costs contribute to the opportunity cost of production? Should an implicit cost be counted as cost? Give three examples of implicit costs. Does the firm's accounting statement take implicit costs into account? Why or why not?

***17.** Consider a machine purchased one year ago for $12,000. The machine is being depreciated $4,000 per year over a three-year period. Its current market value is $5,000, and the expected market value of the machine one year from now is $3,000. If the interest rate is 10 percent, what is the expected cost of holding the machine during the next year?

***18.** Investors seeking to take over a firm often bid a positive price for the business even though it is currently experiencing losses. Why would anyone ever bid a positive price for a firm operating at a loss?

19. Fill in the blanks in the following table:

a. What happens to total product when marginal product is negative?
b. What happens to average product when marginal product is greater than average product?
c. What happens to average product when marginal product is less than average product?
d. At what point does marginal product begin to decrease?
e. At what point does marginal cost begin to increase?
f. Summarize the relationship between marginal product and marginal cost.
g. What happens to marginal costs when total product begins to fall?
h. What is happening to average variable costs when they equal marginal costs?
i. Marginal costs equal average variable costs between what output levels?
j. What is happening to average total costs when they equal marginal costs?
k. Marginal costs equal average total costs between what output levels?

*Asterisk denotes questions for which answers are given in Appendix B

Units of Variable Input	Total Product	Marginal Product	Average Product	Price of Input	Total Variable cost	Average Variable cost	Total Fixed cost	Total cost	Average Total cost	Marginal cost
0	0	___	___	$1	___	___	$2	___	___	___
1	6	___	___	$1	___	___	$2	___	___	___
2	15	___	___	$1	___	___	$2	___	___	___
3	27	___	___	$1	___	___	$2	___	___	___
4	37	___	___	$1	___	___	$2	___	___	___
5	45	___	___	$1	___	___	$2	___	___	___
6	50	___	___	$1	___	___	$2	___	___	___
7	52	___	___	$1	___	___	$2	___	___	___
8	50	___	___	$1	___	___	$2	___	___	___

Price Takers and the Competitive Process

CHAPTER 9

Chapter Focus

- How do firms that are price takers differ from those that are price searchers?
- What determines the output of a price taker?
- How do price takers respond to changes in price in the short run? In the long run?
- How does time influence the elasticity of supply?
- What must firms do in order to make profits? How do profits and losses influence the supply and market price of a product?
- How does competition provide an incentive for producers to supply goods that consumers want at a low cost?

Competition means decentralized planning by many separate persons.

—Friedrich A. von Hayek[1]

[I]t is competition that drives down costs and prices, induces firms to produce the goods consumers want, and spurs innovation and the expansion of new markets. . . .

—President's Council of Economic Advisers[2]

[1]F. A. von Hayek, "The Use of Knowledge in Society," *American Economic Review* 35 (September 1945): 521.

[2]President s Council of Economic Advisers *Economic Report of the President,* 1996 (Washington, D.C.: U.S. Government Printing Office, 1996), 155

In the previous chapter, we focused on cost conditions of business firms. In this chapter and the next two, we will take a closer look at the price and output decisions of the firm, and analyze how they are influenced by market conditions. We will also consider how the structure of a market — for example, the number of firms, the control they have over price, and the ease of entry into the market — influences the decision making of firms. What determines the profitability of firms, and how does the level of profit influence the market supply? When goods and services are allocated by markets, will resources be allocated efficiently? Is there any reason to believe that there will be a link between market allocation and economic prosperity? These are the major questions that we will address in the next several chapters.

PRICE TAKERS AND PRICE SEARCHERS

Price takers
Sellers who must take the market price in order to sell their product. Because each price taker's output is small relative to the total market. price takers can sell all their output at the market price. but they are unable to sell any of their output at a price higher than the market price.

This chapter will focus on markets in which the firms are **price takers:** they simply take the price that is determined in the market. ***In a price-taker market, the firms all produce identical products (for example, wheat, eggs, or regular unleaded gasoline), and each seller is small relative to the total market. Thus, the output of any single firm has little or no effect on the market price. Each firm can sell all its output at the market price, but cannot sell any of its output at a higher price.*** When a firm is a price taker, there is no pricing decision to be made. Price takers try to choose the output level that will maximize profit. given their costs and the price determined by the market.

Price takers, like all other profit-seeking firms, cannot thrive (or even survive) in a competitive environment unless they are sensitive to cost. However, price-taker markets and price-searcher markets have differing degrees of competition, differing ease of entry, and perhaps differing scale economies, too. To compete, each firm has to provide a high level of delivered benefits per dollar, compared to what consumers can find elsewhere. No firm can force consumers to purchase its product, and all products have many substitutes. Successful firms are those that stay ahead of competitors and potential competitors.

APPLICATIONS IN ECONOMICS

The Aalsmeer Flower Auction: An Illustration of a Competitive Market

The Aalsmeer Flower Auction, located near Amsterdam, is a highly competitive market with a large number of both buyers and sellers. Approximately 5,000 growers from the Netherlands and other countries like Israel, Kenya, and Zambia supply their products to this market. On a typical day, several thousand buyers representing wholesale florists around the world participate in the auction. More than 18 million flowers and 2 million plants are sold each day from a building equal in size to 125 football fields.

The 50,000 daily transactions are made possible by the Dutch auction system. Under this system, a clock runs backward from the highest to the lowest price per unit. The buyer pushes a button indicating his or her willingness to purchase when the clock reaches an acceptable per-

© DAVE BARTRUFF/CORBIS

unit price. The purchaser is the first buyer to push the button. Flowers and plants auctioned in the morning at Aalsmeer will be available in shops around the world within twenty-four hours.

In the real world, most firms are not price takers. In most cases, firms that lower their prices are able to attract additional customers. Correspondingly, firms are usually able to increase their prices, at least a little, without losing all their customers. For example, if Nike increased the price of its athletic shoes by 10 percent, the number of shoes sold would decline, but it would not fall to zero. Firms like Nike are not price takers. They are **price searchers**: they choose the price that they will charge for their product, but the quantity that they are able to sell is very much related to that price. To maximize their profits, price searchers must not only decide how much to produce, but also what price to charge. We will examine markets in which the firms are price searchers in the following two chapters.

Price searchers
Firms that face a downward-sloping demand curve for their product. The amount the firm is able to sell is inversely related to the price it charges.

If most real-world firms are price searchers, not price takers, why take the time to analyze the latter? There are several reasons. First, even though most firms are not price takers, there are a number of important markets, particularly in agriculture, in which the firms do essentially take the price determined in the market. Second, the price-taker model helps clarify the relationship between the decision making of individual firms and the market supply in both price-taker and price-searcher markets. Finally, and perhaps most important, the study of markets in which firms are price takers enhances our knowledge of **competition as a dynamic process.** Understanding how the competitive process works when firms are price takers will also contribute to our understanding of the process as it applies to many price searchers.

Competition as a dynamic process
Rivalry or competitiveness between or among parties (for example, producers or input suppliers) *to* deliver a better deal to buyers in terms of quality, price, and product information.

Historically, the term **pure competition** has been used to refer to markets in which firms are price takers. However, these markets are increasingly referred *to* as "price-taker markets" because this expression is more descriptive. Furthermore, this label avoids the implication that competitive forces are necessarily less pure or less intense in price-searcher markets. Often this is not the case. Many price searchers use a broad array of competitive weapons—for example, quality of product, style, convenient location, advertising, and price—all in an effort to attract consumers. When **barriers to entry** are low, the competitive process is just as important in price-searcher markets as it is when the firms are price takers.

Pure competition
A market structure Characterized by a large number of small firms producing an identical product in an industry (market area) that permits complete freedom of entry and exit. Also called price-taker markets.

Nonetheless, it should be noted that price-taker markets and purely competitive markets are really alternative names for the same thing. So when you hear people talk about pure competition or purely competitive markets, they mean markets with characteristics like those analyzed in this chapter.

Barriers to entry
Obstacles that limit the freedom of potential rivals to enter and compete in an industry or market

WHAT ARE THE CHARACTERISTICS OF PRICE-TAKER MARKETS?

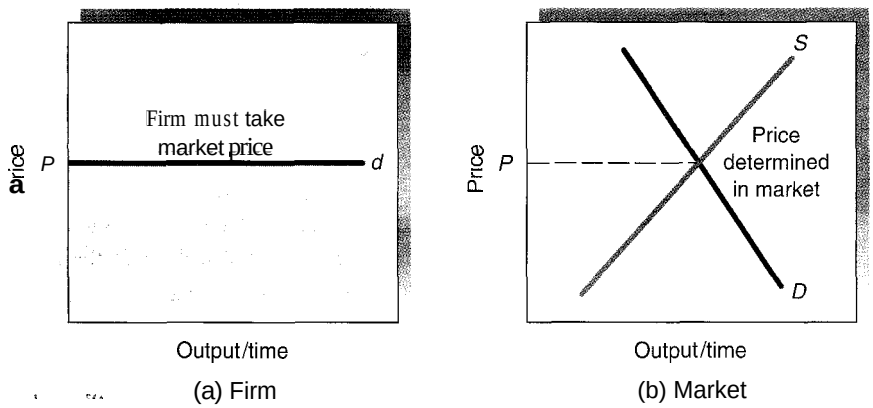
Consider the situation of Les Parrot, a Texas cattle rancher. In the financial pages of the local newspaper, he finds that the current market price of quality steers is 88 cents per pound. Even if his ranch is quite large, there is little that Parrot can do to change the market price of beef cattle. After all, there are tens of thousands of farmers who raise cattle. Thus, Parrot supplies only a small portion of the total cattle market. The amount that he sells will exert little or no effect on the market price of cattle. Parrot is a price taker.

© JACK KURTZ/THE IMAGE WORKS

Producers in the wheat-farming and beef cattle markets are price takers. If they are going to sell their output, they must do so at the price determined by the market. Because individual producers are small relative to the total market, they can sell as many units as they like at the market price.

EXHIBIT 1
The Price Taker's Demand Curve

The market forces of supply and demand determine price (b). Price takers have no control over price. Thus, the demand for the product of the firm is perfectly elastic (a).



The firms in a market will be price takers when the following four conditions are met:

1. All the firms in the market are producing an identical product (for example, beef, or cattle, of a given grade).
2. A large number of firms exist in the market.
3. Each firm supplies only a very small portion of the total amount supplied to the market.
4. No barriers limit the entry or exit of firms in the market.

When these conditions are met, firms selling in the market must accept the market price. This is why they are called price takers. **Exhibit 1** illustrates the relationship between the market forces (part b) and the demand curve facing the price-taking firm (part a). If the firm sets a price above the market level, consumers will simply buy from other sellers. Why pay the higher price when the identical good is available elsewhere at a lower price? For example, if the price of wheat were $5.00 per bushel, a farmer would be unable to find buyers for wheat at $5.50 per bushel. A firm would gain nothing by setting its price below the market level, because any small firm in the market can already sell as much as it wants at the market price. A price reduction would only reduce revenues. A firm that is a price taker thus faces a perfectly elastic demand for its product. (In Exhibit 1, note that a lowercase *d* is used to denote the demand curve faced by the *firm,* whereas an uppercase *D* indicates the ***market*** demand curve.)

HOW DOES THE PRICE TAKER MAXIMIZE PROFIT?

The firm's output decision is based on comparing benefits with costs. A firm that decides to enter a market will expand its output as long as the benefits (additional revenues) from the production and sales of the additional units exceed their marginal costs. How will changes in output change the firm's costs? In the preceding chapter, we discovered that the firm's short-run marginal costs will eventually increase as the firm expands its output by working its fixed plant facilities more intensively. The law of diminishing marginal returns assures us that this will be the case. Eventually, both the firm's short-run marginal and average total cost curves will turn upward.

Marginal revenue (*MR*)
The incremental change in total revenue derived from the sale of one additional unit of a product.

What about the benefits or additional revenues from output expansion? **Marginal revenue** (*MR*) is the change in the firm's total revenue per unit of output. It is the additional revenue derived from the sale of an additional unit of output. Mathematically,

$$MR = \text{Change in total revenue/Change in output}$$

Since the price taker sells all units at the same price, its marginal revenue will be equal to the market price.

In the short run, the price taker will expand its output until marginal revenue (price) just equals marginal cost. This decision-making rule will maximize the firm's profits (or minimize its losses).

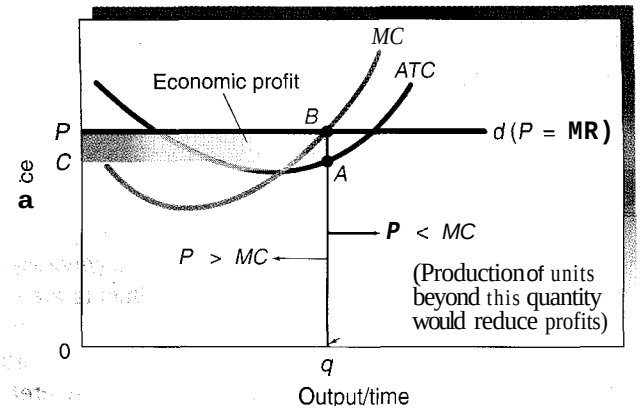


EXHIBIT 2
Profit Maximization When the Firm Is a Price Taker

The price taker will maximize profits by producing the output level q, where $P = MC$.

Exhibit 2 helps explain why. Since the firm can sell as many units as it would like at the market price, the sale of one additional unit will increase revenue by the price of the product. Does the firm gain by producing an extra unit? The answer is "Yes," as long as the marginal revenue (price. for the price taker) is greater than or equal to the marginal cost of that unit. Profit is simply the difference between total revenue and total cost. Profit will increase as long as the production and sale of a unit add more to revenue than to cost. Thus, the firm will gain from an increase in output as long as marginal revenue exceeds marginal cost. Eventually, however, as the firm produces a larger and larger quantity with its fixed-size plant, marginal costs will rise and exceed the price and marginal revenues. When the production of an additional unit adds more to cost than it adds to revenue, the firm's profit will be reduced if the unit is produced. Thus, the profit of the price taker is maximized at the output rate at which $P = MR = MC$. In Exhibit 2. this occurs at output level q.

A profit-maximizing firm with the cost curves shown in Exhibit 2 would produce exactly q units. The total revenue of the firm would be the sales price of each unit. P, multiplied by output sold, q. Total revenue is represented by the area $P0qB$. The firm's total cost, on the other hand, can be found by multiplying the average total cost (ATC) by the output level. Total cost is represented by the area $C0qA$. Because the firm's total revenues exceed its total costs, it is making short-run economic profit (the shaded area).

In the real world, of course, decisions are not made by entrepreneurs who spend time drawing demand and marginal cost curves. Many entrepreneurs have never even heard of these concepts. A business decision maker who has never heard of the $P = MR = MC$ rule for profit maximization, however, probably has another rule that yields approximately the same outcome. For example, the rule might be to produce those units, and only those units, that add more to revenue than to cost. This ensures maximum profit (or minimum loss). It also happens to be the point at which $P = MR = MC$. Why? To stop short of that point would mean that the firm would not be producing some profitable units that would add more to revenue than to cost. Similarly, the firm should not go beyond that point because producing more units would add more to cost than to revenue. This commonsense rule thus leads to the same outcome as our model, even when the decision maker knows none of the technical jargon of economics. No wonder economics is sometimes thought of as "organized common sense."

Just how accurate is the price taker's competitive model in predicting behavior in real markets? Do other models, which assume that sellers collude to eliminate competition, yield better predictions? Direct scientific evidence would help us answer these questions. As the Applications in Economics feature about the significance of competition shows, this evidence has, indeed, been produced repeatedly in recent decades by people working in the relatively new subdiscipline of experimental economics. The evidence indicates that the general implications of the price-taker model are valid under a variety of circumstances.

APPLICATIONS IN ECONOMICS

Experimental Economics: The Significance of Competition

A normal economic event observed in the real world can have more than one potential cause. An economist seeking to isolate the effect of a single variable—a change in price, for example—must therefore try to either hold other potential causal factors constant or adjust for their expected impact. This is not an easy task. In other disciplines, scientists use carefully designed laboratory experiments, which hold other factors constant, to test and verify the basic principles on which their science is built.

In the midtwentieth century, economists also began to conduct laboratory experiments. A good many experiments have been conducted to investigate the predictive power of the price-taker model. In one of the earliest, conducted in 1956 by leading experimental economist Vernon Smith, then at Purdue University, individuals were brought into a laboratory setting and arbitrarily assigned roles as buyers and sellers in a gamelike setting. Each buyer was given a different "limit price" (a maximum price he or she was allowed to pay) for a paper asset. Any buyer who could purchase the paper commodity for less than the limit price received a cash payment equal to the difference between the limit price and the amount actually paid. Each buyer gained financially by purchasing at lower prices—just as buyers and sellers do in real markets. The sellers were treated similarly. Each had a "limit price" (a minimum selling price) and received in cash any extra revenue above that price. In turn, buyers and sellers were free to make verbal offers and enter into exchanges at any mutually agreeable price.

The price-taker model predicts that all mutually advantageous trades among buyers and sellers will occur and that, as trades occur, the price of the good will converge toward a single price—the market price. Prior to the work in experimental economics, many economists thought this model was relevant only under highly restrictive conditions. Experimental researchers, however, have found that outcomes approximating those of the price-taker model generally emerge even when the strict assumptions of the model are absent. For example, even if the number of sellers is relatively small, say ten to fifteen, outcomes similar to those predicted by the price-taker model generally occur.

Participants in these experiments are often startled to discover that their competitive trading generated the largest possible joint income gain and, furthermore, that the competitive model presented in this chapter predicted this occurrence. Vernon Smith, who now directs the Interdisciplinary Center for Economic Studies at George Mason University, cites cases in which, after the experiment, participants describe the experimental market as "unorganized, unstable, chaotic, and confused." Generally, they are amazed when shown that their actions (trades) achieved the maximum income for the group, and that a sealed envelope, given to them prior to the experiment, predicted the approximate amount of their (maximum) joint gain.

At first, Smith himself was surprised that efficient outcomes were achieved under a wide variety of conditions, but after numerous experiments, he stated:

> In many experimental markets, poorly informed, error-prone, and uncomprehending human agents interact through the trading rules to produce social algorithms which demonstrably approximate the wealth maximizing outcomes traditionally thought to require complete information and cognitively rational actors.[1]

In addition to analyzing the exchange process, experimental research has addressed numerous other topics, including the impact of alternative auction rules, the likelihood of a stock market or real estate bubble, and the provision of public goods. In 2002, Vernon Smith was awarded the Nobel Prize for his path-breaking work in this area. Clearly, experimental economics is now one of the most exciting and fruitful areas of current economic research.

[1] Vernon L Smth, "Economics in the Laboratory," *Journal of Economic Perspectives* 8, no. 1 (Winter 1994): 118.

Profit Maximizing—A Numeric Example

Exhibit 3 uses numeric data to illustrate profit-maximizing decision making for a firm that is a price taker. Put yourself in the place of the owner of this firm. Your short-run total and marginal cost schedules have the general characteristics we discussed in the previous chapter. Since the firm confronts a market price of $5 per unit, its marginal revenue is $5. Total revenue thus increases by $5 per additional unit of output produced and sold. You will maximize your profit when you supply an output of 15 units.

There are two ways to look at this profit-maximizing output rate. First, profit is equal to the difference between total revenue and total cost. Thus, profit will be maximized at the output rate at which this difference (*TR* minus *TC*) is greatest. Column 6 of Exhibit 3

EXHIBIT 3
Profit Maximization for a Price Taker: A Numeric Illustration

(1) OUTPUT (PER DAY)	(2) TOTAL REVENUE (*TR*)	(3) TOTAL COST (*TC*)	(4) MARGINAL REVENUE (*MR*)	(5) MARGINAL COST (*MC*)	(6) PROFIT (*TR – TC*)
0	$ 0.00	$ 25.00	$0.00	$ 0.00	$-25.00
1	5.00	29.80	5.00	4.80	-24.80
2	10.00	33.75	5.00	3.95	-23.75
3	15.00	37.25	5.00	3.50	-22.25
4	20.00	40.25	5.00	3.00	-20.25
5	25.00	42.75	5.00	2.50	-17.75
6	30.00	44.75	5.00	2.00	-14.75
7	35.00	46.50	5.00	1.75	-11.50
8	40.00	48.00	5.00	1.50	– 8.00
9	45.00	49.25	5.00	1.25	– 4.25
10	50.00	50.25	5.00	1.00	– 0.25
11	55.00	51.50	5.00	1.25	3.50
12	60.00	53.25	5.00	1.75	6.75
13	65.00	55.75	5.00	2.50	9.25
14	70.00	59.25	5.00	3.50	10.75
15	75.00	64.00	5.00	4.75	11.00
16	80.00	70.00	5.00	6.00	10.00
17	85.00	77.25	5.00	7.25	7.75
18	90.00	85.50	5.00	8.25	4.50
19	95.00	95.00	5.00	9.50	0.00
20	100.00	108.00	5.00	13.00	– 8.00
21	105.00	125.00	5.00	17.00	-20.00

provides this information. For small output rates (less than eleven units). you **and** your firm would actually experience losses. But at fifteen units of output, an $1 1 profit is earned ($75 total revenue minus $64 total cost). A look at the profit figures of column 6 shows that it's impossible to earn a profit larger than $11 at any other rate of output.

In **Exhibit 4,** part (a) presents the total revenue and total cost approach in graph form. (However, the curves are drawn smoothly, as though output could be increased by tin) amounts, not just in whole-unit increments like those shown in Exhibit 3.) Profits will be greatest when the total revenue line lies above the total cost curve by the largest vertical amount. That takes place, of course, at fifteen units of output.

You can also use the marginal approach to determine the profit-maximizing rate of output for the firm. Remember, as long as price (marginal revenue) exceeds marginal cost. production and sale of additional units will add to the firm's profit (or reduce its losses). A look at columns 4 and 5 of Exhibit 3 reveals that *MR* is greater than *MC* for the first fifteen units of output. Producing these units will expand the firm's profit. In contrast. producing any unit beyond fifteen adds more to cost than to revenue. Profit will therefore decline if you expand output beyond fifteen units. Given the firm's cost and revenue schedule, you will maximize profit by producing fifteen, and only fifteen, units per day.

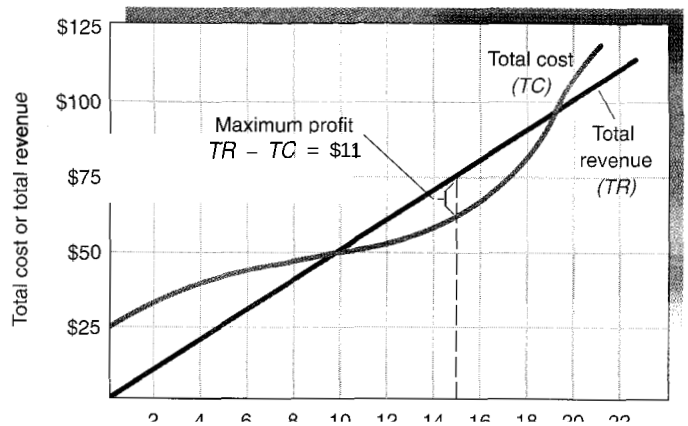
Part (b) of Exhibit 4 graphically illustrates the marginal approach. Note here that the output rate (fifteen units) at which the marginal cost and marginal revenue curves intersect coincides with the output rate in part (a) at which the total revenue curve exceeds the total cost curve by the largest amount. Beyond that output rate, *MR* is less than *MC*, so profit will decline.

Losses and When to Go Out of Business

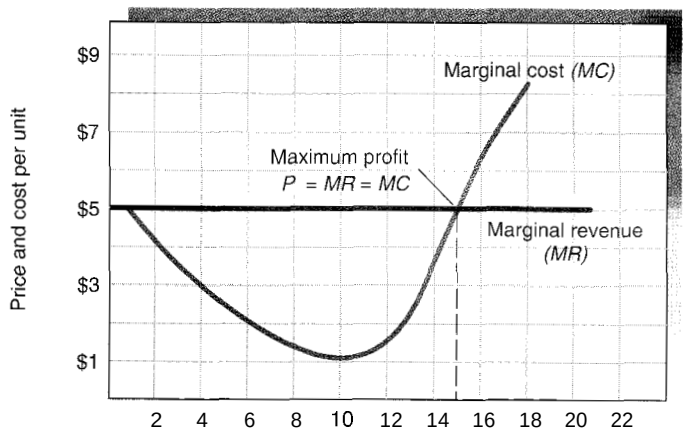
Suppose that market changes cause the price to drop below a firm's average total cost at all possible output levels. How will a profit maximizer (or loss minimizer) respond to this situation? The answer to this question depends both on the firm's current sales revenues relative to its variable cost and on its expectations about the future. The firm's owner has three options: (1) continue to operate in the short run, (2) shut down temporarily, or (3) go out of business.

EXHIBIT 4
The Firm's Profit-Maximizing Output Level—Total and Marginal Approaches

Using the data of Exhibit 3, we can find the profit-maximizing output of a price taker using either the total approach or the marginal approach. Using the total approach, we find that profits are maximized when the firm's total revenues exceed its total costs by the greatest amount (a). Using the marginal approach, we find that profits are maximized by comparing marginal revenue and marginal costs (b).



(a) Total revenue/total cost approach



(b) Marginal revenue/marginal cost approach

If the firm anticipates that the lower market price is temporary, it may want to continue operating in the short run as long as it can cover its variable cost.[3] **Exhibit 5** illustrates why. The firm shown in this exhibit would minimize its loss at output level q, where $P = MR = MC$. But at q, total revenues ($0qBP_1$) are less than total costs ($0qAC$). The firm faces short-run economic losses. Even if it shuts down completely, it will still incur fixed costs, *unless the firm goes out of business.* If it anticipates that the market price will increase enough to allow the firm to cover its average total costs in the future, it may not want to terminate operations and sell its assets. It may choose to produce q units in the short run, even though it will incur losses. At price P_1, producing output q is clearly more advantageous than shutting down, because the firm is able to cover its variable costs and have revenue remaining to pay some of its fixed costs. If it were to shut down but not sell out, the firm would lose the entire amount of its fixed cost.

What if the market price drops below the firm's average variable cost (for example, to P_2)? When this happens, not only is the firm not covering its fixed costs, it's not fully covering its variable, or per-unit, costs either. In other words, the more units the firm produces, the more money it loses. Under these circumstances, a temporary **shutdown** is preferable to short-run operation. Even if the firm's owner expects the market price to increase later, enabling the firm to survive and prosper in the future, shutting down in the short run will

Shutdown
A temporary halt in the operation of a firm. Because the firm anticipates operating in the future, it does not sell its assets and go out of business. The firm's variable cost is eliminated by the shutdown, but its fixed costs continue.

Going out of business
The sale of a firm's assets, and its permanent exit from the market. By going out of business, a firm is able to avoid its fixed costs, which would continue during a shutdown.

[3]Keep in mind the opportunity-cost concept The firm's fixed costs are opportunity costs that do not vary with the level of output They can be avoided if, and only if, the firm goes out of business To specify fixed costs, we need to know (1) how much the firm's fixed assets would bring if they were sold or rented to others and (2) any other costs, such as operating license fees and debts, that could he avoided if the firm declared bankruptcy and/or went out of business Since fixed costs can be avoided if the firm goes out of business, the firm will foresee greater losses from operating even in the short run if it does not expect conditions to improve.

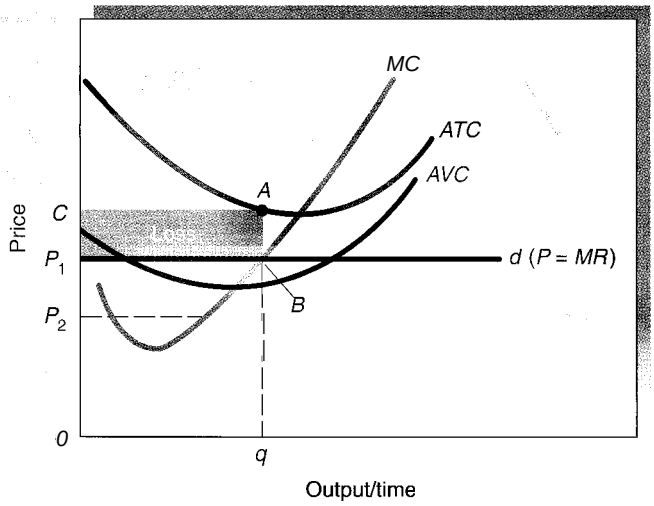


EXHIBIT 5
Operating with Short-Run Losses

A firm making losses will operate in the short run if it (1) can cover its variable costs now and (2) expects price to be high enough in the future to cover all its costs.

hold the line on its losses. The firm will still have to pay its fixed costs, but it won't be taking an additional loss on each unit it produces. Temporary shutdowns are actually regularly planned in some markets. For example, many ski resorts, golf courses, hotels, and restaurants in vacation areas plan to shut down in slow seasons, operating only when tourists or other seasonal purchasers provide enough demand. The price-taker model predicts that these firms will operate only when they expect to cover at least their variable costs.

The firm's third option is **going out of business** immediately. If the firm is sold, even the losses resulting from its fixed costs can be avoided. When market conditions are not expected to change for the better, going out of business is the preferred option.

© Rw S E SOm E

A Temporary Shutdown

A firm will temporarily shut down if it cannot cover its variable cost but does not want to go out of business because it expects to be able to operate profitably in the future. This Ice-cream store in Morgantown, West Virginia, closes for several months each winter and reopens during the summer. During these short-run shutdowns, the store still pays its fixed costs, such as its rent, taxes, and insurance. The store avoids only variable costs during the winter months.

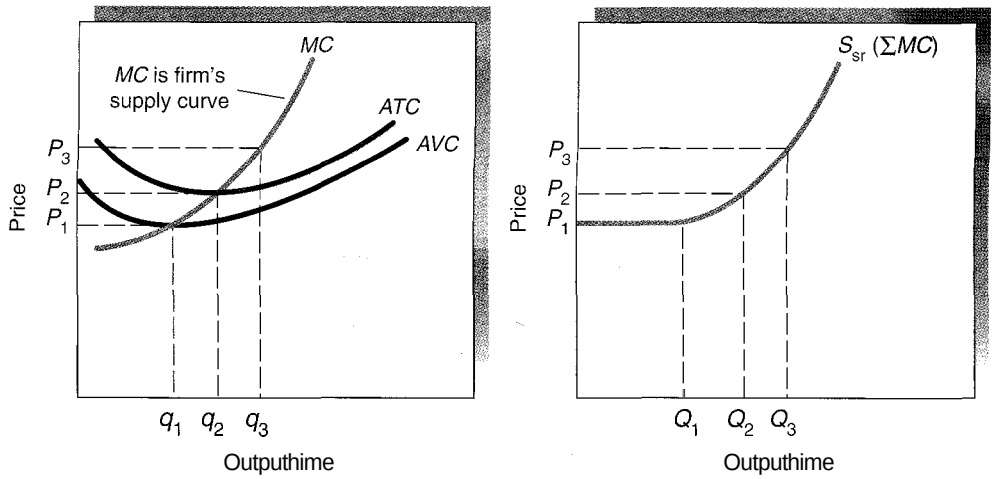
THE FIRM'S SHORT-RUN SUPPLY CURVE

The price taker that intends to stay in business will maximize profits (or minimize losses) when it produces the output level at which* P = MR = MC *and variable costs are covered. Therefore, the portion of the firm's short-run marginal cost curve that lies above its average variable cost is the short-run supply curve of the firm.

Exhibit 6 illustrates that, as the market price increases, the firm will expand output along its *MC* curve. If the market price were less than P_1, the firm would shut down immediately because it would be unable to cover even its variable costs. Shutting down will hold its loss to the payment of fixed costs only. If the market price is P_1, however, a price equal to the firm's average variable cost, the firm may supply output q_1 in the short run. Economic losses will result, but the firm would incur similar losses if it shut down completely. As the market price increases to P_2, the firm will happily expand output along its *MC* curve to q_2. At P_2, price is equal to average total costs. At this point, the firm is making a "normal rate of return," or zero economic profits. Higher prices will result in a still larger short-run output. For example, when the price rises to P_3, the firm will supply Q_3 units. At this price, economic profits will result. At still higher prices, output will be expanded even more. As long as price exceeds average variable cost, higher prices will cause the firm to expand output along its *MC* curve, which therefore becomes the firm's short-run supply curve.

EXHIBIT 6
The Short-Run Supply Curve for the Firm and the Market

As price increases, firms will expand output along their *MC* curve. Thus, the firm's *MC* curve is also its supply curve (a). When resource prices are constant, the short-run market supply (a) is merely the sum of the supply produced by all the firms in the market (b).



THE SHORT-RUN MARKET SUPPLY CURVE

The short-run market supply curve shows the total quantity supplied at alternative market prices. ***When the firms are price takers, the short-run market supply curve is the horizontal summation of the marginal cost curves (ΣMC) above the level of average variable cost for all of the firms in the market. Since individual firms will supply a larger amount at a higher price, the short-run market supply curve will slope upward to the right.***

Exhibit 6 illustrates this relationship. As the price of the product rises from P_1 to P_2 to P_3, the individual firms expand their output along their marginal cost curves. Since the individual firms supply a larger output as the market price increases, the total amount supplied to the market also expands.

Our construction of the short-run market supply curve assumes that the prices of the resources used by the industry do not change with industry output. When the entire industry (rather than just a single firm) expands output, resource prices may rise. The reason, of course, is that when just one firm expands, it has a minuscule effect on the market for resources, but when the entire industry expands output, the rise in resource demand is larger, so resource prices are more likely to rise. When this happens, the short-run market supply curve (reflecting the higher prices of purchased inputs) will be slightly more inelastic (steeper) than the sum of the supply curves of the individual firms.

PRICE AND OUTPUT IN PRICE-TAKER MARKETS

The short-run market supply curve, together with the demand curve for the industry's product, will determine the market price. At the short-run equilibrium market price, each firm will have expanded output until marginal costs are equal to the market price. In the short run, the firms may experience either profits or losses, but they will not have enough time to alter the size of their plants.

In the long run, however, firms will have the opportunity to expand the sizes of their plants. Firms will also be able to enter and exit the industry in the long run. As long-run adjustments are made, output in the whole industry may either expand or contract.

Long-Run Equilibrium

In addition to the balance between quantity supplied and quantity demanded necessary for short-run equilibrium, firms that are price takers can earn the normal rate of return, and only the normal rate, when the market is in long-run equilibrium. Why? ***If economic profit***

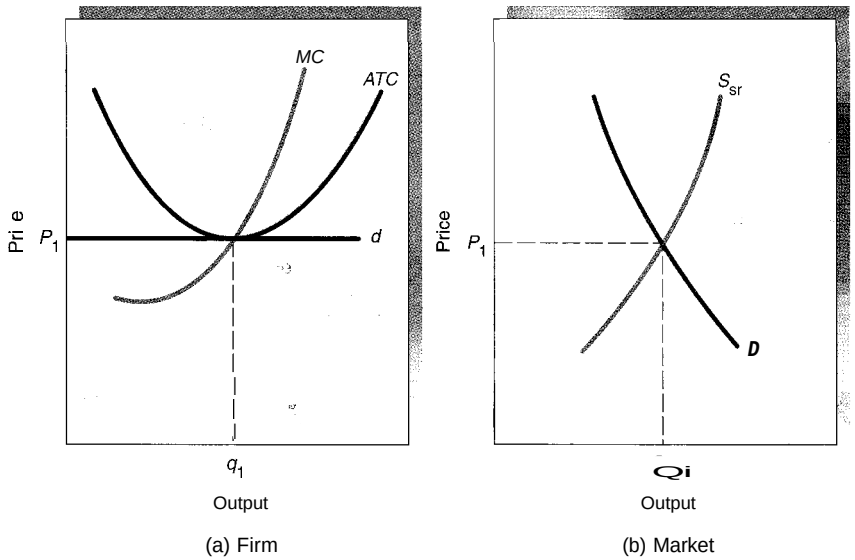


EXHIBIT 7
Long-Run Equilibrium in a Price-Taker Market

The two conditions necessary for equilibrium in a price-taker market are depicted here. First, quantity supplied and quantity demanded must be equal in the market (b). Second, the firms in the industry must earn zero economic profit (the "normal rate of return") at the established market price (a).

is present, new firms will enter the industry to capture some of those profits. Current producers will have an incentive to expand the scale of their operations to capture some of the additional profits, too. This increase in supply will put downward pressure on prices. In contrast, if firms in the industry are suffering economic losses, they will leave the market. This decrease in supply will put upward pressure on prices.

Therefore, as **Exhibit 7** shows, when a price-taker market is in long-run equilibrium, the quantity supplied and the quantity demanded will be equal at the market price (part b), and each firm in the industry will be earning normal (zero) economic profit (that is, its minimum *ATC* will just equal the market price) (part a). Firms will have no incentive to alter either their output levels or their plant sizes. Neither will there be incentive for net investment to flow into or away from the industry.

How Will the Market Respond to an Increase in Demand?

Suppose that a price-taker market is in equilibrium. What will happen if there is an increase in demand? **Exhibit 8** shows us. Suppose that an entrepreneur introduces a fantastic new candy. Consumers go wild over it. However, since it sticks to people's teeth, the market demand for toothpicks increases from D_1 to D,, pushing toothpick prices upward from P, to P,. What impact will the higher market price have on the output level of toothpick-producing firms? It will increase (from q_1 to q_2 in part a of the exhibit) as the firms expand output along their *MC* curves. In the short run, the toothpick producers will make economic profits. The profits will attract new toothpick producers to the industry and cause the existing firms to expand the scale of their plants. Hence. the **market supply** will increase (shift from S_1 to S_2) and eventually eliminate the short-run profits. **If** the prices of resources supplied to the industry are unchanged. the market supply will continue to expand (shift to the right) until the price of toothpicks **returns to** its initial level (P_1), even though output has expanded to Q_3.

How Will the Market Respond to a Fall in Demand?

Economic profits attract new firms to an industry. In contrast, economic losses (when they are expected to continue) encourage capital and entrepreneurship to move out of the industry and into other areas where the profitability potential is more favorable. Economic losses mean that the owners of capital in the industry (or firms that purchase the services of the capital) are earning less than the market rate of return. The opportunity cost of continuing in the industry exceeds the gain.

EXHIBIT 8
Market Response to Increased Demand

A new candy that sticks to people's teeth causes the demand for toothpicks to increase to D_2 (b). Toothpick prices rise to P_2, inducing firms to expand output. Toothpick firms make short-run profits (a), which draw new competitors into the industry. The toothpick supply then expands (shifts from S_1 to S_2). If cost conditions are unchanged, the expansion in supply will continue until the market price of toothpicks declines to its initial level of P_1.

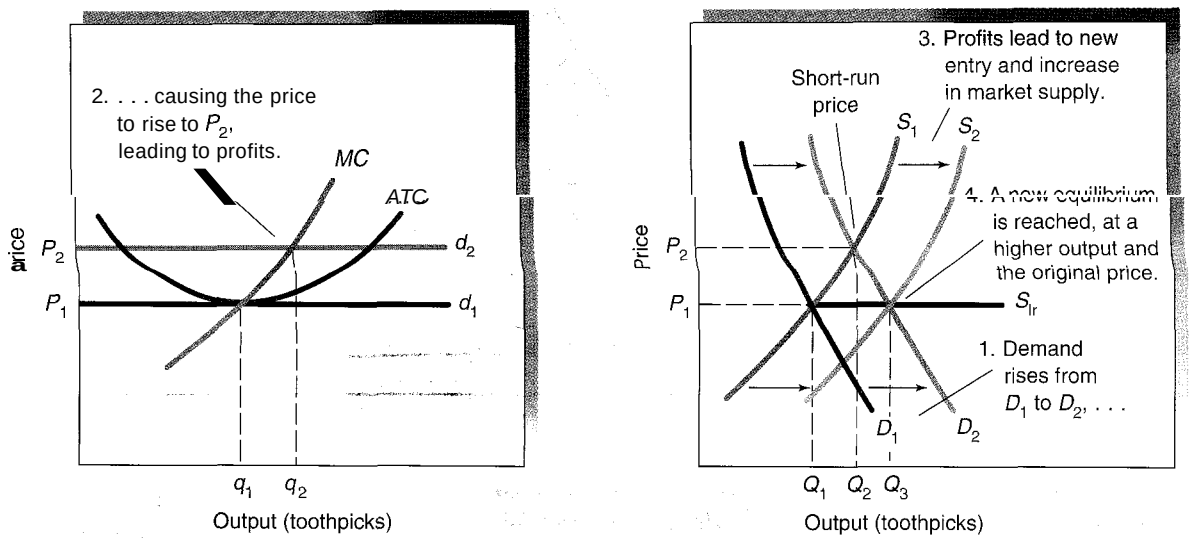


Exhibit 9 shows market forces reacting to economic losses. Initially, assume that an equilibrium price exists in the industry. The firms are able to cover their average costs of production. Now suppose that consumer incomes fall, lowering the product's demand and its market price. At the new, lower price, firms in the industry will not be able to cover their costs of production. In the short run, they will reduce output along their *MC* curves. The lower output by the individual firms results in a lower quantity supplied in the market.

In the face of short-run losses, the inflow of capital will decline and the industry's capital assets will shrink as firms fail to replace equipment when it wears out. Some firms will leave the industry as their fixed costs become variable—when machinery wears out and needs to be replaced, for example—and they are no longer able to cover their variable costs at the prevailing price. Others will reduce the scale of their operations, producing only those units for which the new, lower revenues can still justify the production costs. These factors will cause the industry supply to decline, indicated by the shift from S_1 to S_2. What impact will this have on price? It will rise. Over time, given no other shifts in demand, the short-run market supply curve will decline—continue shifting to the left—until the price rises enough that the firms remaining in the industry can once again earn "normal profits." At that point, a new long-run equilibrium is established. For a real-world example of how a market adjusts to changing demand and cost conditions, see the Applications in Economics feature on coffee production in a price takers' market.

The Long-Run Market Supply Curve

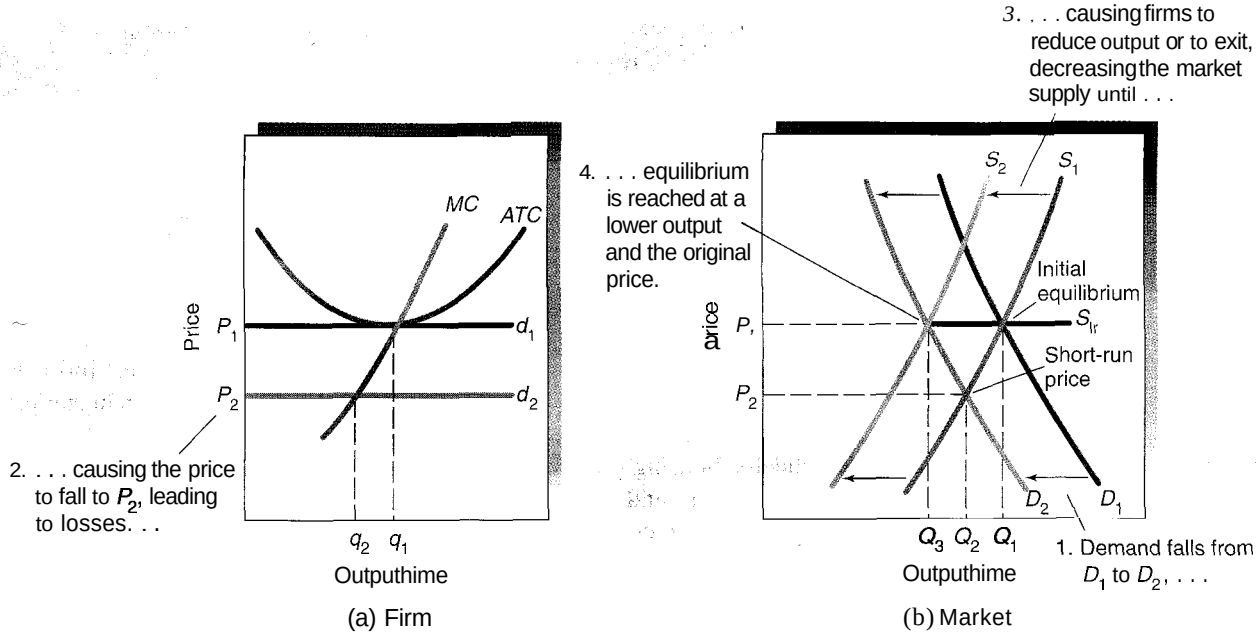
The *long-run market supply curve* shows the minimum price at which firms will supply various market output levels, given enough time to adjust the sizes of their plants (or other fixed factors) and to enter or exit the industry. The shape of the curve depends on what happens to the cost of production as the *industry's* output is altered. Three possibilities emerge, although one is far more likely than the other two.

Constant-cost industry
An industry for which factor prices and costs of production remain constant as market output is expanded. The long-run market supply curve is therefore horizontal in these Industries.

Long-Run Supply in Constant-Cost Industries If resource prices remain unchanged, the long-run market supply curve will be perfectly elastic. In terms of economics, this describes a constant-cost industry. Exhibits 8 and 9 both depict constant-cost industries.

EXHIBIT 9
The Impact of a Fall in Demand

Lower market demand will cause the price to fall and short-run losses to occur. The losses will cause some firms to go out of business and others to reduce their output. In the long run, the market supply will fall, causing the market price to rise. The supply will continue to decline and price will continue to rise until the short-run losses have been eliminated.



APPLICATIONS IN ECONOMICS

Coffee Production in a Price Takers' Market

Sellers in the world's coffee market are price takers. Hundreds of thousands of farmers produce coffee beans, and no grower has a significant impact on the world price. Each grower selling in the world market takes the price as given and is free to respond to any price change. Changes can be dramatic in the coffee market. For example, the market price rose from less than $1 per pound on average early in the 1990s to nearly $2 in early 1997.

As growing coffee became more profitable, coffee production around the world expanded. For example, Vietnamese growers more than quadrupled their coffee production, from 92.000 tons in 1990 to 487,000 tons in 1999. By 2000, Vietnam had become the world's second-largest coffee-producing nation.

But the greater supply of coffee turned out to be more than consumers were willing to buy. In 1997, the market price fell, along with the price paid to growers. By August 2001, coffee bean prices were averaging less than $.60 per pound. Many growers couldn't cover their costs at such low prices and began growing other crops Some landowners in Indonesia, for example, planted rice. Others, especially in Central and Latin America, began to plant organic and other specialty coffees, with lower yields but higher prices. Still other producers, including some in Vietnam, simply abandoned the least-profitable plantations, at least temporarily. In Mexico, more than an estimated 300,000 coffee farmers—unable to cover even their variable costs—left their farms to seek other opportunities.

The world coffee market illustrates very clearly how producers in price-taker markets can quickly expand production when rising prices are seen or expected but also contract production when prices and profits fall.

Source The information presented is from Howard LaFranchi, "Economic Upheaval over Coffee," *Christian Science Monitor* August 15, 2001, and from the BBC Web cite at http //www bbc co uk/worldservice/business/story_fdh200301 shtml (accessed September 17.2001) and from the Web site of the Intemational Coffee Organization at http://www.ico.org/. "Coffee Crisis" section (accessed June 16, 2004).

As Exhibit 8 shows. higher demand causes prices to increase *temporarily*. With time, however. the higher prices and profits will stimulate expansion and additional production, which will push the market price down to its initial level (and profitability to its normal rate). Because resource prices and the minimum per-unit cost remain constant as the market output expands, the larger supply will not require a permanent price increase. Similarly, Exhibit 9 illustrates the impact of a decline in demand in a constant-cost industry. Again, because resource prices and production costs are unaffected by the change in market output, the reduction in demand lowers the price in the short run, but not in the long run. Thus, in a constant-cost industry, the *long-run market supply curve* (S_{lr}) is perfectly elastic.

A constant-cost industry is most likely to arise when the industry's demand for resource inputs is quite small relative to the total demand for these resources. For example, the demand of the matches industry for wood, chemicals, and labor is very small relative to the total demand for these resources. Thus, doubling the output of matches would have very little effect on the price of the resources used by the industry. Matches therefore approximate a constant-cost industry.

Increasing-cost industry
An industry for which costs of production rise as output is expanded. In these industries, even in the long run, higher market prices will be needed to induce firms to expand total output in such industries. As a result, the long-run market supply curve in these industries will slope upward to the right.

Long-Run Supply in Increasing-Cost Industries In most industries, an increase in market demand and *industry* output will lead to higher per-unit production costs for all the firms in the industry. Economists refer to such industries as **increasing-cost industries**. The rising output and expanded resource demand in such industries result in higher prices for at least some resources, causing the firms' cost curves to shift upward. For example, a rising demand for housing places upward pressure on the prices of lumber, window frames, building sites, and construction labor, causing the cost of housing to rise. Similarly, an increase in demand (and market output) for beef will typically cause the prices of feed grains, hay, and grazing land to rise. Thus, the production costs of beef rise as more of it is produced.

For an increasing-cost industry, an expansion in market demand will bid up resource prices, causing the per-unit cost of the firms to rise. As a result, the market price will have to increase in order to induce firms to supply the larger output. The long-run market supply curve for the product will therefore slope upward.

Exhibit 10 depicts an increasing-cost industry. Greater demand causes higher prices and a larger market output. As the industry expands, the price of resources (factors of production) rises so that costs increase. What happens to the firm's cost curves? Both the average and marginal cost curves rise (shift to ATC_2 and MC_2). Greater production costs necessitate a higher long-run price (P_2), so the long-run supply curve slopes upward to the right.

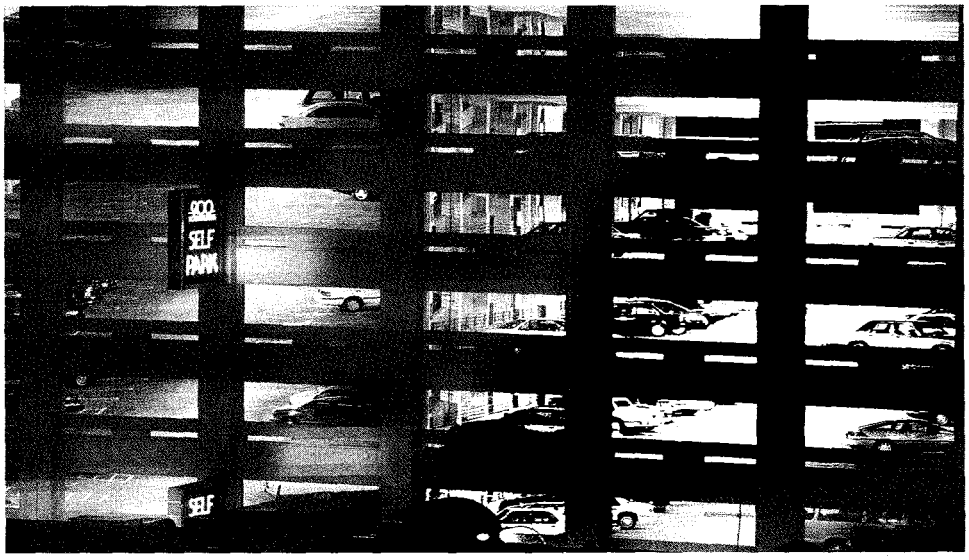
Decreasing-cost industry
An industry for which costs of production decline as the industry expands. The market supply is therefore inversely related to price. Such industries are atypical.

Long-Run Supply in Decreasing-Cost Industries Sometimes, factor prices will decline when the market output of a product is expanded. The lower resource prices will reduce the unit costs of the firms, placing them in a position to supply a larger market output at a lower price. In such **decreasing-cost industries**, the long-run (but not the short-run) market supply curve will slope downward to the right. For example, as the electronics industry expands, suppliers of certain components may be able to adopt large-scale production techniques that will lead to lower component prices. If rising electronics demand leads to reduced component cost (and if other resource prices do not rise to offset the reductions), then the cost curves of the electronics firms will shift downward. Under these circumstances, the industry supply curve for electronics products—reflecting the lower cost—will slope downward to the right.

In most industries, however, increases in demand and expansion in market output cause higher rather than lower input prices. Thus, increasing-cost industries are the norm, and decreasing-cost industries are rare.

Supply Elasticity and the Role of Time

It takes time for firms to adjust to a change in the price of a product. In the short run, firms are stuck with the existing sizes of their plants. If the price increases in the short run, they can expand output only by utilizing their existing plants more intensely. Thus, their output response will be limited. In the long run, however, they will have time to build new plants. This will allow them to expand output by a larger amount in response to an increase in price. Thus, the market supply curve will be more elastic in the long run than in the short run.



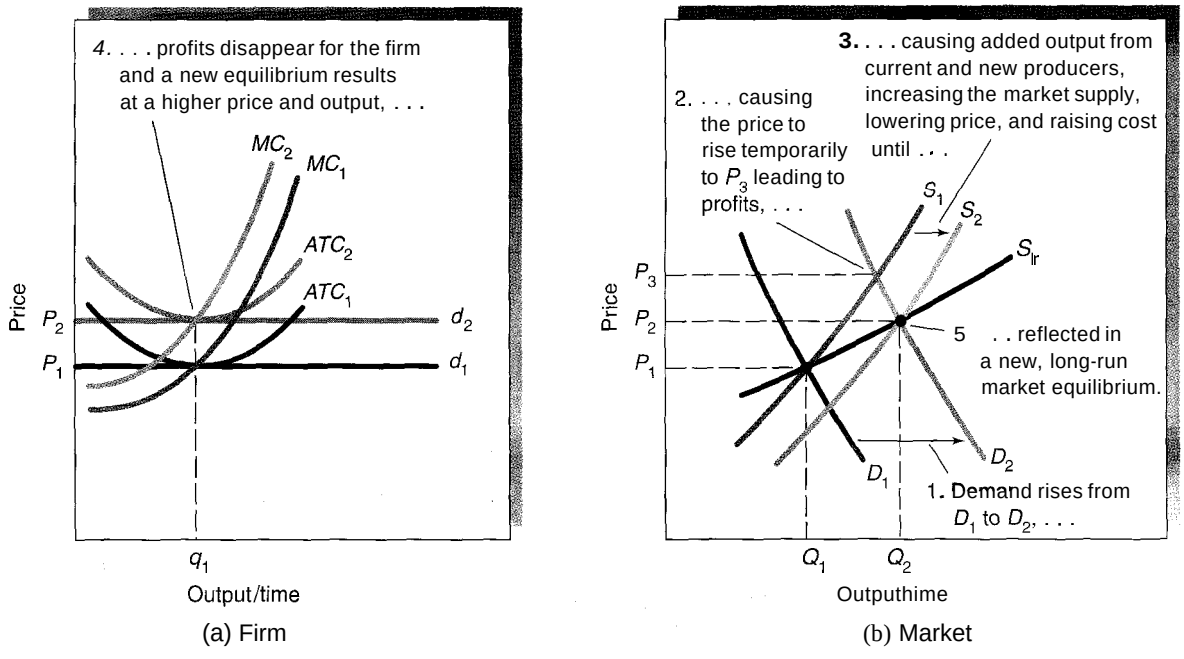
© HOA QUI/INDEX STOCK IMAGES

In central areas of large cities. the supply of parking spaces can be expanded only by using higher-cost space and higher-cost techniques, such as taller parking garages that use a large portion of the building for access ramps As a result, the unit cost of parking spaces increases as the total number is expanded Thus. provision of parking space is an increasing-cost industry

The short- and long-run distinction offers a convenient two-stage analysis, but in the real world there are many intermediate production "runs." The delivery rates for some factor inputs that could not be easily increased in a one-week time period can be increased over a two-week period. It might take two weeks, for example, to hire reliable workers who need no special skills. Expanding other factors might take a month, and still others, six months. Ordering a new custom-made machine tool at Boeing or hiring a competent new manager to expand an eastern Montana ranch operation might take a year. More precisely, the cost penalty for quicker availability is greater for some productive resources than others. In any case, a faster expansion usually means that greater cost penalties are encountered when a firm demands earlier availability of resources it needs for production.

When a firm has a longer time period to plan its output and adjust all of its productive inputs to their desired utilization levels, it will be able to produce any specific new rate of output at a lower cost. ***Because it* is *less costly to expand output slowly in response to a***

EXHIBIT 10
Increasing Costs and Long-Run Supply



demand increase, the expansion of output by firms will increase with time, as long as the price exceeds the cost. Therefore, the elasticity of the market supply curve will be greater when firms have more time to adjust their output.

Exhibit 11 shows the impact of time on producers' responses to an increase in price caused by greater demand. When the price of a product increases from P_1 to P_2, the immediate supply response of the firms is small, reflecting the high cost of hasty expansion. After one week, firms are willing to expand output only from Q_1 to Q_2. After one month, because of cost reductions made possible by the longer production-planning period, firms are willing to offer Q_3 units at the price P_2. After three months, the rate of output expands to Q_4. In the long run, when it is possible to adjust all inputs to the desired utilization levels (after a six-month time period, for example), firms are willing to supply Q_5 units of output at the market price of P_2. The supply curve for products is typically more elastic over a longer time period than over a shorter period.

THE ROLE OF PROFITS AND LOSSES

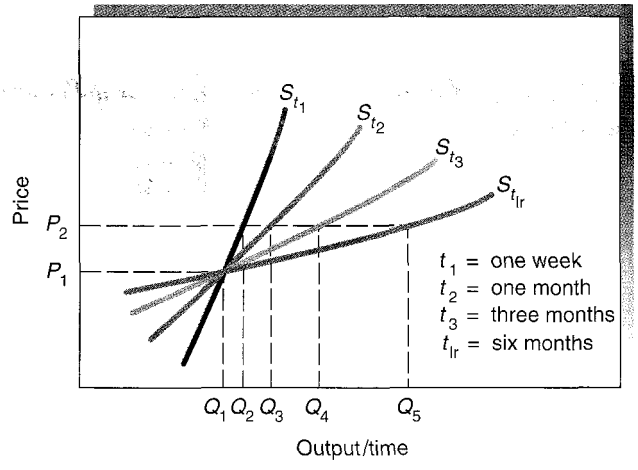
The price-taker model highlights the role of profits and losses: economic profits result when a firm or entrepreneur increases the value of resources. Business firms purchase resources and use them to produce a product or service that is sold to consumers. Costs are incurred as the business pays workers and other resource owners for the resources required to produce the product. If the sales revenue of the business firm exceeds the costs of employing all the resources required to produce the firm's output, then the firm will make a profit. ***In essence, profit is a reward that business owners will earn if they produce a good that consumers value more (as measured by their willingness to pay) than the resources required for the good's production (as measured by the cost of bidding the resources away from their alternative uses).***

For example, suppose that it costs a shirt manufacturer $20,000 per month to lease a building, rent the required machines, and purchase the labor, cloth, buttons, and other materials necessary to produce and market 1,000 shirts per month. Thus, the average cost of the shirts is $20 (the $20,000 monthly cost divided by the 1,000 monthly output). If the manufacturer sells the 1,000 shirts for $22 each, its actions create wealth. Consumers value the shirts more than they value the resources required for their production. The manufacturer's $2 profit per shirt is a reward received for increasing the value of the resources.

In contrast, losses are a penalty imposed on businesses that reduce the value of resources. Losses indicate that the value of the resources used by the firm (as measured by their cost) exceeds the price consumers are willing to pay for the product supplied. Losses, along with bankruptcies, are the market's way of providing signals and incentives to bring such wasteful activities to a halt.

EXHIBIT 11
Time and the Elasticity of Supply

The elasticity of the market supply curve usually increases when suppliers have more time to adjust to a change in the price.



We live in a world of changing tastes and technology, imperfect knowledge, and uncertainty. Business decision makers cannot be sure of either future market prices or costs of production. Their decisions must be based on expectations. Nonetheless, the reward-penalty structure of a market economy is clear. ***Firms that anticipate correctly the products and services for which future demand will be most urgent (relative to production costs), and produce and market them efficiently will make economic profits. Those that are inefficient and allocate resources incorrectly to areas of weak demand (relative to costs) will be penalized with losses.***

The firms most adept at giving consumers value for their money thrive and expand. Those less successful in doing so tend to shrink or even to disappear from the market. Free entry and the competitive process will protect the consumer from arbitrarily high prices. When profits are present, profit-seeking entrepreneurs will put more resources into these markets, supply will expand, and eventually the price will be driven down to the unit cost.

The Market Process in Action

When entry barriers are low, even small firms are often able to challenge and compete successfully against rivals that are much larger. Michael Dell's success in the personal computer market vividly illustrates this point. In 1984, Dell began producing and marketing personal computers while he was still a student at the University of Texas. At the age of nineteen, he began operations from his dormitory room with an investment of $1,000.[4] Then, as now, Dell Computer took orders for PCs, bought component parts from competing sellers, and custom-built the machines almost immediately to fill each order for quick delivery directly to consumers. This business strategy made it possible for the firm to keep inventories low and build the latest technical innovations into its products as soon as they were available. In contrast, IBM, Compaq, and Apple, the dominant firms during the 1980s, built specific PC models by the thousands and marketed them through retailers.

Dell's sales grew year after year and the firm eventually became the industry leader. By the first quarter of 2004, it was selling computers at a rate of $46 billion per year and Michael Dell was one of the world's richest individuals. On the other hand, IBM had exited from the PC market, a struggling Compaq had been bought out by Hewlett-Packard, and Apple's market share had declined substantially.

Although Dell Computer is not a price taker, it operates in a market with low entry barriers. Just as the price-taker model indicates, the low entry barriers will keep prices low and make sure that Dell and other producers in the industry stay on their toes. If Dell's cost efficiency were to fall or if it raised its prices significantly, rivals would expand, and new firms would be lured into the market. A higher price would be an open invitation for others to step in and steal Dell's customers, just as Dell has taken customers from others during the past two decades. And who knows, there may be another young college student among you with the engineering and an entrepreneurial genius needed to discover an even better way to make PCs.

COMPETITION PROMOTES PROSPERITY

Competition motivates businesses to produce efficiently, cater to the views of consumers, and search for innovative improvements.

Competition

The price-taker model highlights the importance of the competitive process. Competition puts pressure on producers to operate efficiently and use resources wisely. Each competing firm will have a strong tendency to produce its products as cheaply as possible. Holding quality constant, pursuit of profit will encourage each firm to minimize the cost of production—to use the set of resources least valued in other uses that can produce the desired output. Firms that fail to keep costs low will be driven from the market.

[4]These and other facts about Dell are from Michael Dell and Andrew Fisher, "It's Crunch Time for Your Competitors," *Financial Times* (London), Sept. 5, 2001, and also from the Dell Web site at `http://wwwl.us.dell.com/content/topics/global.aspxl` corp/pressoffice/en/2004/2004_05_13_rr_000?c=us&l=en&s=corp.

OUTSTANDING ECONOMIST

Friedrich A. von Hayek (1899–1992)

Remarkably, the writings of this 1974 Nobel Prize recipient spanned seven decades. Friedrich von Hayek's work on knowledge and markets helped us understand the competitive process and the fatal defects of central planning. Von Hayek also made important contributions to our understanding of monetary theory, law and economics, business cycles, and capital theory. He is also well known for the idea that market exchanges create an "extended order" where perfect strangers who never meet nonetheless cooperate with one another to achieve their individual goals. The market process described in this chapter is an integral part of that cooperative network.

Similarly, firms in competitive markets will be motivated to discover and produce goods that are valued more highly than the resources required to produce them. Thus, resources are drawn to those uses where they are most productive, as judged by the consumers' willingness to pay. The ability of firms freely to expand or contract their businesses and enter or exit markets means that resources will not be trapped unproductively in a particular industry when they're valued more highly elsewhere.

If firms are going to be successful in competitive markets, they must also be innovative and forward looking. The production techniques and product offerings that lead to success today will not necessarily pass the competitive market test tomorrow. Producers who survive in a competitive environment cannot become complacent. On the contrary, they must be willing to experiment and quick to adopt improved methods.

In competitive markets, business firms must serve the interests of consumers. As Adam Smith noted more than 200 years ago, competition harnesses personal self-interest and channels it into activities that enhance our living standards. Smith stated:

> It is not from the benevolence of the butcher, the brewer, or the baker, that we expect our dinner, but from their regard to their own self-interest. We address ourselves, not to their humanity but to their self-love, and never talk to them of our own necessities, but of their advantages.[5]

Sellers, of course, profit when they supply products valued more highly than the resources needed to make them. But their actions also help the rest of us get more value from those resources than we would otherwise. Paradoxical as it may seem, personal self-interest—a characteristic many view as less than admirable—is a powerful source of economic progress when it is directed by the competitive market process.

LOOKING AHEAD

Consumers often seek variety in product design, style, durability, service, and location. Diferentiating products slightly gives firms some control over the prices they can charge for them. In the next two chapters, we will examine markets in which the firms are price searchers. We will also consider the impact entry barriers have on markets and pricing.

[5]Adam Smith, *An Inquiry into the Nature and Causes of the Wealth of Nations* (1776; Cannan's ed.. Chicago: University of Chicago Press, 1976), 18.

KEY POINTS

- A firm facing a perfectly elastic demand for its product is a price taker. A firm that can raise its price without losing all of its customers (and that must lower its price in order to sell more units) is a price searcher.
- To maximize profit, a price taker will expand its output as long as the sale of additional units adds more to revenues than to costs. Therefore, the profit-maximizing price taker will produce the output level at which marginal revenue (and price) equals marginal cost.
- The price taker's short-run marginal cost curve (above its average variable cost) is its supply curve. The short-run market supply curve is the horizontal summation of the marginal cost curves (above AVC) of the firms in the industry.
- A firm that experiences losses but that anticipates being able to cover its costs in the long run will operate in the short run if it can cover its average variable costs. Conversely, the firm will shut down if it cannot cover its average variable costs. A firm that does not anticipate being able to cover its average total cost even in the long run will minimize losses by immediately going out of business.
- When the market price exceeds the firm's average total cost, it will earn an economic profit. When entry barriers are absent, profits will attract new firms into the industry and stimulate the existing firms to expand. This increasing market supply continues and puts downward pressure on the price until it reaches the level of average total cost, eliminating the economic profit.
- When the market price is less than the firm's average total cost, the resulting losses imply that the resources could be used elsewhere to produce more value. Losses will cause firms to leave the industry or to reduce the scale of their operations. This declining market supply continues and puts upward pressure on price until the firms remaining in the market are able to earn normal returns (zero economic profit).
- As the output of an industry expands in response to rising demand, "fixed" resources like the size of the firm's plant will make it costly for firms to expand output quickly. The diminishing returns and rising marginal costs of firms explains why the short-run market supply curve slopes upward to the right.
- Normally, as industry output expands, rising factor prices push the costs of each firm upward, causing the long-run market supply curve to also slope upward to the right. However, in the long run, firms can alter the size of their plants and other resources that are fixed in the short run. As a result, the market supply curve is generally more elastic in the long run than the short run.
- Firms earn economic profit by producing goods that can be sold for more than the cost of the resources required to produce them. Profit is a reward people get for increasing the value of resources. Conversely, losses are a penalty imposed on those who use resources in a way that reduces their value.
- Competition motivates producers to operate efficiently and heed the views of consumers. Competition and the market process harness self-interest and direct producers toward wealth-creating activities.

CRITICAL ANALYSIS QUESTIONS

***1.** Farmers are often heard to complain about the high costs of machinery, labor, and fertilizer, suggesting that these costs drive down their profits. Does it follow that if, for example, the price of fertilizer fell by 10 percent, farming (a highly competitive industry with low barriers to entry) would be more profitable? Explain.

***2.** If the firms in a price-taker market are making short-run profits, what will happen to the market price in the long run? Explain.

3. "In a price-taker market, if a business operator produces efficiently—that is, if the cost of producing the good is minimized—the operator will be able to make at least a normal profit." True or false? Explain.

4. Suppose the government of a large city levies a 5 percent sales tax on hotel rooms. How will the tax affect (a) prices of hotel rooms, (b) the profits of hotel owners, and (c) gross (including the tax) expenditures on hotel rooms?

"5. If coffee suppliers are price takers, how will an unanticipated increase in demand for their product affect each of the following, in a market that was initially in long-run equilibrium?
   a. The short-run market price of the product
   b. Industry output in the short run
   c. Profitability in the short run
   d. The long-run market price in the industry
   e. Industry output in the long run
   f. Profitability in the long run

*6. Suppose that the development of a new drought-resistant hybrid seed corn leads to a 50 percent increase in the average yield per acre without increasing the cost to the farmers who use the new technology. If the producers in the corn production industry were price takers, what would happen to the following?
   a. The price of corn
   b. The profitability of corn farmers who quickly adopt the new technology
   c. The profitability of corn farmers who are slow to adopt the new technology
   d. The price of soybeans, a substitute product for corn

7. "When the firms in the industry are just able to cover their cost of production, economic profit is zero. Therefore, if demand falls, causing prices to go down even a little bit, all of the firms in the industry will be driven out of business." True or false? Explain.

8. Why does the short-run market supply curve for a product slope upward to the right? Why does the long-run market supply curve generally slope upward to the right? Why is the long-run market supply curve generally more elastic than the short-run supply curve?

*9. How does competition among firms affect the incentive of each firm to (a) operate efficiently (produce at a low per-unit cost) and (b) produce goods that consumers value? What happens to firms that fail to do these two things?

10. Will firms in a price-taker market be able to earn profits in the long run? Why or why not? What determines profitability? Discuss.

*11. During the summer of 1988, drought conditions throughout much of the United States substantially reduced the size of the corn, wheat, and soybean crops, three commodities for which demand is inelastic. Use the price-taker model to determine how the drought affected (a) prices of the three commodities, (b) revenue from the three crops, and (c) the profitability of those farming the three crops.

12. Why is competition in a market important? Is there a positive or negative effect on the economy when strong competitive pressures drive various firms out of business? Discuss.

13. Do business firms in competitive markets have a strong incentive to serve the interests of consumers? Are they motivated by a strong desire to help consumers? Are "good intentions" necessary if individuals are going to engage in actions that are helpful to others? Discuss.

*14. The accompanying table presents the expected cost and revenue data for the Tucker Tomato Farm. The Tuckers produce tomatoes in a greenhouse and sell them wholesale in a price-taker market.
   a. Fill in the firm's marginal cost, average variable cost, average total cost, and profit schedules.
   b. If the Tuckers are profit maximizers, how many tomatoes should they produce when the market price is $500 per ton? Indicate their profits.
   c. Indicate the firm's output level and maximum profit if the market price of tomatoes increases to $550 per ton.
   d. How many units would the Tucker Tomato Farm produce if the price of tomatoes fell to $450 per ton? What would be the firm's profits? Should the firm stay in business? Explain.

COST AND REVENUE SCHEDULES FOR TUCKER TOMATO FARM, INC.

OUTPUT (TONS PER MONTH)	TOTAL COST	PRICE PER TON	MARGINAL COST	AVERAGE VARIABLE COST	AVERAGE TOTAL COST	PROFITS (MONTHLY)
0	$1,000	$500	____	____	____	____
1	1,200	500	____	____	____	____
2	1,350	500	____	____	____	____
3	1,550	500	____	____	____	____
4	1,900	500	____	____	____	____
5	2,300	500	____	____	____	____
6	2,750	500	____	____	____	____
7	3,250	500	____	____	____	____
8	3,800	500	____	____	____	____
9	4,400	500	____	____	____	____
10	5,150	500	____	____	____	____

15. In the accompanying table, you are given information about two firms that compete in a price-taker market. Assume that fixed costs for each firm are $20.

a. Complete the table.

b. What is the lowest price at which firm A will produce?

c. How many units of output will it produce at that price? (Assume that it cannot produce fractional units.)

d. What is the lowest price at which firm B will produce?

e. How many units of output will it produce?

f. How many units will firm A produce if the market price is $20?

g. How many units will firm B produce at the $20 price? (Assume that it cannot produce fractional units.)

h. If each firm's total fixed costs are $20 and the price of output is $20, which firm would earn a higher net profit or incur a smaller loss?

i. How much would that net profit or loss be?

FIRM A

QUANTITY	TOTAL VARIABLE COST	MARGINAL COST	AVERAGE VARIABLE COST
1	$ 24	____	____
2	30	____	____
3	38	____	____
4	48	____	____
5	62	____	____
6	82	____	____
7	110	____	____

FIRM B

QUANTITY	TOTAL VARIABLE COST	MARGINAL COST	AVERAGE VARIABLE COST
1	$ 8	____	____
2	10	____	____
3	16	____	____
4	24	____	____
5	36	____	____
6	56	____	____
7	86	____	____

'Asterisk denotes questions for which answers are given in Appendix B

CHAPTER

10

Price-Searcher Markets with Low Entry Barriers

[T]he price to the price searcher is not determined for him as if by some impersonal market mechanism. Instead he must search out the optimal (wealth-maximizing) price. And, not knowing the demand schedule exactly, he will have to resort to retrial-and-error search processes.

—*Armen A. Alchian and William R. Allen*[1]

Chapter Focus

- What are the characteristics of competitive price-searcher markets? How are price and output determined in such markets?
- Is there a linkage between innovative actions by some firms and business failures by others? Are business failures bad for the economy?
- Can markets be competitive when they consist of only one or two suppliers?
- Why is the role of the entrepreneur left out of economic models? How do the actions of entrepreneurs influence economic progress?
- Why do some economists criticize price-searcher behavior when entry barriers are low, while others like the results?
- Is price discrimination bad?

[1]Armen A. Alchian and William R. Allen, *University Economics*, 2d ed. (Belmont, Calif.: Wadsworth Publishing, 1967), 113.

As we noted in the previous chapter, markets set the price in price-taker markets. Price-taker firms decide what output to produce, but they do not have a pricing decision to make. In this chapter, we will focus on markets in which entry barriers are low, but firms can alter the price charged for their output. If firms reduce their price, they will be able to sell more units. Correspondingly, they can raise their price and lose some, but not all, of their customers. Thus, the firm has to decide what price to charge for the goods and services that it sells.

COMPETITIVE PRICE-SEARCHER MARKETS

Markets that are characterized by (1) low entry barriers and (2) firms that face a downward-sloping demand curve are called **competitive price-searcher markets.**[2] The low entry barriers ensure that these markets are competitive, and the firm's downward–sloping demand curve means that the sellers in these markets have to search for the price and output combination that will maximize their profits.

Competitive price-searcher market
A market in which the firms have a downward-sloping demand curve, and entry into and exit from the market are relatively easy.

In contrast with price-taker markets, in which the firms produce identical products, price searchers produce **differentiated products.** For example, ice cream from Häagen-Dazs is not identical to ice cream from Ben and Jerry's, Breyers, or Baskin-Robbins. The products supplied by the alternative sellers may differ in their design, dependability, location, packaging, and a multitude of other factors. This product differentiation explains why the firms confront a downward-sloping demand curve. Because some consumers are willing to pay more to get the specific product they like best, the firm will not lose all its customers to rivals if it raises its price.

Differentiated products
Products distinguished from similar products by characteristics like quality, design, location, and method of promotion.

Rival firms, however, supply products that are quite similar. Therefore, when a price searcher raises its price, some of its customers will switch to the substitutes. ***Because good substitutes are readily available from other suppliers, the demand curve faced by the firms in competitive price-searcher markets will be highly elastic.***

Given the low entry barriers, new entrants will be attracted if an activity is profitable. Thus, sellers in competitive price-searcher markets face competition from both firms that are already in the market and potential new entrants. Clearly, competition in these markets will be intense. Sometimes economists use the term **monopolistic competition** to describe markets quite similar to those of the competitive price-searcher model, but because there is nothing "monopolistic" about these markets, we believe that this term is misleading. Competitive price searcher is much more descriptive of the conditions in these markets. However, students should be aware that the expression "monopolistic competition" is often used to describe markets very much like those analyzed in this chapter.

Monopolistic competition
A term often used by economists to describe markets characterized by a large number of sellers that supply differentiated products to a market with low barriers to entry. Essentially, it is an alternative term for a competitive mice-searcher market.

Although the price searcher can set the price for its products, the quantity sold at alternative prices will be determined by market forces. Therefore, to find the profit-maximizing price and quantity combination, price searchers must try to estimate not just one market price, but how buyers will respond to the various prices that might be charged. In essence, the firm must try to figure out what the demand curve for its product looks like. Clearly, the price searcher confronts a more complex set of decisions than the price taker. And the complexity does not end with the firm's pricing decision.

For the price searcher, demand is not simply a given. The firm, by changing product quality, style, location, and service (among many other factors), and by advertising, can alter the demand for its products. It can increase demand by drawing customers from rivals if it can convince consumers that its products provide more value. When an airline adopts a more generous frequent-flier program or a soap manufacturer provides "cents-off" coupons, each is trying to make its product a little more attractive than rival products to certain potential customers. The precise effects of these decisions cannot easily be predicted, but they nonetheless can make the difference between profit and loss for the firm. In the real world, most firms occupy the complex and risky territory of the price searcher.

[2] It is important to distinguish between the demand curve faced by the firm and the market demand curve. The competitive price-searcher model focuses on the *firm's* demand curve.

In the highly competitive market for custom-fitted golf clubs, firms use competitive weapons such as style, swingweight, variable length and club-shaft flexibility, advertising, and celebrity endorsements. Firms in this market are price searchers.

© ANNETTE COOLIDGE/PHOTOEDIT

Price and Output in Competitive Price-Searcher Markets

How does a price searcher decide what price to charge and what level of output to produce? For the price searcher, reducing prices in order to expand output and sales has two conflicting influences on total revenue. As **Exhibit 1** illustrates, the increase in sales (from q_1 to q_2) due to the lower price will, by itself, add to the revenue of the price searcher. The price reduction, however, also applies to units *that would otherwise have been sold at a higher price* (P_1, rather than the lower price, P_2). This factor by itself will cause a reduction in total revenue. As the price is gradually lowered to sell additional units, these two conflicting forces will eventually result in marginal revenue (a change in total revenue) that is less than the selling price of the additional units. Because the price of units that could have been sold at the higher price will now be lower, the price searcher's marginal revenue will be less than the price of its product. As Exhibit 1 shows, the marginal revenue curve of the price searcher will always lie below the firm's demand curve.[3] (Remember, the lowercase *d* is used when the reference is to the *firm's* demand curve.)

Any firm can increase profits by expanding output as long as marginal revenue exceeds marginal cost. Therefore, a price searcher will lower price and expand output until marginal revenue is equal to marginal cost.

Exhibit 2 illustrates the profit-maximizing price and output. The price searcher will increase profit by expanding its output to q, where marginal revenue is equal to marginal cost and price P can be charged. Beyond q, the price reduction required to sell additional units will reduce the firm's profit. For any output level less than q (for example, R), a price reduction and sales expansion will add more to total revenues than to total costs. At output R, marginal revenues exceed marginal costs. Therefore, profits will be greater if the price

[3]For a straight-line demand curve, the marginal revenue curve will bisect any line from the y-axis to the demand curve, drawn parallel to the **x-axis.** For example, the MR curve will divide the line P_1F into two equal parts, P_2E and EF.

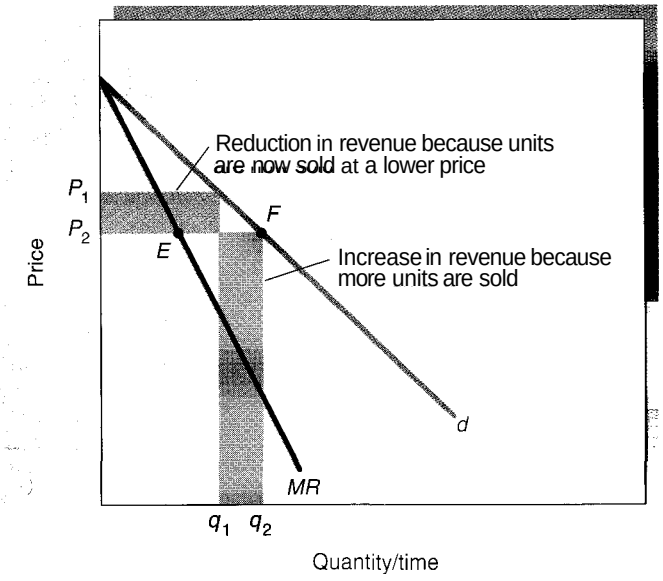


EXHIBIT 1
Marginal Revenue of a Price Searcher

When a firm faces a downward-sloping demand curve, a price reduction that increases sales will exert two conflicting influences on total revenue. First, total revenue will rise because of an increase in the number of units **sold** (from q_1 to q_2). However, revenue losses from the lower price (P_2) on units that could have been sold at a higher price (P_1) will at least partially offset the additional revenues from increased sales. Therefore, the marginal revenue curve will lie inside the firm's demand curve.

is reduced and output can be expanded. On the other hand, if output exceeds q (for example, S), selling additional units beyond q will add more to costs (MC) than to revenues (MR). At output S, the firm will gain by raising the price to P, even though the price rise will result in the loss of some customers. Profits will be maximized by charging price P and producing the output level q, where $MC = MR$.

The firm in Exhibit 2 is making an economic profit. Total revenues $PAq0$ exceed the firm's total costs $CBq0$ at the profit-maximizing output level. Given the low barriers to entry, profits will attract rivals. Other firms will enter this market and attempt to produce a similar product (or provide a similar service).

What will the entry of new rivals do to the demand for the products of the firms already in the market? These new rivals will draw customers away from existing firms and thereby reduce the demand for their output. As long as new entrants expect to make economic profits, additional competitors will be attracted to the market. This entry of new

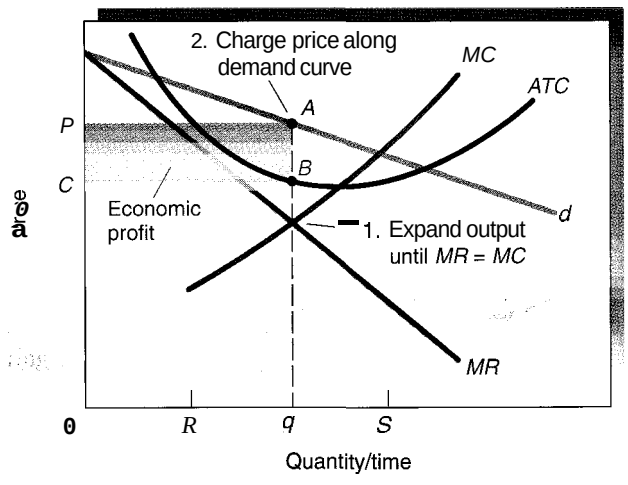


EXHIBIT 2
The Price Searcher's Price and Output

A price searcher maximizes its profits by producing output q, for which $MR = MC$, and charging price P. The firm is making economic profits. What impact will these profits have if this is a typical firm?

ECONOMICS AT THE MOVIES

WARNER BROS/THE KOBAL COLLECTION/HAMIL

You've Got Mail (1998)

Tom Hanks is opening a huge bookstore in Manhattan that could put Meg Ryan's small children's bookstore out of business. Can you show how the demand curve facing Ryan's bookstore would be affected by the opening of the new store in this competitive price-searcher market?

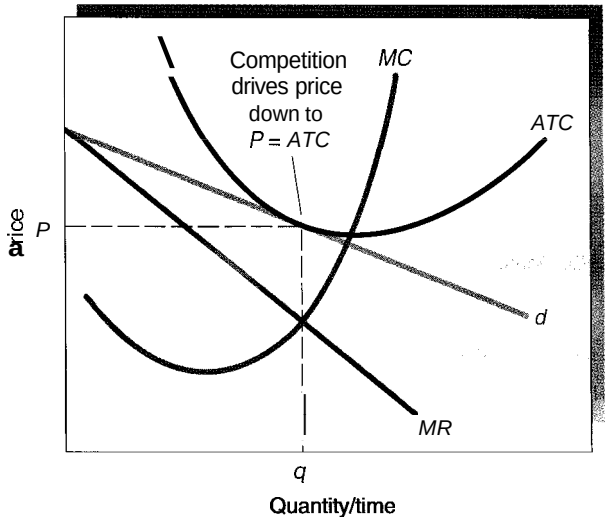
firms into the market will continue until competition among rivals shifts the demand curve in far enough to eliminate the economic profit. As **Exhibit 3** illustrates, when long-run equilibrium is present in a competitive price-searcher market, the price will equal per-unit cost. Firms will produce at the $MR = MC$ output level, but they will be unable to earn economic profits because competitive pressures will force the price down to the per-unit cost level.

If losses are present in a specific market, with time, some of the firms in the market will go out of business. As firms leave, some of their previous customers will buy from other firms. The demand curve facing the remaining firms in the industry will then shift outward until the economic losses are eliminated and the long-run, zero-profit equilibrium illustrated by Exhibit 3 is restored.

Consider the role of profits and losses in competitive price-searcher markets. Profits will attract both new firms and additional investment capital into the market. In turn, the increased availability of the product (and similar products) will drive down the price until the profits are eliminated. Conversely, economic losses will cause competitors to exit the market and investment capital to move elsewhere. This will reduce the supply of the product and eventually make it possible for the firms remaining in the market to charge a price sufficient to cover their unit costs. ***Thus, firms in competitive price-searcher markets can make either economic profits or losses in the short run. But, after long-run adjustments occur, only a normal profit (that is, zero economic profit) will be possible because of the competitive conditions.***

EXHIBIT 3
Competitive Price-Searcher and Long-Run Normal Profit

Because firms are free to enter and exit the market, competition will eventually drive down prices to the level of average total cost for price searchers.



DYNAMIC COMPETITION, INNOVATION, AND BUSINESS FAILURES

Business failures and cutbacks in output and employment are generally reported under the heading of negative news about the economy. Clearly, they are often painful for the parties directly involved. Owners often lose a substantial amount of their wealth, and employees are forced to search for alternative job opportunities. From the standpoint of the entire economy, however, business failures play an important role that is often overlooked: they weed out inefficient, high-cost producers and release resources so they can be employed more productively elsewhere. The assets and workers of firms that fail become available for use by others supplying goods that are more highly valued by consumers. Without this release of resources, the expansion of both profitable firms and the entire economy would be slowed.

Business failures occur for a variety of reasons, including poor management and misjudgment of market conditions. One of the most important sources of business failure is innovative behavior by rival firms. Suppose that an innovative firm introduces a new improved product or comes up with a lower-cost method of supplying a good. The actions of the innovator will reduce the demand for goods supplied by rival firms. Facing intense competition from the innovator, rival firms that are unable to do as well for buyers will experience losses and will eventually be driven out of business. Viewed from this perspective, business failures are often merely the flip side of innovative actions that improve products and reduce costs for consumers.

The great Austrian economist Joseph Schumpeter referred to the competition that occurs when new products, technologies, and better forms of organization are developed as "creative destruction." Dynamic competition of this type is both widespread and highly important in competitive price-searcher markets. It is constantly leading to the expansion of some firms and the contraction of others. Examples abound. In early 2001, Lucent, a high-technology firm, announced that its sales had fallen 28 percent, while rival Nortel Networks experienced a 34 percent increase. Lucent's failure to adopt a new generation of optical-technology production quickly forced it to lay off 16,000 of its employees. As a result, these workers were available to other firms and pursuits.

Important changes were also taking place in the retail sector that year. J. C. Penney announced it would shut forty-seven stores eleven months after it had closed forty-five others. In 2002, Kmart filed for bankruptcy. Meanwhile, Wal-Mart, Costco, and Target were rapidly expanding across the United States, "the way the Huns blew into Europe," as one news report put it?

Numerous businesses come and go. Each year, newly incorporated businesses account for about 10 percent of the total number of businesses. At the same time, approximately 1 percent of businesses file for bankruptcy, and many others close their doors or sell their assets to other, more successful (or at least more optimistic) operators. ***It is important to recognize that business failures do not destroy either the assets owned by the firm or the talents of its workers. Instead, they release these resources for more productive use by other firms.***

CONTESTABLE MARKETS AND THE COMPETITIVE PROCESS

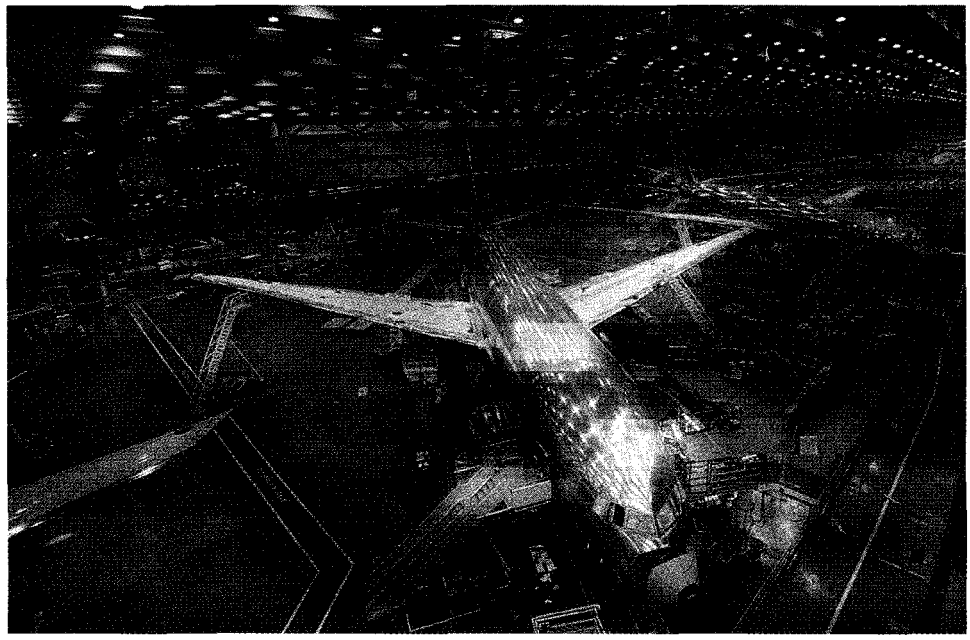
Contestable market
A market in which the costs of entry and exit are low, so a firm risks little by entering. Efficient production and zero economic profits should prevail in a contestable market.

If firms can easily enter a market with little risk of suffering a substantial loss of wealth, competition is likely to be intense, even if there are only a few firms operating within it. Markets in which firms can enter and exit with minimal risk are known as **contestable markets**.[5] Even if entering a contestable market requires a substantial amount of capital, this won't be

[4]Steve Syre and Charles Stein, "Capitalism's Messy Ways Still Come Up Rosy," *Boston Globe*, January 26, 2001, and David Hayes, "Sprint Will Slice 1,100 More Jobs," *Knight-Ridder/Tribune Business News*, June 17, 2004.

[5]The classic article on this topic is William J. Baumol's "Contestable Markets: An Uprising in the Theory of Industry Structure," *American Economic Review 72* (March 1982): 1–15.

Airliners are extremely expensive, so their cost might be a barrier to entry. But airplanes can also be leased and are highly mobile among markets, as are pilots, flight attendants, and office workers. These highly mobile resources allow firms to enter and exit a given transportation market with relative ease. Such a market is contestable.



© DAVE BARTRUFF/INDEX STOCK IMAGERY

a major deterrent if the funds can largely be recovered should the firm need to exit. This will be the case when the assets necessary to compete in a market can be either leased for short time periods or resold at something near their original purchase price.

When airport landing rights and facilities are available, an airline route is a classic example of a contestable market. Consider the case of the airline route between Salt Lake City, Utah, and Albuquerque, New Mexico. Only Delta Airlines provides nonstop service on this route, since there is so little traffic. Furthermore, there would seem to be high barriers to entry, since it takes multimillion-dollar airplanes to compete, as well as facilities for reservations, ticketing, baggage handling, and so on. Given that Delta is the only supplier on the route, one might expect the price to be high and service relatively poor. However, if firms can easily rent facility space at airports, then the barriers to entry are much lower than the equipment costs suggest. The nonstop Salt Lake City–Albuquerque market, for example, can be entered by other airlines simply by shifting aircraft, personnel, and equipment from other routes and locations. Moreover, airlines often rent or lease their planes. This reduces both the amount of initial capital needed to enter a market and the risk of losing it should the venture fail.

Delta is well aware that other firms can easily enter this market. Some of its rivals already serve both Salt Lake City and Albuquerque, and, in fact, provide service between the two via connecting flights. Delta knows that if it raises prices much above cost on the route, other firms will enter the market. The very threat of competition encourages Delta to provide quality service at an attractive price.

In a contestable market, potential competition, as well as actual entry, will discipline sellers. When entry and exit are not expensive, even a firm alone in a market faces the serious prospect of competition. Two important conditions will be present in contestable markets: (1) prices above the level necessary to achieve zero economic profits will not be sustained, and (2) the costs of production will be kept to a minimum. This is because production inefficiencies and prices that are above costs present a profitable opportunity for new entrants. Predictably, profit-seeking rivals will enter and drive the price down to the level of per-unit costs.

The contestable market model has important policy implications. If policy makers are concerned that a market is not sufficiently competitive, they should take a close look at its entry barriers and what might be done to make it contestable. Much of economists' enthusiasm for deregulation can be traced to the fact that regulation often impedes market entry. Many economists believe that deregulating markets to make entry easier can lead to lower prices and better performance than other alternatives, including directly regulating prices.

COMPLEX DECISION MAKING AND THE ENTREPRENEUR

Scientific models simplify conditions in order to help us understand complex relationships. Economic models like the price-taker and price-searcher models are no exception. But these models leave out some important elements of the business decision-making process. Furthermore, they gloss over the complexity of other decisions that must be made by real-world entrepreneurs.

Will profits increase if prices are raised (or lowered)'?To get an answer to this question, real-world business decision makers cannot go into a back room and look at a demand-cost diagram. Instead, they must search for clues, experiment with actual price changes, and interpret what they see, often using a great deal of "seat-of-the-pants" judgment. Our model doesn't reveal precisely how this is done, but it highlights the fact that entrepreneurs are strongly motivated to find the profit-maximizing price. Those that are most successful will at least approximate this objective, and the outcome will be as if they had deliberately chosen the $MR = MC$ price and output combination of our model.

Although the competitive price-searcher model explains how entrepreneurs will react to profits and losses in a specific market, it does not explain how and when new products will be developed or new production techniques applied. How will consumers react to a potential new product? Can it be produced profitably? Will a new production process or alternative technology reduce cost? Can per-unit costs be reduced if the firm offers a different combination of products and services? Here, again, the marginal principle applies: if the change adds more to revenue than it does to cost, it should be made. But how much of a change should be made? Up to the point where $MC = MR$, of course. Identifying this point for each potential change, however, is difficult. Such decisions generally involve an important variable that is omitted from our economic models: entrepreneurship.

The Left-Out Variable: Entrepreneurship

Entrepreneurial judgment is necessary when there is no decision rule that can be applied using only information that is freely available. For this reason, we are unable to incorporate fully the function of the entrepreneur into economic models. There simply is no way to model these complex decisions that involve uncertainty?discovery, and business judgment. All we can do is note the importance of entrepreneurial judgment and recognize that our models are limited because they are unable to capture this critical element of both business success and economic progress.

If we cannot put entrepreneurship into our models, what can we say about its function? One way to answer this question is to consider a generalized job description for an entrepreneurial position. An investor who lacks the desire, or perhaps the skill. to be an entrepreneur but nonetheless wants to be in business, might provide some of the capital to launch a business and hire someone else to act as the entrepreneur. A hypothetical newspaper ad to find such a person might read like this:

> *Wanted: Entrepreneur.* Diverse skills required. Must be (1) alert to new business opportunities and new problems before they become obvious; (2) willing to back judgments with hard work and creative effort before others recognize the correctness of judgments; (3) able to make correct decisions and convince others of their validity, so as to attract additional financial backing; and (4) able to recognize one's own inevitable mistakes and to back away from incorrect decisions without wasting additional resources. Exciting, exhausting, high-risk position. Pay will be very good for success, and very poor for failure.

Entrepreneurship is not for the fainthearted or the lazy. Entrepreneurs are at the center of the action in the real world, even if they do not have a place in most economic models.[6]

[6] For a more complete overview of entrepreneurship and references on the topic, see Mark Crosson, "Entrepreneurship," in ***The New*** *Palgrave: A Dictionary of Economics*, ed. John Eatwell et al. (New York Stockton Press, 1987), 151-153.

APPLICATIONS IN ECONOMICS

Entrepreneurs at the Helm of Some Bizarre Occupations[7]

These innovative people see a need and find a niche in the market

BY DAVID YOUNG (*CHICAGO TRIBUNE*)

Mike Turk lost a coin toss and wound up with a thousand teak trophy bases he couldn't get rid of. So he bolted Army surplus hand grenades to them and sold them for $15 apiece as desk ornaments. Joseph Tokarski, a postal worker contemplating a post-retirement business, developed a machine to clean up Canada goose droppings.

"There are two important dimensions to being an entrepreneur: The first is risk and the second is innovation," said Sumaria Mohan-Neill, professor of entrepreneurship at Roosevelt University's College of Business Administration. "To succeed, a person has to be innovative. Some would call it bizarre," she added.

Turk and Tokarski are among what may be scores of entrepreneurs around Chicago who have found unusual business niches and somehow make their living at bizarre occupations. They manufacture odd products, they provide unusual services and some sell the grotesque. Sometimes these entrepreneurs accidentally find their callings; sometimes out of desperation they gravitate to odd occupations after losing their regular jobs; and in some cases, they are able to turn their hobbies into going business concerns.

"People leverage their knowledge," said Steven C. Michael, business administration professor at the University of Illinois' College of Commerce in Urbana. According to Michael, entrepreneurs have a vision of their mission and are able to sell that to other people, but they also have to be flexible enough to change course.

"Entrepreneurs don't spot markets. They spot needs and assemble the skills to fill them," says Michael, and that's the genesis of Magic Mound Mover, Joseph Tokarski's business. Tokarski said he developed a vacuum system to clean up the backyard scat from his own dogs because no one else in the family would do the chore. The device worked so well that he decided to mount it on a golf cart and go after a problem common in Midwestern open spaces: Canada goose droppings. A little advertising in local newspapers and by word of mouth has resulted in jobs with apartment complexes, park districts and school systems cleaning up after the ubiquitous geese.

Similarly, coming up with a clever idea to solve a problem played a big role for Turk, who was working in the family-owned surplus business Surplus Trading Corp. in Benton Harbor, Mich. Only for Turk, the problem was what to do with 1,000 teak trophy bases he acquired on the toss of a coin that determined their price. "We have all kinds of ways of doing deals in the surplus business," concedes Turk, who over the years has bought and sold 85 dozen women's panties from Victoria's Secret, 60 cases of olive oil and 500 pounds of hard candy. But he couldn't unload the teak at any price, so he decided to turn them into novelty items. He bolted Army surplus hand grenades to the bases and attached plaques that read, "Complaint Department: Take a number." The number is attached to the detonating pin. He has sold hundreds of them.

Turk and Tokarski are lucky ones; the failure rate for entrepreneurs is high, possibly as much as 70 to 80 percent, according to business scholars.

© MIKE TURK/SURPLUS TRADING COMPANY

"Complaint Department: Take a number." (Pull the pin of the hand grenade.)

[7] David Young, "Risky, Odd, Bizarre Even–Entrepreneur's Spot Needs" Chicago *Tribune*, Jan 30, 1999, 1 Repnnted by permission

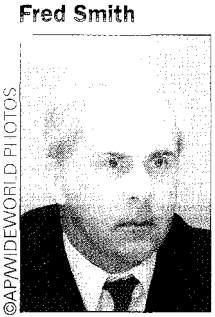
APPLICATIONS IN ECONOMICS

Five Entrepreneurs Who Have Changed Our Lives

Compared to movie stars and athletes, business entrepreneurs seldom command quite as much respect in our society. Often, their contributions are either overlooked or misunderstood. Of course, everyone knows that highly successful entrepreneurs make a lot of money. Remember, however, that trade helps both the buyers and the sellers. Consumers benefit, too, by purchasing the products and services of successful entrepreneurs. If they didn't, they wouldn't buy them.

Entrepreneurship often comes from unusual sources—people who have an ability to think about and institute unconventional ideas. Many entrepreneurs have had an enormous impact on our lives, far more than most of us realize. Let's take a closer look at the impact of five leading entrepreneurs.

Fred Smith



©AP/WIDEWORLD PHOTOS

In 2004, Fred Smith was named "CEO of the Year" by *Chief Executive* magazine. Smith's business career began when he was in college in the 1960s, with the observation that computers were going to become a major part of people's lives. More important, he predicted a logical consequence: if computers were going to replace people for many tasks, the computers would have to be in good working order every day. If something went wrong, they had to be fixed immediately, and parts had to be available right away. As Smith describes it now, computer companies needed a "logistics system that provided the parts and pieces wherever that computer was located whenever it was needed."

Smith developed the idea and described the system in a paper for a business strategy course while studying in the MBA program at Yale. Perhaps because it seemed far-fetched at the time, he reports that the professor gave him a "gentleman's C." In the following decade, however, Smith founded Federal Express—now officially FedEx—and actually created the express delivery system outlined in his paper. Beginning in 1973, FedEx created a giant network of airplane and truck transportation capable of providing overnight delivery around the world. Smith's chief innovation was the development of a transportation system with a "central switch" or "central junction." For FedEx, that was Memphis, Tennessee. When one looks at moving an item from Detroit to Minneapolis via Memphis, it seems inefficient, says Smith. "But when you take all of the transactions on the network together, it's tremendously efficient."

FedEx has revolutionized the express mail and package delivery system, consistently earning a profit. It has absorbed several companies in related businesses, and its annual revenue now exceeds $25 billion. Furthermore, FedEx has forced competitors, from the U.S. Postal Service to United Parcel Service, to become more efficient and make major changes in their operations. In the process, Smith transformed the lives of millions of shippers, who now use the term "FedEx" as a verb and know they can get a package anywhere "absolutely positively overnight"—to quote one of the company's past advertising slogans. He can take great pride in the fact that the market gave him a much higher grade than his old MBA professor.

For more detail, see Gretchen Morgenson, ed., "Fred Smith," in *Forbes Great Minds of Business* (New York: John Wiley, 1997), 35–71; and William J. Holstein, "Leader of the Pack," *Chief Executive* (June 2004).

Pleasant Rowland

©NICK BERARO

Even though she had never run a business or studied in an MBA program, Pleasant Rowland used her life savings to start a company selling a popular line of dolls and books known as The American Girl. Today, the Pleasant Company is a $250-million-a-year business.

Rowland was motivated by a mission: to give girls ages 7 to 12 the chance to experience the love of dolls and the fun of reading she remembered from her childhood. Rowland's idea began to take shape one Christmas season when she tried to find some dolls for her nieces, then ages 8 and 10. The dolls she found—Cabbage Patch babies and Barbie dolls—seemed of mediocre quality, and they failed to evoke "what it meant to be a girl growing up in America." She had already been thinking about American history While visiting Colonial Williamsburg, Virginia, she fell in love with the historic buildings and streets. She began to wonder, "Isn't there some way that I can make the magic of this historic place come alive for little girls?"

Her idea was a subtle one, creating a distinctive line of dolls that reflect authentic American history and serve as appealing role models. She had to find designers and craftspeople who could create dolls that captured the imagination and love of young girls. dolls with often poignant stories—a girl who pioneered in the West or one who escaped slavery in the South Rowland shaped her products as she went along, drawing on all her talents and unswayed by those who thought she was on the wrong track. She had been a teacher and a writer, so she knew how to communicate with children. She knew that her kind of doll wasn't suited for garish Saturday morning television commercials, where she would be battling against giant companies like Hasbro and Mattel. She didn't have enough money to compete with those companies anyway. So she marketed her dolls through a catalog, writing her own copy in a "softer" voice.

(continued)

APPLICATIONS IN ECONOMICS

(continued)

Rowland turned out to be right; there was a void in the market, and her dolls helped fill it. Books and a magazine, *The American Girl,* naturally followed. Rowland put it this way: "I don't think I'm in the doll business or the book business or the direct-mail business. I'm in the little girl business." By following her convictions, she found an audience ready to purchase what she wanted to provide.

Mattel bought into Rowland's approach as well. In 1998, she sold Pleasant Company and American Girl to Mattel for $700 million and took on a new career—philanthropy. By 2004, Pleasant Company (recently renamed American Girl) had sold more than 8 million American Girl dolls and 90 million books.

For more detail, see Gretchen Morgenson, ed., "Pleasant Rowland," in *Forbes Great Minds of Business* (New York: John Wiley, 1998), 121–55, and Julie Sloan, "How We Got Started: Pleasant Rowland," *Fortune Small Business* (October 2002).

William "Bill" Gates III

©ROMMEL PECSON/THE IMAGE WORKS

Even as a child, Bill Gates loved computers. At the early age of fifteen, he earned $20,000 by writing a computer program to help manage traffic. He entered Harvard at the age of seventeen, and, while still an undergraduate, developed BASIC for the first microcomputer, the MITS Altair. Convinced that eventually most offices and homes would have a microcomputer, Gates, along with his friend Paul Allen, created Microsoft in 1975. The focus of the business was the development of software designed to make microcomputers both more useful and more user-friendly.

Gates quickly earned a reputation as a programming genius. When IBM decided to enter the personal computer (PC) market in 1980, it contracted with Gates to provide the basic operating software for its computers, a system now known as MS-DOS. Gates reserved the rights to sell his MS-DOS software to other firms. As IBM (and IBM clones) grew to dominate the PC market during the 1980s, Gates's fortune rose as well. Microsoft developed into the dominant firm in the computer software industry, and today, its programs run more than 90 percent of the world's computers. Microsoft has moved beyond MS-DOS, of course, but its Windows software continues to be among the most popular in the world. Near the end of 2001, Gates's personal fortune was estimated at more than $50 billion. This was after giving $21 billion to his charitable foundation, and losing more than $9 billion that year in the falling stock market.

Probably still the world's richest man in 2004, Bill Gates attributes his success to his workaholic nature. He continues to work long hours and is still actively involved in the operation and strategic decision making of Microsoft, even as his charitable activities have increased. If you have ever used a computer, you have almost certainly benefited from programs Gates created. His remarkable programming genius has changed the way we use personal computers both at home and in the workplace.

For more detail, see Gene N. Landrum, *Profiles of Genius* (Buffalo, N.Y.: Prometheus Books, 1993); and Robert F. Hartley, *Management Mistakes and Successes* (New York: John Wiley, 1994).

Sergey Brin and Larry Page

©STEPHEN HIGGINS/ BLOOMBERG NEWS/LONDON

For years, Internet users have struggled to find billions of bits of information scattered across the World Wide Web. More than anyone, two Stanford computer science graduate students—Sergey Brin and Larry Page—have changed that by getting information to users more quickly and easily on Google.com.

Brin and Page were in their early twenties when they dropped out of graduate school. Using their credit cards and those of their parents and friends, they formed Google, Inc. in a garage in Menlo Park, California. ("Googol" is the number 1 followed by 100 zeros.)

Google's computers are constantly searching the Web, downloading a huge number of files into the company's databases, and updating them regularly. The sophisticated programs Brin and Page developed sort and quickly bring the most important and reliable information to the top of users' screens. Information seekers love it. The keywords they type in for their searches also target the advertising that generates the firm's revenue. Unlike, say, a newspaper, which delivers the same ads to every reader, Google links advertising to the user's particular interests according to the topic being searched. Users see fewer ads, but advertisers get more targeted exposure of their messages.

By 2000, Google had become the world's top search engine, handling 18 million hits a day. The following year, the firm turned its first annual profit: $7 million on sales of $86 million. In 2002, Google was processing 70 million hits per day, and Google, Inc. had expanded its business to sell additional Web-related and wireless business and personal services.

By 2004, first-quarter profits alone were $64 million, and Google announced plans to sell stock to the public. That plan was innovative also. Instead of simply hiring a firm to sell its shares, Page and Brin decided to sell them by means of a "Dutch auction," whereby anyone, including

APPLICATIONS IN ECONOMICS

individual investors, can bid on the shares. The bidder specifies a price as high or low as he or she wants, and the number of shares to be purchased at that price. Market analysts predicted that total share values, and thus the capital raised by Google, would reach $20 billion to $30 billion; a few estimates ran well above that.

The two founders of Google have not yet earned their doctorates, but they *have* earned a fortune. They did it by designing a system that benefits millions of people every day.

For more detail, see "Google Shows Wall Street a Better Way to Sell Stock," Today's Debate, *USA Today*, sec. A, May 7, 2004.

ENTREPRENEURSHIP AND ECONOMIC PROGRESS

The entrepreneurial discovery and development of improved products and production processes is a central element of economic progress.

Entrepreneurship

Discovery and development of improved products and production methods propel economic progress. Think of the new products that have been introduced during the last fifty years; microwave ovens, videocassette recorders, color television sets, personal computers, fax machines, cellular telephone service, DVD and MP3 players, and better coronary artery bypass techniques come to mind. Innovations like these have had an enormous impact on our lives. But no one knows what the next innovative breakthrough will be or precisely which production techniques will minimize per-unit costs. Better ways of doing things do not just happen; they must be discovered and developed by entrepreneurs.

Is the entrepreneur's new, visionary idea the greatest thing since the development of the fast-food chain? Or is it simply another dream that will soon vaporize? In a market economy, it is relatively easy to try new business ideas. A person needs only to win the support of a few investors willing to finance the innovative new product or production technology. However, competition holds entrepreneurs and the investors who support them accountable: their ideas face a "reality check" imposed by production costs and consumers' willingness to pay. Consumers are the ultimate judge and jury. If they do not value an innovative new product or service enough to cover its cost, that product or service will not survive in the marketplace.

Furthermore, today's successful product may not pass tomorrow's competitive test. Therefore, entrepreneurs must be good at anticipating, identifying, and quickly adopting improved ideas, be they their own or others'. They must constantly face the ongoing reality of dynamic change in a competitive world.

Entrepreneurial Decision Making and the Structure, Size, and Scope of Firms

Entrepreneurs must also figure out the type of business structure, scale of operation, and scope of activities that keeps per-unit costs low. Their pursuit of profit encourages them to do so. Unlike other economic systems, a market economy does not mandate or limit the types of firms that are permitted to compete. Any form of business organization is permissible. An owner-operated firm, a partnership, a corporation, an employee-owned firm, a consumer cooperative, a commune, or any other form of business is free to enter the market. To be successful, however, a business structure must be cost-effective. A form of business organization that results in high per-unit cost will be driven from a competitive market by lower-cost rivals.

The same is true for the size of a firm. For some products, a business must be quite large to take full advantage of economies of scale. When per-unit costs decline as output increases, small businesses tend to have higher production costs (and therefore higher prices) than their larger counterparts. When this is the case, consumers interested in getting the maximum value for their money will tend to buy from the lower-priced, larger

firm. In contrast, when personalized service and individualized products are valued highly, small firms, often organized as individual proprietorships or partnerships, are likely to be more cost-effective. It is up to the entrepreneur to discover the scale of operation that best fits the circumstances of each market.

The scope of a business is another variable that requires entrepreneurial decision making. Should a gasoline station stand alone, or should its scope be expanded to include auto repair services? Or should it instead be combined with a convenience store to create more value relative to cost? Unlimited combinations of business activities are possible. Differences in locations and other circumstances will often influence how well each is received. A new and better combination of products and services in the right place can generate a profit, at least until others catch on and provide close substitutes. In contrast, a business that chooses the wrong scope of operation will earn less and is likely to be driven out of the market by rivals offering combinations buyers like better. Once again, buyers have the final say.

To be successful, entrepreneurs must consistently offer consumers at least as much value for their dollar as they can get elsewhere. Put another way, they must figure out how to supply consumers with goods and services that are valued highly relative to their costs. If they do, their actions will create wealth and increase productivity. Many of the activities of entrepreneurs are omitted from our basic economic models. Nonetheless, these people play a *vital* role as agents of economic progress.

EVALUATING COMPETITIVE PRICE-SEARCHER MARKETS

Because the long-run equilibrium conditions in price-taker markets are generally consistent with ideal economic efficiency, it is useful to compare and contrast them with long-run conditions in competitive price-searcher markets. There are both similarities and differences.

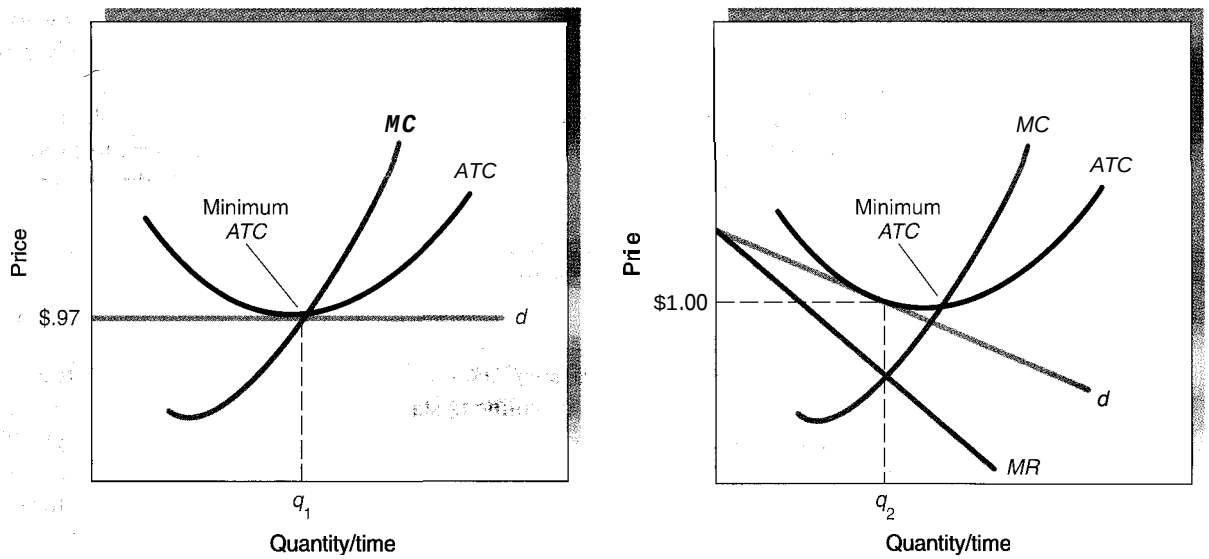
First, let us consider the similarities. Because of the low entry barriers, both price takers and competitive price searchers will have a strong incentive to serve the interests of consumers. Neither will be able to earn long-run economic profits. In the long run, competition will drive prices down to the level of average total cost in both price-taker and competitive price-searcher markets. Furthermore, entrepreneurs in both price-taker and price-searcher markets have a strong incentive to manage and operate their businesses efficiently. In both cases, operational inefficiency will lead to higher costs, losses, and forced exit from the market. Similarly, price takers and competitive price searchers alike will be motivated to develop and adopt new, cost-reducing procedures and techniques because lower costs will mean higher short-run profits (or at least smaller losses).

The responses to changing demand conditions in price-taker and competitive price-searcher markets are also similar. In both cases, any increase in market demand that was not already expected leads to higher prices, short-run profits, expanded output, and the entry of new firms. With the entry of new producers and the expanded output of existing firms, the market supply will increase, putting downward pressure on the price. This process will continue until the market price falls to the level of average total cost, squeezing out all economic profit. Correspondingly, lower demand will lead to lower prices and short-run losses, causing output to fall and some firms to exit. As the market supply declines, prices will rise until the short-run losses are eliminated and the firms remaining in the market can again cover their costs. Thus, profits and losses direct the output decisions and market supply in both price-taker and competitive price-searcher markets.

What are some of the differences between the two market structures? As **Exhibit 4** shows, the price taker confronts a horizontal demand curve, while the demand curve of the price-searcher firm is downward sloping. This is important: it means that the marginal revenue of the price-searcher will be less than, not equal to, the price charged for the product sold. So, when the price searcher expands output until $MR = MC$, the price will still exceed the marginal cost (part b). In contrast, the price charged by a profit-maximizing price taker will equal the marginal cost (part a). In addition, when a price searcher is in long-run equilibrium, the firm's output rate will be less than the rate that minimizes average total cost. The price searcher would have a lower per-unit cost (97 cents rather than $1) if a larger output were produced, but MR would be even less.

EXHIBIT 4
Comparing Price-Taker and Price-Searcher Markets

Here we illustrate the long-run equilibrium conditions of a price taker and a price searcher when entry barriers are low. In both cases, the price is equal to the average total cost, and economic profit is zero. However, because the price searcher confronts a downward-sloping demand curve for its product, its profit-maximizing price exceeds marginal cost, and output is not large enough to minimize average total cost when the market is in long-run equilibrium. For identical cost conditions, the price of the product in a price-searcher market will be slightly higher than in a price-taker market. Some argue that this higher price is indicative of inefficiency, whereas others believe that it merely reflects the higher cost accompanying greater variety and convenience.



(a) Price taker (b) Price searcher

Some argue that the slightly higher prices (and unit costs) in competitive price-searcher markets indicate that these markets allocate goods and services inefficiently. However, it is important to recognize that the higher prices reflect the fact that it is costly to provide the greater product variety and diversity found in competitive price-searcher markets. But a broader range of choices also provides benefits. The greater variety makes it possible for a wider range of consumers to obtain the quality, style, and accompanying service that best fit their preferences and specific situation.

Are the higher prices accompanying the greater diversity worth the cost? This question can't be answered with certainty, but in some cases it is clear that consumers are willing to pay for it. For example, many consumers patronize small, conveniently located stores, even though they know that the same items are available at lower prices at large supermarkets.

The two market structures also differ with regard to advertising: price searchers often advertise, but price takers do not. Like variety, advertising is costly, and in the long run, consumers cover this cost in the form of higher prices. Do advertised products provide enough information and variety to justify the higher cost from the consumer's viewpoint? We cannot know for certain. However, we do know that producers are always free to use low-priced, more uniform, and less-advertised products to lure buyers away from firms that advertise and provide variety at a higher price. Some firms do, and sometimes the strategy works; yet many consumers choose to buy advertised goods and pay the higher prices. This result suggests that many customers do find that advertised products and greater variety are worth more.

Perhaps most important, competitive price-searcher markets provide entrepreneurs with a strong incentive to innovate and discover better ways of doing things. In a dynamic setting with constantly changing opportunities for producers and buyers, this is vitally important for the efficient use of resources.

A SPECIAL CASE: PRICE DISCRIMINATION

So far, we have assumed that all sellers of a product will charge each customer the same price. Sometimes, though, price searchers can increase their revenues (and profits) by charging different prices to different groups of consumers. Businesses like hotels, restaurants, and drugstores often charge senior citizens less than other customers. Students and children are often given discounts at movie theaters and athletic events. Grocery stores commonly give discounts to customers who clip "cents-off' coupons from newspapers or magazines. Colleges often give financial aid (reduced tuition) to students from less-wealthy families. These practices are called **price discrimination**. ***To gain from such a practice, price searchers must be able to do two things: (1) identify and separate ut least two groups with differing elasticities of demand, and (2) prevent those who buy at the low price from reselling to the customers charged higher prices.***

Price discrimination
A practice whereby a seller charges different consumers different prices for the same product or service.

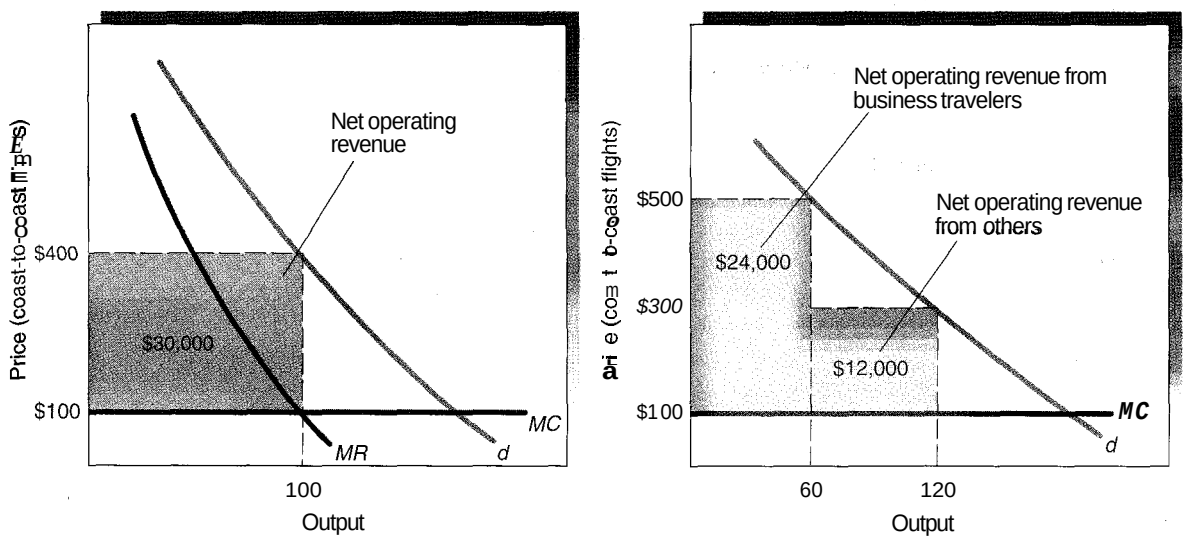
Let us take a closer look at how sellers may gain from price discrimination. Suppose that a seller has two groups of customers: one with an inelastic demand for its product and the other with an elastic demand. An increase in the price charged the first group will increase the total revenue derived from that group. At the same time, a reduction in price will increase revenues derived from the latter. Thus, a seller may be able to increase total revenue and profit by charging the first group a higher price than the second.

The pricing of airline tickets illustrates the potential of price discrimination. The airline industry has found that the demand of business fliers is substantially more inelastic than the demand of vacationers, students, and other travelers. Thus, airlines often charge higher fares to persons who are unwilling to stay over a weekend, who spend only a day or two at their destination, and who make reservations a short time before their flight. These high fares fall primarily on business travelers who are less sensitive to price. In contrast, discount fares are offered to fliers willing to make reservations well in advance, travel during off-peak hours, and stay at their destinations over a weekend before returning home. Such travelers are likely to be vacationers and students, who are highly sensitive to price.

Exhibit 5 illustrates the logic of this policy. Part (a) shows what would happen if a single price were charged to all customers. Given the demand, the profit-maximizing firm ex-

EXHIBIT 5
Price Discrimination

As part (a) illustrates, a $400 ticket price will maximize profits on coast-to-coast flights if an airline charges a single price. However, the airline can do still better if it raises the price to $500 for passengers with a highly inelastic demand (business travelers) and reduces the price to $300 for travelers with a more elastic demand (for example, students and vacationers). When sellers can segment their market, they can gain by (1) charging a higher price to consumers with a less elastic demand and (2) offering discounts to customers whose demand is more elastic.

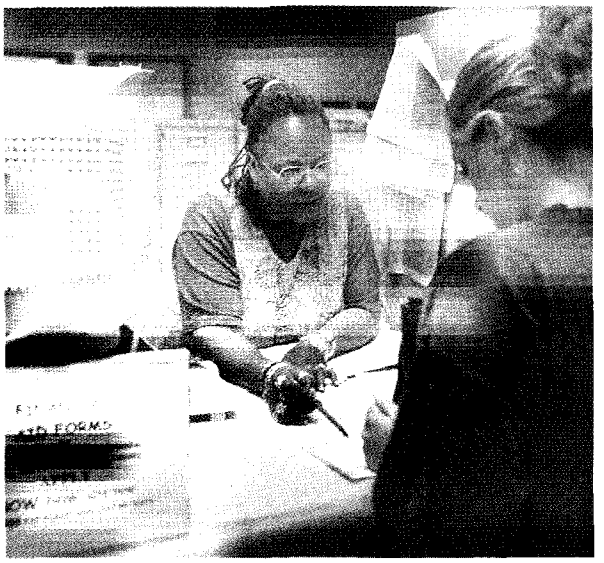


(a) Seller charges a single price

(b) Seller uses price discrimination

pands output to 100, where *MR* equals *MC*. The profit-maximizing price on coast-to-coast flights is $400, which generates $40,000 of revenue per flight. Since the marginal cost per passenger is $100, this provides the airline with net operating revenue of $30,000 with which to cover other costs.

However, as part (b) shows, although the market demand schedule is unchanged, the airline can do even better if it uses price discrimination. When it charges business travelers $500, most of these passengers continue to use the airline, since their demand is highly inelastic. On the other hand, a $100 price cut generates substantial additional ticket sales from price-sensitive vacationers, students, and others whose demands are more elastic. Therefore, with price discrimination, the airline can sell 60 tickets (primarily to business travelers) at $500 and 60 additional tickets to others at $300. Total revenue jumps to $48,000 and leaves the airline with $36,000 ($48,000 minus 120 times the $100 marginal cost per passenger) of revenue in excess of variable cost. Compared to the single-price outcome (part a), the price discrimination strategy expands profit by $6,000.



© ANDREW HOLBROOKE/THE IMAGE WORKS

Students from low-income families are likely to be more sensitive to tuition charges than those from high-income families. Colleges often make scholarships and other forms of financial aid more readily available to low-income students. This price discrimination makes it possible for them to generate larger revenues from tuition than would otherwise be the case.

When sellers can segment their market (at a low cost) into groups with differing price elasticities of demand, price discrimination can increase profits. For *each group,* the seller will maximize profit by equating marginal cost and marginal revenue, This rule will lead to higher prices for groups with a more inelastic demand and lower prices for those with a more elastic demand. Compared to the single-price situation, price discrimination increases profitability because a higher price increases the net revenue from groups with **an** inelastic demand, while a lower price increases the net revenue from price-sensitive customers. With price discrimination, the number of units sold also increases (compare part a with part b) because the discounts provided to price-sensitive groups increase the quantity sold more than the higher prices charged the less-price-sensitive groups reduce sales. Of course, this will work only if the two groups can't trade with each other. Movie theaters sell children's tickets to kids, and the tickets cannot be used by adults. Notice, though, that they don't try to sell popcorn at lower prices to children because popcorn could be too easily resold to adults. Airline tickets are nontransferable, so they cannot be resold. Price discrimination can work in both of these situations because differing categories of buyers can be identified, and trade between them is prevented.

Sometimes price discrimination is subtle. Colleges engage in price discrimination by charging a high standard tuition to get additional revenue from high-income students with a more inelastic demand, while providing low-income students with scholarships based on need (tuition "discounts"). At the University of Maine, for example, when tuition at the undergraduate level ran between $12,998 and $20,738 (depending on the degree). financial aid amounted to 54 percent of tuition for students from low-income families. while for high-income students, it was only 17 percent of tuition. Partial-tuition scholarships given to students whose parents are less wealthy enable a school to attract students who have a more elastic demand. In this way, low-income students are not as frequently priced out of the market by the high standard tuition.

How do buyers fare when a seller can price discriminate? Some buyers pay more than they would if a single, intermediate price were offered. They purchase fewer units, and they are worse off. In contrast, those for whom the price discrimination process lowers the price are better off. Of course, with some products, such as airline transportation, a single buyer might be better off with some purchases and worse off with others.

On balance, however, we can expect that output will be greater with price discrimination than it would be with a single price. The market is not as understocked as it would have been in the absence of the price discrimination. So from an allocative efficiency standpoint, price discrimination gets high marks; it allows more trades, reducing the inefficiency due to price being set above marginal cost. Some of the gains that would accrue to consumers with an inelastic demand are transferred to the price searcher as increased revenue, but additional gains from trade are created by the increased output of goods that would be lost if price discrimination were prevented.

In some markets, there is an additional gain from price discrimination: production may occur that would be lost entirely if only a single price could be charged. With price discrimination, some otherwise unprofitable firms may be able to generate enough additional revenue to operate successfully in the marketplace. For example, some small towns in Montana might not provide enough revenue at a single price to enable a local physician to cover her opportunity costs. However, if she is able to discriminate on the basis of income, charging higher-income patients more than normal rates and lower-income patients less, the resulting revenues from practicing in the small town may enable her to stay in the community. In this case, all residents of the town may be better off because of the price discrimination, since it makes it possible for them to access a local physician. After all, even those being charged the highest prices are not disadvantaged if the price discrimination keeps the physician in town. They are just as able to seek physician services elsewhere as they would have been in the absence of the price-discriminating local doctor. With or without price discrimination, access to competing sellers (or buyers) protects market participants from unfair treatment.

In summary, if potential customers can be segmented into groups with different elasticities of demand and retruding can be controlled at a low cost, sellers can often gain by charging higher prices to those with the less elastic demand (and lower prices to those with a more elastic demand). Price discrimination can also increase the total gains from trade and thereby reduce allocative inefficiency. Sometimes it even allows production where none would have otherwise occurred.

LOOKING AHEAD

In this chapter, we have analyzed the choices facing price searchers and their behavior when barriers to entry are low. The following chapter will analyze the performance of price-searcher markets when the barriers to entry are relatively high.

KEY POINTS

- Firms in price-searcher markets with low barriers to entry face a downward-sloping demand curve. They are free to set the prices for the products that they sell but face strong competitive pressure from existing and potential rivals.
- The mark of the price-searcher market with low barriers to entry is product differentiation. Price searchers use product quality, style, convenience of location, advertising, and price as competitive weapons. Because each firm competes with rivals offering similar products, each confronts a highly elastic demand curve for its products.
- A profit-maximizing price searcher will expand output as long as marginal revenue exceeds marginal cost, lowering its price in the process, until $MR = MC$. The price charged by the profit-maximizing price searcher will be greater than its marginal cost.
- Firms in competitive price-searcher markets can experience either profits or losses in the short run. Profits will attract rival firms into the market until supply increases and the profit-maximizing price falls to the level of per-unit price. Losses will cause firms to exit the market until the price increases enough that the remaining firms can once again cover their per-unit costs. Because of the low entry barriers, the firms in competitive price-searcher markets will earn only normal returns (zero economic profit) when long-run equilibrium is present.
- Competition can come from potential as well as actual rivals. If entry and exit can be arranged at low cost, and if there are no legal barriers to entry, the theory of contestable markets indicates that competitive results will be approximated, even if there are only a few firms actually in the market.
- Although standard economic models do not include the central role of entrepreneurial decision making in a world of uncertainty, economists recognize its importance. Entrepreneurs who discover and introduce

lower-cost production methods and new products that are highly valued relative to cost promote economic progress. Entrepreneurs also have a strong incentive to discover the type of business structure, firm size, and scope of operation that can best keep the per-unit cost of products or services low.

- Competitive price-searcher markets provide more variety but may raise costs relative to price-taker markets because (1) price exceeds marginal cost at the profit-maximizing output level: (2) long-run average cost is not minimized: and (3) advertising is costly. When barriers to entry are low, however, price searchers have an incentive to (1) produce efficiently; (2) undertake production if, and only if, their actions will increase the value of resources used; and (3) be innovative in offering new product options.
- When a price searcher can (1) identify groups of customers that have different price elasticities of demand and (2) prevent customers from retrading the product, price discrimination may emerge. Sellers may be able to gain by charging higher prices to groups with a less elastic demand and lower prices to those with a more elastic demand. The practice generally leads to a larger output and more gains from trade than would otherwise occur.

CRITICAL ANALYSIS QUESTIONS

1. Price searchers can set the prices of their products. Does this mean that they will charge the highest possible price for their products? What price will maximize the profits of a price searcher? How will the firm's marginal cost compare with its price at the profit-maximum output?

2. Since price searchers can set their prices, does this mean that their prices are unaffected by market conditions? In price-searcher markets with low barriers to entry, will the firms be able to make economic profit in the long run? Why or why not? What do competitive price searchers have to do to make economic profit?

***3.** What determines the *variety* of styles, designs, and sizes of different products? Why do you think there are only a few different varieties of toothpicks but lots of different types of napkins on the market?

***4.** How would imposing a per-unit tax of $2,000 on each new U.S. automobile affect (a) higher-quality, higher-priced cars (those selling for more than the median price) compared to (b) lower-quality, lower-priced cars? What would happen as a result to the average *quality* of automobiles if the proceeds of the tax were used to subsidize a government-operated lottery?

5. What must an entrepreneur do in order to introduce a new innovative product? What determines whether the new product will be a success or failure?

6. Is quality and style competition as important as price competition? Would you like to live in a country where government regulation restricted the use of quality and style competition? Why or why not? Do you think you would get more or less for your consumer dollar under restrictions like these? Discuss.

***7.** Suppose that a price searcher is currently charging a price that maximizes the firm's total revenue. Will this price also maximize the firm's profit? Why or why not? Explain.

***8.** **a.** What determines whether corporations, individual proprietorships, employee-owned firms, consumer cooperatives, or some other form of business structure will dominate in a market?

b. What determines whether small firms, medium-size firms, or large firms will dominate in a market?

9. Would our standard of living be higher if the government "bailed out" troubled businesses? If a firm goes out of business, what happens to the firm's assets, workers, and customers? Are business failures bad for the economy? Why or why not?

***10.** Suppose that a group of investors wants to start a business operated out of a popular Utah ski area. The group is considering either building a new hotel complex or starting a new local airline serving that market. Each new business would require about the same amount of capital and personnel hiring. The group believes each endeavor has the same profit potential. Which is the safer (less likely to result in a substantial capital loss) investment? Why? Is there an offsetting advantage to the other investments?

11. Is price discrimination harmful to the economy? How does price discrimination affect the total amount of gains from exchange? Explain. Why do colleges often charge students different prices, based on their family income?

***12.** "When competition is really severe, only the big firms survive. The little guy has no chance." True or false? Explain.

13. What is the primary requirement for a market to be competitive? Is competition necessary for markets to work well? Why or why not? How does competition

influence the following: (a) the cost efficiency of producers. (b) the quality of products, and (c) the discovery and development of new products? Explain your answers.

"14. What keeps McDonald's, Wal-Mart, General Motors, or any other business firm from raising prices, selling shoddy products, and providing lousy service?

15. "The superiority of the competitive market is the positive stimuli it provides for constantly improving efficiency, innovating, and offering consumers diversity of choice." This quotation is from Alfred Kahn, the architect of transportation deregulation during the 1970s. Evaluate the statement. Is it true? Discuss.

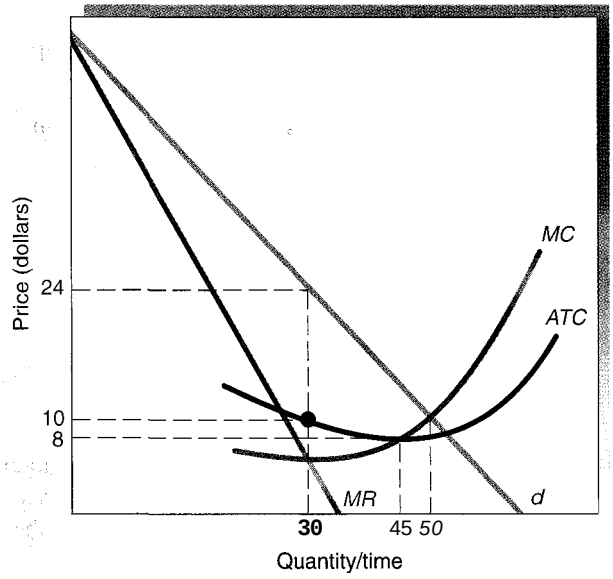
16. The accompanying graph shows the short-run demand and cost situation for a price searcher in a market with low barriers to entry.

a. What level of output will maximize the firm's profit level?

b. What price will the firm charge?

c. How much revenue will the firm receive in this situation? How much is total cost? Total profit?

d. How will the situation change over time?



***17.** Rod N. Reel owns a dealership that sells fishing boats in an open, price-searcher market. To develop his pricing strategy, Rod hired an economist to estimate his demand curve. Columns (1) and (2) of the following chart provide the data for the expected weekly quantity demanded for Rod's fishing boats at alternative prices. Rod's marginal (and average) cost of supplying each boat is constant at $5,000 per boat no matter how many boats he sells per week in this range. This cost includes all opportunity costs and represents the economic cost per boat.

a. Find Rod's economic profits at each alternative price by calculating the difference between total revenue and total cost.

b. Find Rod's marginal revenue and marginal cost from the sale of each additional boat.

c. If Rod wants to maximize his profits, what price should he charge per boat?

d. How many boats will Rod sell per week at the profit-maximizing price?

e. What will Rod's profits be per week at this price and sales volume?

f. At the price and sales level where profits are maximized, has Rod sold all boats that have higher marginal revenue than marginal cost?

g. If Rod's profits are typical of all firms in the boat sales business, what might be expected to happen in the future? Will more boat dealers open in the area, or will some of the existing ones go out of business? What will happen to the profitability of the boat dealers in the future once the entry/exit has occurred?

h. Challenge Question: Recall the relationship between elasticity of demand, price changes, and their impact on total revenues. As Rod lowers his price from $9,000 to $5,000, his total revenues keep increasing. Is demand in this price range elastic, inelastic, or unit elastic? When Rod lowers his price from $5,000 to $4,000, his total revenues stay the same. Is demand in this price range elastic, inelastic, or unit elastic? Can you guess what might happen at prices below $4,000? Explain.

*Asterisk denotes questions for which answers are given in Appendix B.

PRICE OF FISHING BOATS (1)	NUMBER OF FISHING BOATS SOLD PER WEEK (2)	TOTAL REVENUE PER WEEK (3)	TOTAL COST PER WEEK (4)	ECONOMIC PROFIT PER WEEK (5)	MARGINAL REVENUE (6)	MARGINAL COST (7)
$9,000	0	____	____	____	____	____
8,000	1	____	____	____	____	____
7,000	2	____	____	____	____	____
6,000	3	____	____	____	____	____
5,000	4	____	____	____	____	____
4,000	5	____	____	____	____	____

Price-Searcher Markets with High Entry Barriers

CHAPTER 11

Chapter Focus

- What are the barriers to entry that protect some firms against competition from potential rivals?
- What is a monopoly? Does it guarantee the ability to make a profit?
- What is an oligopoly? When are oligopolists likely to collude? Why is it impossible to construct a general theory of output and price for an oligopolist?
- Will competition keep large firms in check in markets that have high barriers to entry? What problems may arise in such markets?
- What policy alternatives might improve the operation of markets with high barriers to entry?

"When the monopolist raises prices above the competitive level in order to reap his monopoly profits, customers buy less of the product, less is produced, and society as a whole is worse off. In short, monopoly reduces society's income."

". . .[C]ompetition is a tough weed, not a delicate flower."

—*George Stigler*[1]

[1]Both quotes are from George J. Stigler, "Monopoly" in *The Concise Encyclopedia of Economics*, ed. David R. Henderson (Indianapolis, Ind.: Liberty Fund, Inc., Library of Economics and Liberty), http://www.econlib.org/library/Enc/Monopoly.html (accessed August 31, 2004).

In the previous two chapters, we analyzed the way firms behave in competitive markets with low barriers to entry. We now turn to the analysis of firm behavior when entry barriers are high and there are few, if any, rival firms offering the same or similar products. This chapter focuses on factors that make entry into a market more difficult, and how this can reduce the ability of markets to keep firms in check. We will also consider policy alternatives that might improve the efficiency of these markets.

WHY ARE ENTRY BARRIERS SOMETIMES HIGH?

What makes it difficult for potential competitors to enter a market? Under certain conditions, four factors can be important: economies of scale, government licensing, patents, and control over an essential resource. Let's consider each.

Economies of Scale

As we have learned, when the fixed costs in an industry are large, bigger firms can achieve lower average total per-unit costs than smaller firms can over the full range of output consumers are willing to buy. **As** our prior analysis of contestable markets indicated, however, economies of scale and high fixed costs are not a significant barrier to entry when the assets needed to enter the market can be leased, transferred to another location, or resold later without a major loss of value.

In some markets, however, this will not be the case. Airline routes may be contestable, but airplane manufacturing is not because many items of equipment are specialized for the production of a specific airplane. The cost of that equipment is a sunk cost. Except in its current use, the specialized equipment is worth little. If necessary, the incumbent firm will continue operating for a long period of time as long as it can cover its variable (operating) costs. Potential new entrants will recognize that this incentive structure will make it extremely difficult to displace the current incumbent. Thus, they will be reluctant to attempt entry, even if the current incumbent is earning economic profit. Under these circumstances, a dominant firm will tend to emerge in the industry, and if the market is not contestable, the cost advantages resulting from its size will protect it from potential rivals.

Government Licensing and Other Legal Barriers to Entry

Legal barriers are the oldest and most effective method of protecting a business firm from potential competitors. Kings once granted exclusive business rights to favored citizens or groups. Tariffs that raise the price of imported goods and quotas that limit the quantity of imports have restricted entry and reduced the competitiveness of various markets for centuries. Today, governments continue to establish barriers that restrict the right to buy and sell goods. To compete in certain parts of the communications industry in the United States (for example, to operate a radio or television station), one must obtain a government franchise. Similarly, local governments often grant exclusive franchises for the operation of cable television systems. Each of these measures reduces competition and protects the profits of favored firms.

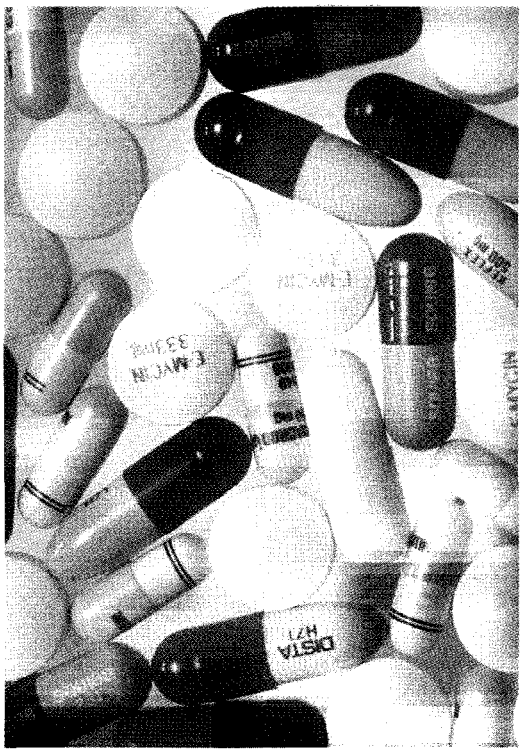
Licensing
A requirement that one obtain permission from the government in order to perform certain business activities or work in various occupations.

Licensing can also limit entry and reduce competitiveness. States and cities often require the operators of liquor stores, hairstyling shops, taxicabs, funeral homes, and many other businesses to obtain a license. Sometimes these licenses cost little and are designed to ensure certain minimum standards. Often, however, they are costly to obtain and a major deterrent to the entry of potential rivals.

Patents

Most countries have patent laws to give inventors a property right to their inventions. **A** patent grants the owner the exclusive legal right to the commercial use of a newly invented product or process for a limited period of time, twenty years in the United States. Others are

not allowed to copy the product or procedure without the permission of the patent holder. For example, when a pharmaceutical company is granted a patent for a newly developed drug, other potential suppliers must obtain permission from the patent holder before producing and selling the product. Others might be able to supply the drug more economically. If they want to do so, however, they will have to purchase the production and marketing rights from the originating firm as long as the patent is in effect.



©MARK ANTMAN/THE IMAGEWORKS

A firm that develops a new drug can use patent protection to restrain production by others for twenty years from the time the patent is issued. Although consumers will pay higher prices than if open competition were permitted, they may benefit from additional investment in research and more rapid development of new products by firms seeking a market with patent protection.

Both costs and benefits for consumers come with a patent system. The entry barrier created by the grant of a patent generally leads to higher consumer prices for products that have already been developed. On the positive side, however, patents increase the potential returns to inventive activity, thus encouraging scientific research and technological improvements. According to a recent study, the average cost of developing and testing a new pharmaceutical drug is approximately $900 million.[2] Without the profit potential accompanying patents, the incentive to engage in research and develop new products would be slowed.

Control over an Essential Resource

Control over a resource essential for production of a product may also insulate a firm from direct competitors. An example often cited is the Aluminum Company of America, which before World War II controlled the known supply of bauxite conveniently available to American firms. Without this critical raw material, potential competitors could not produce aluminum. Over time, however, other supplies of bauxite were discovered, and the Aluminum Company of America lost its advantage. Resource monopolies are seldom complete. Profit opportunities provide challengers with an incentive to search for mineral deposits, new technologies, and substitute resources. Over time, they are usually found.

Barriers to entry like the different ones we've just described are often temporary, but they do exist. Let's see what happens when, at least temporarily, there is a barrier to entry high enough to limit the market to only one seller.

CHARACTERISTICS OF A MONOPOLY

The word *monopoly,* derived from two Greek words, means "single seller." We will define **monopoly** as a market characterized by (1) high barriers to entry and *(2)* a single seller of a well-defined product for which there are no good substitutes. Even this definition is ambiguous, because "high barriers" and "good substitutes" are both relative terms. Are the barriers to entry into the automobile or steel industries high? Many observers would argue that they are. After all, it takes a great deal of capital to operate at the least-cost scale of output in both of these industries. However, there are no *physical* or *legal* restraints that prevent an entrepreneur from producing automobiles or steel. If price is well above cost and profit potential

Monopoly
A market structure characterized by (1) a single seller of a well-defined product for which there are no good substitutes and (2) high barriers to the entry of any other firms into the market for that product.

[2] Tufts Center for the Study of Drug Development, "Total Cost to Develop a New Prescription Drug, Including Cost of Post-Approval Research, Is $897 Million," News Release May 13, 2003, downloaded July 6, 2004, at: http://csdd.tufts.edu/NewsEvents/RecentNews.asp?newsid529.

is present, it should not be too difficult to find the necessary investment capital. Thus, some would argue that entry barriers into these industries are not particularly high.

"Good substitute" is also a subjective term. There is always some substitutability among products, even those produced by a single seller. Is a letter a good substitute for a telephone or e-mail message? For some purposes — correspondence between law firms, for example — a letter delivered by mail is a very good substitute. In other cases, when the speed of communication and immediacy of response are important, the telephone and e-mail are far superior forms of communication. Are there any good substitutes for electricity? Most of the known substitutes for electric lighting (candles, oil lamps, and battery lights, for example) are inferior to electric lights in most uses. Natural gas, fuel oil, and wood, though, are often excellent substitutes for electric heating. Also, electric utilities have found that many of their largest industrial customers can generate their own electricity, sometimes at low cost, in conjunction with other fuel-burning operations.

Monopoly, then, is always a matter of degree. Only a small fraction of all markets are served by only one seller. In a few markets, governments have allocated specific markets to a single seller. This is the case with cable television and providers of electricity in many communities. The monopoly model we discuss next will illuminate the operation of these markets. It will also help us understand markets in which there are just a few sellers and little active rivalry. When there are only two or three producers in a market, firms may seek to collude rather than compete, and thus, together, they may behave much like a monopoly.

Price and Output under Monopoly

Suppose you invent, patent, and produce a microwave device that locks the hammer of any firearm in the immediate area. This fabulous invention can be used to immobilize the weapons of terrorists or armed robbers. Since you own the exclusive patent right to the firearm lock device, you are not concerned about a competitive supplier in the foreseeable future. Although other products are competitive with your invention, they are poor substitutes. In short, you are a monopolist.

What price should you charge for your product? Because you are the only seller of this device, the demand for your product is also the market demand curve. It will be downward sloping because consumers will cut back on their purchases as the price of the firearm lock increases. As a price searcher, you will have to search out the most profitable price to charge. An expansion in output will increase profit as long as the production and sale of additional units add more to revenue than to cost. ***Like other price searchers, the monopolist will expand its output until marginal revenue equals marginal cost. This profit-maximizing output rate can be sold at the price indicated on the firm's demand curve.***

Exhibit 1 graphically illustrates how a monopolist derives the profit-maximizing output rate.[3] The firm will continue to expand output as long as marginal revenue exceeds marginal cost. Therefore, output will be expanded to Q, where $MR = MC$. The monopolist will be able to sell the profit-maximizing output Q at price P, the height of the demand curve at Q. At any output less than Q, the benefits (marginal revenue) of producing the *additional* units will exceed their costs. In this range, the monopolist will gain by reducing the price and expanding output toward Q. For any output greater than Q, the monopolist's costs of producing additional units will be greater than the benefits (marginal revenue). Producing these units will reduce profits. Output rate Q and price P will maximize the firm's profit.

Exhibit I also depicts the profits of a monopolist. At output Q and price P, the firm's total revenue is equal *to PAQO,* the price times the number of units sold. The firm's total cost would be *CBQ0,* the average per-unit cost multiplied by the number of units sold. The firm's profits are merely total revenue less total cost, the shaded area of Exhibit 1.

Exhibit 2 provides a numeric illustration of decision making by a profit-maximizing monopolist. At low output rates, marginal revenue exceeds marginal cost. The monopolist will continue expanding output as long as MR is greater than MC. Thus, an output rate of

[3]In this chapter we assume that firms are unable to use pnce discrimination to increase their revenues. If they could, the analysis of price discrimination from the previous chapter would help describe their decisions.

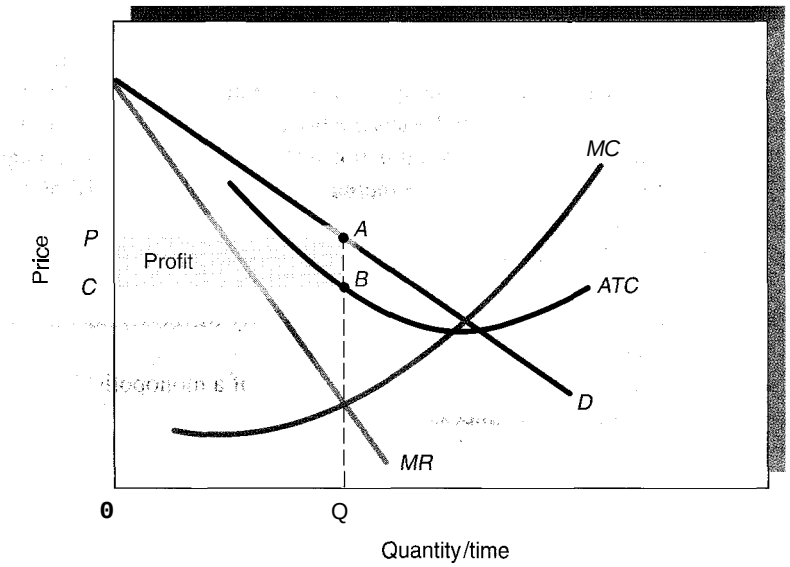


EXHIBIT 1
The Short-Run Price and Output of a Monopolist

The monopolist will reduce its price and expand its output as long as *MR* exceeds *MC*. Output Q will result. When the price exceeds the average total cost at any output level, a profit will be earned.

eight units per day will be chosen. *(Note:* If tiny portions of a unit could be produced and sold, then production would increase to where $MR = MC$.*)* Given the demand for the product, the monopolist can sell eight units at a price of $17.25 each. Total revenue will be $138, compared to a total cost of $108.50. The monopolist will make a profit of $29.50. The profit rate will be smaller at all other output rates. For example, if the monopolist reduces the price to $16 in order to sell nine units per day, revenue will increase by $6. However, the marginal cost of producing the ninth unit is $6.25. Since the cost of producing the ninth unit is greater than the revenue it brings in, profits will decline.

When high barriers to entry are present, they will insulate the monopolist from direct competition with rival firms producing a similar product. In markets with high entry barriers, monopoly profits will not attract—at least not quickly—rivals who will expand supply, cut prices, and spoil the seller's market.

EXHIBIT 2
Profit Maximization for a Monopolist

RATE OF OUTPUT (PER DAY) (1)	PRICE (PER UNIT) (2)	TOTAL REVENUE (1) × (2) (3)	TOTAL COST (PER DAY) (4)	PROFIT (3) − (4) (5)	MARGINAL COST (6)	MARGINAL REVENUE (7)
0	—	—	$50.00	$–50.00	—	—
1	$25.00	$25.00	60.00	–35.00	$10.00	$25.00
2	24.00	48.00	69.00	–21.00	9.00	23.00
3	23.00	69.00	77.00	–8.00	8.00	21.00
4	22.00	88.00	84.00	4.00	7.00	19.00
5	21.00	105.00	90.50	14.50	6.50	17.00
6	19.75	118.50	96.75	21.75	6.25	13.50
7	18.50	129.50	102.75	26.75	6.00	11.00
8	17.25	138.00	108.50	29.50	5.75	8.50
9	16.00	144.00	114.75	29.25	6.25	6.00
10	14.75	147.50	121.25	26.25	6.50	3.50
11	13.50	148.50	128.00	20.50	6.75	1.00
12	12.25	147.00	135.00	12.00	7.00	–1.50
13	11.00	143.00	142.25	.75	7.25	–4.00

Protected by high entry barriers, a monopolist may be able to continue earning a profit, even in the long run. Does this mean that monopolists can charge as high a price as they want? Monopolists are often accused of price gouging. In evaluating this charge, however, it is important to recognize that, like other sellers, monopolists will seek to maximize *profit,* not *price.* Consumers will buy less as the price increases. Thus, a higher price is not always best for a monopolist. Exhibit 2 illustrates this point. What would happen to the profit of the monopolist if the price were increased from $17.25 to $18.50? At the higher price, only seven units would be sold, and total revenue would equal $129.50. The cost of producing seven units would be $102.75. Thus, when the price is $18.50 and output seven units, profit is only $26.75—less than it would be at the lower price ($17.25) and larger output (eight units). The highest price is not usually the best price for the monopolist. Sometimes lower prices will mean more profit.

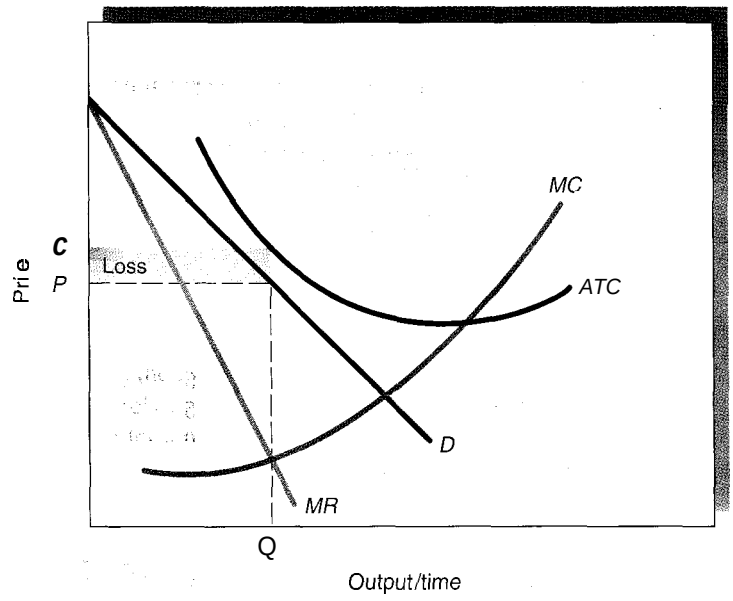
Will a monopolist always earn a profit? The profitability of a monopolist is limited by the demand for the product that it produces. In some cases, a monopolist—even one protected by high barriers to entry—may be unable to sell its product for a profit. For example, there are thousands of clever, patented items that are never produced because demand-cost conditions are not favorable. **Exhibit 3** illustrates this possibility. Here, the monopolist's average total cost curve is above its demand curve at every level of output. Even when operating at the $MR = MC$ rate of output, economic losses (shaded area) will occur. While output Q could be sold at price P, this price is too low to cover the per-unit cost of the monopolist. This model explains why many small towns lack even a single bookstore, gourmet coffee shop, or any number of other specialty stores. Under these circumstances, not even a monopolist would want to operate—at least not for long.

The Monopoly Model and Decision Making in the Real World

Until now, we have proceeded as if monopolists always knew exactly what their revenue and cost curves looked like. In reality, a monopolist cannot be sure of the ever-changing demand conditions for a product. How many sales will be lost if the price is raised? How many sales will be added if the price is lowered? Monopolists often experiment with price changes, attempting to find the price at which profit will be maximized. Thus, the revenue and cost data illustrated in Exhibits 1, 2, and 3 represent *expected* revenues and costs associated with various output levels. To complicate matters further in real life, the monopolist searching for additional profits can undertake quality changes of various sorts. Buyer re-

EXHIBIT 3
How a Monopolist Can Lose Money

Even a monopolist will lose money if its average total cost curve lies everywhere above the demand curve.



actions to potential quality changes and resulting shifts in demand and cost must also be considered by the profit-seeking monopolist.

A monopolist, like other business decision makers, seldom calculates what we have called demand, marginal revenue, and cost curves. With imperfect information, the profit-maximizing price can only be approximated. However, the monopolist has a strong incentive to find the profit-maximizing price and output rate and, when that occurs, the outcome will be as if the $MR = MC$ rule had been used to maximize profit.

As we have indicated, it is difficult for a monopolist to predict demand conditions and consumer response to quality and price changes. But life is even more complex for the next class of price searchers we consider: large firms with few rivals in a market protected by high barriers to entry.

THE CHARACTERISTICS OF AN OLIGOPOLY

In the United States, most of the output in certain industries, such as the automobile, cigarette, and aircraft industries, is produced by five or fewer dominant firms. These industries are characterized by **oligopoly**. *Oligopoly* means "few sellers." ***The distinguishing characteristics of an oligopolistic market are (1) a small number of rival firms, (2) interdependence among the sellers because each is large relative to the size of the market, (3) substantial economies of scale, and (4) high entry barriers to the market.***

Oligopoly
A market situation in which a small number of sellers constitutes the entire industry It is competition among the few.

Interdependence among Oligopolistic Firms

Since the number of sellers in an oligopolistic industry is small, the supply decisions of one firm will significantly influence the demand, price, and profit of rivals. This adds to the complexity of the firm's decision making. **A** firm that is deciding what price to charge, output to produce, or quality of product to offer must consider not only the currently available substitutes, but also the potential reactions of rival producers. The business decisions of each firm depend on the policies it expects its major rivals to follow because those will influence the demand for all the rivals' products.

Substantial Economies of Scale

In an oligopolistic industry, large-scale production (relative to the total market) is generally required to achieve minimum per-unit cost. Economies of scale are present, so a small number of large-scale firms can produce enough of the product to meet the entire market demand. Using the automobile industry in the recent past as an example, **Exhibit 4** illustrates the importance of economies of scale as a source of oligopoly. It has been estimated

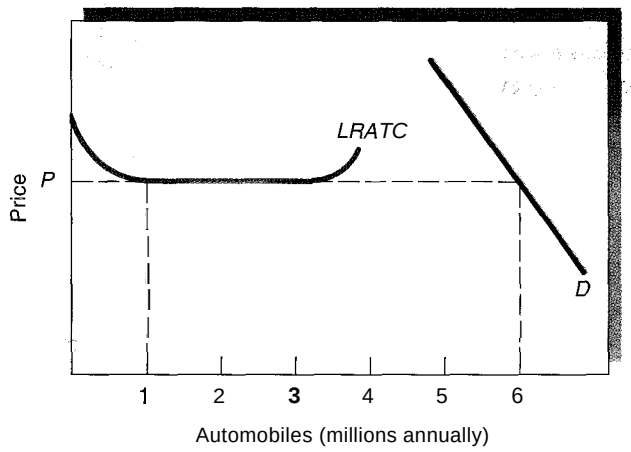


EXHIBIT 4
Economies of Scale and Oligopoly

An oligopoly exists in the automobile industry because firms do not fully achieve the lower costs of large-scale operations until they are able to produce approximately one-sixth of the total market.

that each firm must produce approximately 1 million automobiles annually before its per-unit cost of production is minimized. However, when the selling price of automobiles is barely sufficient to cover costs, the total quantity demanded from these producers is only 6 million. To minimize costs, then, each firm must produce at least one-sixth (1 million of the 6 million) of the output demanded. If this is true, the industry can support no more than five or six domestic firms of cost-efficient size.

Significant Barriers to Entry

As with monopoly, barriers to entry limit the ability of new firms to compete effectively in oligopolistic industries. Except where a market is contestable, economies of scale are probably the most significant entry barrier protecting firms in an oligopolistic industry. A potential competitor may be unable to start out small and gradually grow to the optimal size, since it must gain a large share of the market before it can minimize its per-unit cost. The manufacture of refrigerators, diesel engines, and automobiles seems to fall into this category. Other factors, including patent rights, control over an essential resource, and government-imposed entry restraints, can also prevent new competitors from entering profitable oligopolistic industries. The presence of high entry barriers is what distinguishes oligopoly from a competitive price-searcher market.

Identical or Differentiated Products

The products of sellers in an oligopolistic industry may be either similar or differentiated. When firms produce identical products, like milk or gasoline, there is less opportunity for nonprice competition. On the other hand, when rival firms produce differentiated products they are more likely to use style, quality, and advertising as competitive weapons. Each firm attempts to convince buyers that other products would be poor substitutes.

PRICE AND OUTPUT UNDER OLIGOPOLY

Unlike a monopolist or a price taker, an oligopolist cannot determine the product price that will deliver maximum profit simply by estimating its own costs and the *existing* market demand. The demand facing an oligopolistic firm depends also on the pricing behavior of its close rivals. Suppose, for example, that General Motors is considering how to price next year's new Chevrolet model when it comes on the market. The number of Chevrolets it can sell at each price will depend not only on buyer preferences but also on the prices (and quality) of substitutes available from Ford, Toyota, and other GM rivals. What pricing strategy will each rival use if GM puts a low price on its new model? How will each rival firm react to a higher GM price?

Economics cannot predict these complex interactions among rivals without making some strong, simplifying assumptions about how each firm reacts when another firm makes a change in price or quality. These simplifying assumptions will not be realistic for the general case, and therefore we cannot precisely predict the price and output that will emerge in an oligopoly. We can, however, zero in on a potential range of prices and the factors that will determine whether prices in the industry will be high or low relative to costs of production.

Consider a typical oligopolistic industry in which seven or eight rival firms produce the entire market output because substantial economies of scale are present. The firms produce nearly identical products and have similar costs of production. **Exhibit 5** depicts the market demand conditions and long-run costs of production of the individual firms for such an industry.

What price will prevail? We can answer this question for two extreme cases. First, suppose that each firm sets its price independently of the other firms. There is no **collusion** (no agreement among the firms to limit output and keep the price high), so each competitive firm is free to seek additional customers and profits by offering buyers a slightly better deal than its market rivals. Under these conditions, the market price would be driven down to P_c. Firms would be just able to cover their per-unit costs of production. If a single firm raised its price,

Collusion
Agreement among firms to avoid various competitive practices, particularly price reductions. It may involve either formal agreements or merely tacit recognition that competitive practices will be self-defeating in the long run. Tacit collusion is difficult to detect. In the United States, antitrust laws prohibit collusion and conspiracies to restrain trade.

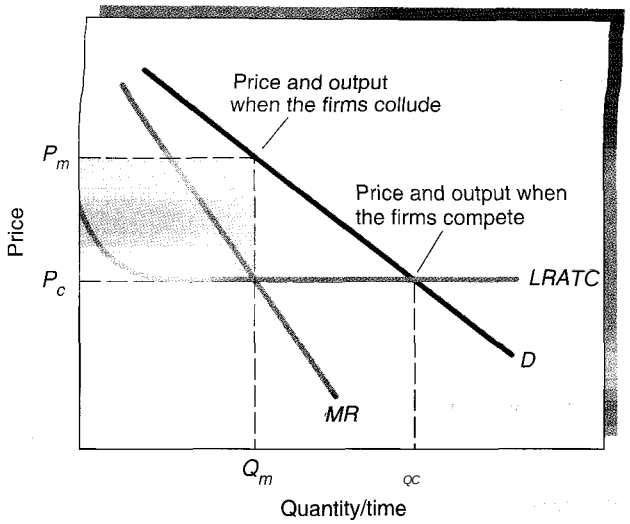


EXHIBIT 5
Price and Output in an Oligopoly

If oligopolists were to act independently and compete with one another, price cutting would drive price down to P_c. In contrast, perfect cooperation among firms would lead to a higher price of P_m and a smaller output (Q_m rather than Q_c). The shaded area shows profit if firms collude. Demand here is the market demand.

its customers would switch to rival firms, which would now expand to accommodate the new customers. The firm that raised its price would lose out. It would be self-defeating for any one firm to raise its price if the other firms didn't also raise theirs.

What happens if the current price is greater than P_c, perhaps because the market has not yet adjusted fully to a recent reduction in resource prices that lowered production costs? As a result, price is above per-unit cost and the firms are making economic profit. However, any *individual firm* that reduces its price slightly, by 1 or 2 percent, for example, will gain numerous customers if the other firms maintain the higher price. The price-cutting firm will attract some new buyers to the market, but, more important, it will also lure many buyers away from rival firms charging higher prices. Thus, if the rival sellers act independently, each will have a strong incentive to reduce price in order to increase sales and gain a large share of the total market. Price will be driven down to P_c and the economic profit will be eliminated.

However, oligopolists don't always compete with one another. They have a strong incentive to collude—to raise prices and restrict output in order to maximize their joint profits. The collusion may be either formal or informal. Formal collusion might even involve the formation of a **cartel,** like the Organization of Petroleum Exporting Countries (OPEC). Under federal antitrust laws, both formal and informal collusive agreements to restrict output and raise prices are illegal in the United States. Nonetheless, let's see what would happen if oligopolists followed this course.

Cartel
An organization of sellers designed to coordinate supply decisions so that the joint profits of the members will be maximized. A cartel will seek to create a monopoly in the market.

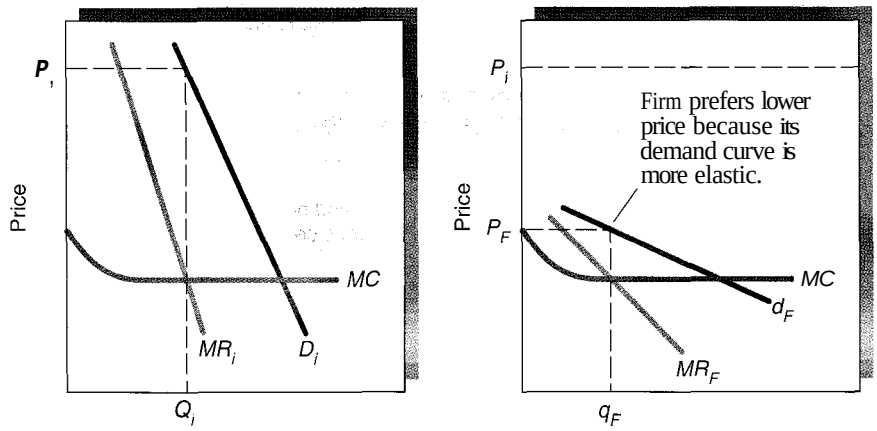
Exhibit 5 shows the marginal revenue curve that would accompany the market demand D for the product. Under perfect cooperation, the oligopolists would refuse to produce units for which marginal revenue to the group was less than marginal cost. Thus, they would restrict joint output to Q_m, where $MR = MC$. Market price would rise to P_m. With collusion, substantial joint profits (the shaded area of Exhibit 5) can be attained. If the collusion were perfect, the outcome would be identical to the one that would occur under monopoly.

Real-world outcomes are likely to be between the two polar case extremes of pnce competition and perfect collusion. Oligopolists generally recognize their interdependence and try to avoid vigorous price competition, because they know it will drive down the price to the level of per-unit costs. But there are also obstacles to collusion, which is why prices in oligopolistic industries are unlikely to rise to the monopoly level. ***As a result, oligopoly prices are typically above per-unit cost but below those a monopolist would charge.***

The Incentive to Collude and to Cheat

Collusion comes at the expense of buyers. Rival oligopolists can profit as a group by colluding to restrict output and raise price. To achieve the higher price, however, they must reduce total output. Output must be cut because consumers will purchase a smaller quan-

EXHIBIT 6
Gaining from Cheating



(a) Industry (b) Individual firm

The industry demand (D_i) and marginal revenue (MR_i) curves show that the joint profits of oligopolists would be maximized at Q_i, where $MR_i = MC$. Price P_i would be best for the industry as a whole (a). However, as part (b) shows, the demand curve (d_i) facing each firm (under the assumption that no other firms cheat) will be much more elastic than D_i. Given the greater elasticity of its demand curve, an individual firm will have an incentive to cut its price to P_F and expand output to q_F, where $MR_F = MC$. In other words, an individual oligopolist can gain by secretly cutting its price and cheating on the collusive agreement—until the other firms find out and react by cutting their price.

tity at the higher price. This raises a point of conflict: none of the oligopolists want to reduce their output. Although the firms have an incentive to collude in order to achieve higher prices, each of them also has an incentive to cheat on the agreement.

Exhibit 6 explains why this is the case. An undetected price cut will enable a firm to attract (1) customers who normally wouldn't buy from any firm at the higher price and, more important, (2) customers who would normally buy from the other firms. Therefore, the demand facing each *individual firm* will be considerably more elastic than the market demand curve. As Exhibit 6 shows, the price P_i that maximizes the industry's profits will be higher than the price P_F that is best for a single oligopolist when the others stay with the higher price. If a firm can use secret rebates or other ways to undercut the price set by the collusive agreement while the other sellers maintain the higher price, it can expand its sales (beyond the level agreed upon by the cartel). This will more than make up for the reduction in per-unit profit margin from the price cut.

In oligopolistic industries, there are two conflicting tendencies. An oligopolistic firm has a strong incentive to cooperate with its rivals so that joint profit can be maximized. But it also has a strong incentive to cheat—to secretly expand output in order to increase its profit. Oligopolistic agreements, therefore, tend to be unstable. This instability exists whether the cooperative behavior is formal, as in the case of a cartel, or informal.

Obstacles to Collusion

There are certain situations in which it is more difficult for oligopolists to collude. Five major obstacles limit collusive behavior.

1. As the number of firms in an oligopolistic market increases, the likelihood of effective collusion declines. Other things being constant, an increase in the number of major firms in an industry will make it more difficult for the oligopolists to communicate, negotiate, and enforce agreements among themselves. Furthermore, when there are more firms involved, the objectives of individual firms are more likely to conflict with those of others in the industry. Opinions about the best collusive price arrangement are likely to differ because marginal costs, unused plant capacity, and estimates of market demand elasticity will often differ among the firms. Aggressive, less mature firms might be especially interested in expanding their share of total output. As the number of firms expands, conflict

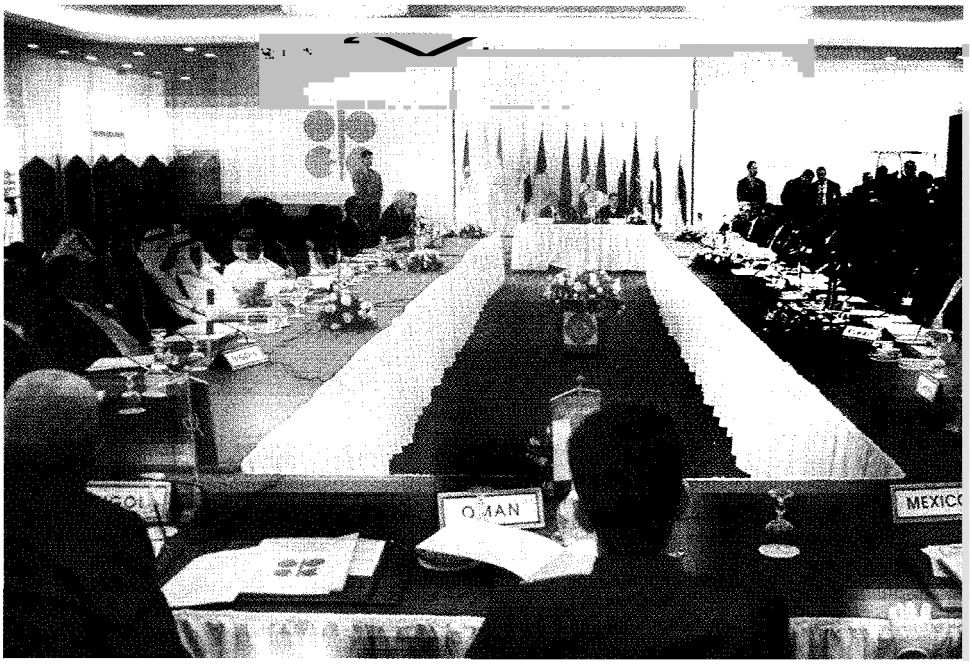
arising from these and other sources will increase and make it more difficult to maintain collusive agreements.

2. When it is difficult to detect and eliminate price cuts, collusion is less attractive. Unless a firm has a way of policing the pricing activities of rivals, it may be the "sucker" in a collusive agreement. Firms that secretly cut prices can gain a larger share of the market, while those maintaining higher prices lose customers and profits. The history of OPEC highlights this point. Cheating by members on production and sales quota agreements is not easy to monitor by the others, and it has been a persistent problem for the cartel.

Price cutting can sometimes be accomplished in ways that are difficult for the other firms to identify. For example, a firm might freely provide better credit terms, faster delivery, and other related services that improve the package offered to the buyer. In industries in which differentiated products are supplied, oligopolists can hold money prices constant and still use improvements in quality and style to provide consumers with more value. "Price cuts" like this are particularly attractive to oligopolists because they cannot be easily and quickly duplicated by rivals. As a result, collusive agreements on price are largely ineffective when quality and style are important competitive weapons. The bottom line is this: when cheating (price cutting) is profitable and difficult for rivals to detect and police, collusive agreements—both formal and informal—will be difficult to maintain.

3. Low entry barriers are an obstacle to collusion. Unless potential new rivals can be excluded, oligopolists will be unable to sustain economic profits. Profitable collusion will attract new competitors into the market. But, if entry cannot be blocked, over time the benefits of higher prices accompanying a collusive agreement will have to be shared with more and more firms. High prices will also make it more attractive for potential rivals to develop substitutes, which will reduce the demand for the product provided by the oligopolists. Of course, these factors will not work instantaneously, but over time they will gnaw away at abnormally high prices and profits resulting from collusion.

Once again, OPEC illustrates the point. As OPEC pushed prices sharply upward in the 1970s: non-OPEC producers developed new oil fields and expanded their output from existing sources. The higher prices also encouraged energy conservation and greater use of alternative energy sources, such as coal, natural gas, and nuclear energy. By the mid-1980s, OPEC's share of world oil production had fallen to about 30 percent, down from 55 percent in 1973 when it first began to restrict output and raise oil prices. At that point,



SAMIR SID/BLOOMBERG NEWS/LANDOV

Collusive agreements among oligopolists and cartels like OPEC are difficult to maintain because any individual cartel member would be better off if it could cheat on the agreement and charge a slightly lower price, so long as other members do not.

OPEC changed its course, expanding its output in order to recapture a larger share of the world market. As a result, oil prices declined throughout the 1985–2000 period. Despite recent increases, adjusted for inflation, the price of a barrel of crude oil in 2004 was still well below that of the mid-1980s. Furthermore, OPEC's share of world production stood at 40 percent in 2004, considerably less than during the early 1970s.

4. Unstable demand conditions are an obstacle to collusion. Honest differences of opinion among oligopolists about what is best for the industry always exist, but unstable demand increases the likelihood of conflict. One firm might want to expand because it anticipates a sharp increase in future demand, whereas a more pessimistic rival may want to hold the line on existing industrial capacity. The larger the differences in expectations about future demand, the greater the potential for conflict among oligopolistic firms. Successful collusion is more likely when demand is relatively stable.

5. Vigorous antitrust action increases the cost of collusion. Under existing antitrust laws in many nations, collusive behavior is prohibited. Secret agreements are, of course, possible. Simple informal cooperation might be conducted without discussions or collusive agreements. However, like other illegal behavior, such agreements are not legally enforceable by any firm. Vigorous antitrust action can discourage firms from making illegal agreements. As the threat of getting caught increases, participants will be less likely to attempt collusive behavior.

Uncertainty and Oligopoly

Uncertainty and imprecision characterize the theory of oligopoly. We know that firm will gain if they successfully restrict output and raise prices. However, collusion is fraught with conflicts and difficulties. In some industries, these difficulties are so great that the **market power** of the oligopolists is relatively small. In other industries, oligopolistic cooperation, although not perfect, can raise prices significantly, indicating a higher degree of market power, which is, in effect, a degree of monopoly power. The use of **game theory** to analyze the costs and benefits of collusive behavior, under varying degrees of cooperation and conflict, has become an important part of economics during the last two decades. (See the accompanying Applications in Economics feature on this topic.) Although this developing field of economics does not yield precise predictions on oligopoly pricing and output, it does suggest the conditions that make it more likely that an oligopolist will be kept in check by competitive pressures.

Market power
The ability of a firm that is not a pure monopolist to earn persistently large profits, indicating that it has some monopoly power. Because the firm has few (or weak) competitors, it has a degree of freedom from vigorous competition.

Game theory
A tool used to analyze the strategic choices made by competitors in a conflict situation like the decisions made by firms in an oligopoly.

MARKET POWER AND PROFIT—THE EARLY BIRD CATCHES THE WORM

Our analysis of both monopoly and oligopoly indicates that because entry barriers into these markets are often high, firms can sometimes earn economic profits over lengthy periods of time. Suppose a well-established firm, such as Disney or Microsoft, could use its market power to persistently earn above-normal returns. Would the current stockholders gain? Surprisingly, the answer is "No." The ownership value of a share of corporate stock in such a corporation would long ago have begun to reflect the firm's market power and expected future profitability. Many of the present stockholders paid high prices for their stock because they expected these firms to continue to be highly profitable. In other words, they have paid for any above-normal economic profits the firm is expected to earn because of its market power.

Do not expect to get rich buying the stock of monopolistic or oligopolistic firms known to be highly profitable. You are already too late. Those who owned the stock when these firms initially developed their market position have already captured the gain. The value of their stock increased at that time. After a firm's future prospects are widely recognized, subsequent stockholders fail to gain a higher-than-normal rate of return on their financial investment.

APPLICATIONS IN ECONOMICS

Oligopolistic Decision Making, Game Theory, and the Prisoner's Dilemma

In an oligopoly, a few firms compete in the same market, and the actions of each influence the demand faced by the others. One firm can gain customers at the expense of the others by cutting its price or increasing its advertising. The firms that fail to follow suit will lose customers. But if the other firms react competitively by doing the same, then all the firms lose. Thus, each firm finds itself on the horns of a dilemma—this is an example of the classic "prisoner's dilemma."

Economists have increasingly used game theory like the prisoner's dilemma to analyze strategic choices made by competitors, particularly those in an oligopoly. To understand the prisoner's dilemma, consider the hypothetical case of Al and Bob, two touring Americans who just met at a train station in a small foreign country. Al and Bob are taken prisoner and hauled into the local police station to be questioned separately. They are suspected of being the two men who robbed a local merchant, and each is told that if he makes the job of the police easier by confessing immediately, he will get only a six-month sentence. But each is also told that if he says nothing while the other confesses, then the one who did not confess immediately will get a twelve-month sentence. If neither confesses, both will be held for three months while the investigation continues. Al and Bob will not be allowed to communicate with each other. Will they confess?

To analyze such a situation, the game theorist begins by laying out the alternative outcomes and showing how they are related to choices made by the players of the game, as we do in **Exhibit 7.** For each man individually, in this version of the dilemma, the box reveals that the best choice depends on what the other does. If Al confesses, then Bob can save six months of jail time by also confessing. The same holds for Al, if he thinks Bob will confess. But if neither confesses, both serve only three months in jail.[1] The proper strategy for each prisoner depends heavily on his estimate of the likelihood that the other will confess. The problem becomes more complex if we consider prisoners who face a series of such decisions over time and each prisoner learns what the other has done before making his or her next choice.

[1]If the numbers were slightly different from those in the accompanying chart, both Al and Bob would have an incentive to confess regardless of what they thought the other would do. For example, if each were told that he would get only a one-month sentence if he confessed and the other party did not, then confession would become the dominant strategy for both Bob and Al Under these circumstances. if Bob thought Ai would remain dent, then Bob would spend only one month in jail if he confessed (compared to three months if he also remained silent) Al would be in a similar situation if he believed Bob would remain silent On the other hand if Bob thought Al would confess, then Bob's best option would be to confess also since confession would lead to only a six month sentence whereas silence would result in a twelve-month jail term The same would also be true for Al if he thought Bob would confess. Thus, in this classic prisoner's dilemma case, both men have an incentive to confess even though confession leads to more jail time for both—six months rather than three months. What is best for the individual (or firm) does not always lead to the best outcome from the viewpoint of the group (or industry).

EXHIBIT 7
The Prisoner's Dilemma

		Al's Choice	
		Confess	Not Confess
Bob's Choice	Confess	6 months each	Al: 12 months Bob: 6 months
	Not Confess	Al: 6 months Bob: 12 months	3 months each

Al and Bob must each decide, without communicating with each other, whether to confess. Al knows that if Bob does not confess, then he can either confess and spend six months in jail, or not confess and spend three months. But if Bob does confess, then Al's failure to confess will cost him an additional six months in jail. Bob is in a similar situation, facing the same options under the same assumptions about whether Al confesses. Each man has an incentive to confess if he thinks the other one will, but an incentive not to confess if he thinks the other will also remain silent.

Firms in an oligopoly must make decisions somewhat like those of the prisoners: Should a firm cut its price in order to attract more customers—luring some *of* them away from its competitors—or should it keep its price high and risk losing customers to competitors who cut *their* prices? If all firms keep the price high, then as a group they will reap more profit. If all of them cut prices. then as a group they will reap less profit. But the firm that fails to cut price when the others do will lose many customers The decision about whether or not to advertise has similar characteristics.

For example. consider the pricing policies or advertising strategies of large autornakers. like Ford, General Motors, and Chrysler. Suppose the profit rate of each would be 15 percent if all the major automobile producers raised their prices (or cut their advertising expenditures) by a similar amount. In contrast. each would earn a profit rate of only 10 percent if intense competition led to lower prices and/or larger advertising expenditures. Industry profits are highest when all firms decide to charge high prices. However, if one automaker cuts its price or advertises more heavily, it will be able to steal customers away from its rivals and increase its

(continued)

APPLICATIONS IN ECONOMICS

(continued)

profit rate to 20 percent. Thus, if the firm thinks its rivals will continue to charge higher prices (or fail to match its advertising expenditures), it will be tempted to do it. If the other firms follow a similar course, however, the strategy will backfire and the profit rate of all the firms in the industry will be less than the level it would have achieved had they all charged higher prices (or spent less on advertising).

Although our simplified analysis highlights the interdependence among the firms and the importance of probability estimates concerning the strategy of rivals, the real world is much more complex. The choices of the rival firms will be repeated over and over again, although often in modified forms. A prior strategy will be modified in light of previous reactions on the part of rivals. How attractive a strategy will be will depend on its likelihood of being detected by rivals and the speed with which they will respond. In addition, the strategies of oligopolistic firms will be influenced by market conditions and the threat of foreign competition—factors that change with the passage of time. As you can see, additional assumptions will need to be made to account for these and other complex factors—factors that often change in the real world. But if the assumptions incorporated into game theory models are not consistent with real-world conditions, the implications of the game theory analysis may well be invalid.

Economists have used game theory extensively to show how results change when the "rules of the game" change for the firms in an oligopolistic market, as well as for auction-bidding markets and other business decision-making situations. When the rules of the game are carefully defined and enforced, as they are in economic experiments held in laboratories, game theory has yielded interesting and important testable conclusions. But in open markets in the real world, empirical work using game theory has been less successful. The use of models is always difficult when the problem to be solved is complex and when human expectations about the changing strategic choices of other human beings must be taken into account. There is no question, however, that game theory can be useful for scholars and business practitioners alike, by helping them frame the issues involved in strategic decision making.[2]

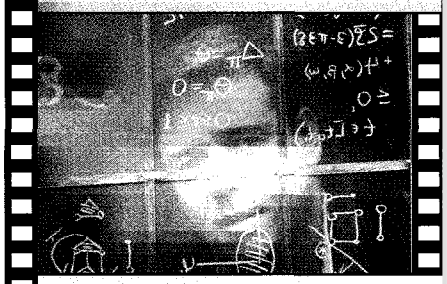
[2] For a further explanation of game theory and the prisoner's dilemma, see Avinash Dixit and Barry Nalebuff, "Game Theory," in *The Concise Encyclopedia of Economics*, ed David R Henderson (Indianapo is, Ind.: Liberty Fund, Inc., 2002), http //www econlib org/library/CEE html

DEFECTS OF MARKETS WITH HIGH ENTRY BARRIERS

From Adam Smith's time to the present, economists have generally considered monopoly to be a problem. The attitude toward markets with few sellers and high entry barriers has been only slightly more tolerant. Generally, economists level three major criticisms at concentrated markets with high entry barriers. We will outline each of them.

1. When entry barriers are high and there are few, if any, alternative suppliers, the discipline of market forces is weakened. When entry barriers are low, consumers are well equipped to direct the behavior of suppliers. Consumer choices will

ECONOMICS AT THE MOVIES



DREAMWORKS/UNIVERSAL/THE KOBAL COLLECTION/REED, E

A Beautiful Mind (2001)

A Beautiful Mind, starring Russell Crowe, tells the story of Nobel Prize–winning economist John Nash's struggle with schizophrenia. The prisoner's dilemma presented in this chapter was fundamentally shaped by Nash. His early ideas and how he came up with them are depicted in the movie, although he claims he didn't come up with the "Nash equilibrium" solution to the prisoner's dilemma in a bar, as the movie portrays.

drive firms from the market if they charge high prices, supply unattractive goods, or otherwise fail to serve the interest of consumers. If you do not like the food at a local restaurant, you eat elsewhere. If you think prices are high and the selection poor at a department store, you patronize rival firms. But consider your alternatives if you do not like the local cable television service. What can you do when the service is bad? You can voice a complaint to the company or your legislative representative, but unlike a competitive situation, you have no exit option other than sacrificing your cable service, perhaps to rely instead on satellite dish service. Cable and satellite service are imperfect substitutes, each with its own disadvantages. To the extent that consumers do not have good exit options, they must either take what the seller offers or do without. Consequently, their ability to keep sellers in check is greatly reduced.

2. Reduced competition results in allocative inefficiency. Allocative efficiency requires that additional units be produced when they are valued more highly than what it costs to produce them. When barriers to entry are low, profit-seeking producers of each good expand their output until its price is driven down to the level of per-unit costs. With high barriers to entry, however, this will not generally be the case. A monopolist or cartel that does not have to worry about rivals entering the market can often gain by restricting output and raising the price. Some units might not be produced, even though consumers value them more than their costs. Prices may exceed not only marginal costs, but also average total cost for a long period of time when entry barriers into a market are high.

3. Government grants of monopoly power will encourage rent seeking; resources will be wasted by firms attempting to secure and maintain grants of market protection. Special favors granted by the government will lead to costly activities for those who seek such favors. As we noted in Chapter 6, economists refer to such activities as rent seeking. When licenses or other entry barriers erected by the government enhance the profitability and provide protection from the rigors of market competition, people will expend scarce resources to secure and maintain these political favors. From an efficiency standpoint, these rent-seeking costs related to getting and keeping monopoly power add to the welfare losses resulting from the allocative inefficiencies we previously mentioned.

Suppose the government issues a license providing a seller with the exclusive right to sell liquor in a specific market. If this grant of monopoly power permits the licensee to earn monopoly profit, potential suppliers will expend resources trying to convince government officials that they should be granted the license. The potential monopolists will lobby government officials, make political contributions, hire representatives to do consulting studies, and undertake other actions designed to obtain the monopoly grant. Any firm that expects its rent-seeking activities to be successful will be willing to spend up to the present value of the future expected monopoly profits, if necessary. to obtain the monopoly protection. Other suppliers, of course, will also be willing to invest in rent-seeking activities. When several suppliers believe they can win, the total expenditures of all firms on rent-seeking activities can actually consume resources worth more than the economic profit expected from the monopoly enterprise.

POLICY ALTERNATIVES WHEN ENTRY BARRIERS ARE HIGH

What government policies might be used to counteract the problems that result from high barriers to entry? Economists suggest four policy options:

1. Control the structure of the industry to ensure the presence of rival firms.
2. Reduce artificial barriers that limit competition.
3. Regulate the price and output of firms in the market.
4. Supply the market with goods produced by a government firm.

Each of these policies has been used to either reduce entry barriers or counteract their negative results. We will briefly consider each of them and analyze both their potential and limitations as tools that can be used to improve the efficiency of resource use.

Antitrust Policy and Controlling the Structure of an Industry

The major problems associated with high barriers to entry and monopoly power are not present when rival sellers are present and compete. The United States, to a greater degree than most Western countries, has adopted antitrust laws designed to prevent monopoly and promote competition. Antitrust legislation began in 1890 with the passage of the Sherman Antitrust Act. Additional legislation, including the Clayton Act and the Federal Trade Commission Act of 1914, has buttressed policy in this area.

Antitrust laws give the U.S. Department of Justice the power to prosecute firms engaging in collusive behavior or other actions designed to create a monopoly or cartel. They also give the government the power to break up an existing monopoly and prevent mergers that significantly reduce competition. The Federal Trade Commission (FTC) is is allowed to bring charges under the Sherman Act. Private firms can also bring suits charging antitrust violations by a rival under the Sherman and Clayton acts and collect up to three times any actual damages caused by the rival firm. Each year, far more cases are brought by private firms against their competitors than by the two federal agencies. Rival firms can also ask the two agencies to bring antitrust cases against successful competitors who may have, or threaten to gain, monopoly power.

Economists are not in total agreement about the usefulness of antitrust action to increase the number of rival firms in an industry. Those who support policies of this type believe that only by preserving numerous competitors in a market can we be confident that monopoly has been controlled. Opponents point to the danger of protecting high-cost firms at the expense of both consumers and the more aggressive and successful firm(s) in the market. Protection of less efficient firms from successful competitors, they say, is all too common in the actual practice of antitrust policy.

Economists agree that there is one clear situation in which the application of antitrust action would be inappropriate. When substantial economies of scale are present and unit cost will be minimized only if the entire industry output is produced by a single firm, a **natural monopoly** is present. In that case, breaking up a large firm in order to increase the number of rivals in the industry would lead to higher per-unit costs. Because of their higher costs, the prices charged by the smaller firms might even exceed those of a monopolist. Therefore, using antitrust laws to expand the number of firms is not an attractive option when economies of scale in an industry produce a natural monopoly.

Natural monopoly
A market situation in which the average costs of production continually decline with increased output. In a natural monopoly, the average costs of production will be lowest when a single, large firm produces the entire output demanded by the marketplace.

Reduce Artificial Barriers to Trade

Government-imposed restraints such as licensing requirements, tariffs, and quotas often reduce the competitiveness of markets. For example, local governments often require potential entrants into the taxicab business to prove that "additional capacity is needed to serve the market." Hearings must be held to address this issue, and existing firms are allowed to explain why the new entrant is not needed. Governments also restrain entry into various occupations through the use of substantial licensing fees, complex examinations, and lengthy training requirements. When government-imposed restrictions are the source of the monopoly power, the appropriate policy action is straightforward: remove the restraints. But this is easier said than done. The restraints generally reflect the political influence of special-interest groups—the producers currently in the industry—who have a vested interest in restricting competition and keeping prices high.

If a product is traded in international markets, foreign suppliers can enhance the competitiveness of the domestic market. When the domestic market is dominated by a small number of firms, removing tariffs, quotas, and other obstacles that limit competition from abroad is particularly helpful. The competition of foreign rivals will help ensure that the domestic firms improve quality and keep their costs low.

The U.S. automobile industry illustrates the importance of this approach. Profits for U.S. producers in the highly competitive automobile segment of the industry have been low. The 2.5 percent tariff on imported autos places foreign producers at only a modest disadvantage. In contrast, there is a 25 percent tariff on imported trucks. (The tariff was implemented in 1963 to retaliate against restrictions imposed by foreign governments on U.S. frozen-chicken exports.) The higher tariff protects the same U.S. producers of cars in the truck segment of the industry. There, profits per vehicle are much higher. When General Motors began selling its newly designed line of full-size pickups in 1998 (the first full updating of those models in ten years), the profit on a well-equipped $30,000 vehicle was reported to be $8,000. The tariff keeps the price of foreign imports high, so the level of competition is lower in the U.S. market for new trucks than in the market for cars. Clearly, though, the competition from abroad has helped keep domestic auto producers on their toes. As a result, most observers believe that the overall quality of automobiles available to the American consumer, including those produced by domestic manufacturers, is higher than it would have been without the competition from abroad.

As with domestically imposed entry restraints, political considerations often make it difficult to remove obstacles that retard competition from abroad. Trade restraints that limit competition from foreign suppliers benefit primarily the owners, managers, and workers in the protected industries. These interest groups are highly concentrated and generally well organized politically. In contrast, consumers are poorly organized, and the costs imposed on them by the restraints are nearly always thinly spread among them. As we discussed in Chapter 6, the political process often works to the advantage of the concentrated interests. As a result, even when lower trade barriers are needed to increase the competitiveness of an industry, political factors often prevent them.

Regulate the Protected Producer

Can government regulation improve the allocative efficiency of a monopoly or an oligopoly? In theory, the answer to this question is clearly "Yes." **Exhibit 8** illustrates why ideal government price regulation, in the case of a monopolist, can improve resource allocation. The firm illustrated here is a natural monopoly. Its *LRATC* falls over the entire relevant range of output. Antitrust action dividing the firm into several smaller rivals will not help consumers because the unit costs of the smaller firms will be higher than if the output were produced by the monopolist. Nonetheless, from the viewpoint of allocative efficiency, the profit-maximizing monopolist will produce an output that is too small and charge a price that is too high. It will set price at P_0 and produce output Q_0, where

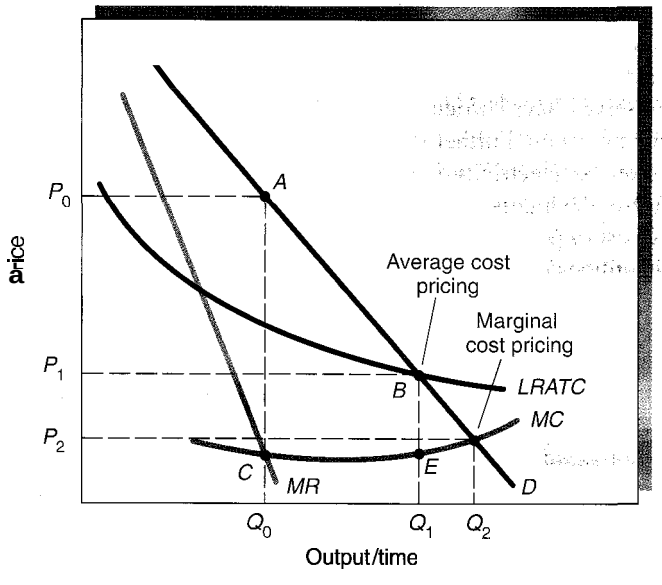


EXHIBIT 8
Regulating a Monopolist

If unregulated. a profit-maximizing monopolist with the costs indicated here will produce Q_0 units and charge P_0. If a regulatory agency forced the monopolist to reduce price to P_1, then the monopolist would expand output to Q Ideally, we would like output to be expanded to Q_2, where $P = MC$, but regulatory agencies usually do not attempt to keep prices as low as P_2. Can you explain why?

$MR = MC$. Consumers, however, will value the additional units more than the opportunity cost of producing them. Let's consider the potential of two regulatory options and also analyze some of their real-world limitations.

Average Cost Pricing

If a regulatory agency forces the firm in Exhibit 8 to reduce the price to P_1, at which point the *ATC* curve intersects the market (and firm) demand curve, then the firm will expand output to Q_1. Since it cannot charge a price above P_1, the firm cannot increase its revenues by selling a smaller output at a higher price. Once the price ceiling is instituted, the firm can increase its revenues by P_1, and by only P_1, for each unit it sells. The regulated firm's ***MR*** is constant at P_1 for all units sold until output is increased to Q_1. Since the firm's *MC* is less than P_1 (and therefore less than *MR*), the profit-maximizing regulated firm shown here will expand output from Q_0 to Q_1. The benefits from the consumption of these units (ABQ_1Q_0) clearly exceed their costs (CEQ_1Q_0). Social welfare has improved as a result of the regulatory action (we will ignore the impact on the distribution of income). At this output level, revenues are sufficient to cover costs. The firm is making zero economic profit (or "normal" accounting profit).

Marginal Cost Pricing

Even at the Q_1 output level, marginal cost is still less than price. Additional welfare gains could be achieved if output were increased to Q_2. However, if a regulatory agency forced the monopolist to reduce the price to P_2 (so the price would equal the marginal cost at the output level Q_2), economic losses would result. Even a monopolist, unless it is subsidized, wouldn't produce the product at P_2 or any price below P_1. But how large should a subsidy be? No one has the knowledge to set this subsidy, partly because prices, technologies, and the value of output change daily. Also, who has the incentive to fight the pressures on both sides for higher (or lower) subsidies? Marginal cost pricing is seldom practical.

Problems with Regulation

Although regulating a monopoly might seem like a way to improve the efficiency of markets, both economic analysis and the history of regulation suggest that it has serious shortcomings. Why? Three factors tend to reduce the effectiveness of regulations and the probability that regulators will act on behalf of all citizens to control monopolies. Let's look at each of these factors.

1. Lack of information. When we discussed ideal regulation, we assumed that we knew what the firm's *ATC*, *MC*, and demand curves looked like. In reality, of course, we don't. The firms themselves have difficulty knowing their costs, and especially their demand curves, with any precision. Furthermore, both supply and demand will vary with time and place. The market for electricity is an excellent example. Because electricity cannot be stored for later use, the highly variable supply and demand conditions in that market cause the opportunity cost of power to change dramatically from one day, or even one hour, to the next. California utilities made headlines when they experienced this early in the summer of 2001. Low rainfall reduced hydropower generation, rising natural gas prices increased gas-fired generation costs, and some hot weather early in the summer added to an already growing demand. These factors combined to sharply raise the marginal cost of delivering electric power. The regulated retail power rates were far too low to cover costs in this situation, and both policy makers and regulators were slow to adjust. Later in the summer, these demand and supply factors had shifted back toward their previous levels, and the marginal cost of power had declined substantially. Meanwhile, however, the resulting power outages and utility revenue shortfalls caused serious problems.

Because estimates of demand and marginal costs are difficult to obtain, regulatory agencies usually gauge profits (or the rate of return) to determine whether the regulated

price is too high or too low. The regulatory agency seeks to impose a "fair" or "normal" rate of return on the firm. If the firm is making profits (that is, an abnormally high rate of return), the regulated price must be higher than P, in Exhibit 8, and that price should be lowered. Alternatively, if the firm is incurring losses (less than the fair, or normal, rate of return), this indicates that the regulated price is less than P_1 and the firm should be permitted to charge a higher price.

The actual existence of profits, though, is not easily identifiable. Accounting profit, or the rate of return, is not the same as economic profit, which includes opportunity cost. Also, regulated firms have a strong incentive to adopt reporting techniques and accounting methods that conceal profits. Because of this, a regulatory agency will find it difficult or impossible to identify and impose a price consistent with allocative efficiency.

2. Cost shifting. Regulation changes incentives, so it can affect costs. If demand is sufficient, the owners of the regulated firm can expect the long-run rate of profit to be essentially fixed, regardless of whether efficient management reduces costs or inefficient management allows costs to increase. If costs decrease, the "fair return" rule imposed by the regulatory agency will force a price reduction; if costs increase, the "fair return" rule will allow a price increase. Compared to the managers of unregulated firms, managers of regulated firms have more freedom to pursue personal objectives and less incentive to seek lower costs. Predictably, they will often fly first-class, entertain lavishly at company expense, grant unwarranted wage increases, and so on. This will lead to higher unit cost. Yet the firm's monopoly position leaves consumers unable to switch to substitute products.

3. Special-interest influence. The difficulties of government regulation discussed so far are practical limitations a regulatory agency confronts. Unfortunately, the political element of government regulation exacerbates the problem. Regulated firms have a strong incentive to see that "friendly," "reasonable" people serve as regulators, and they will invest political and economic resources toward this end. Just as rent-seeking activities designed to gain monopoly privileges in the first place can be expected, so, too, can activities to influence regulators' decisions. Few consumers will invest the time and resources required to understand regulatory policy and make the political contributions needed to counteract the influence peddling. Over time, regulatory commissions are therefore likely to reflect the views and interests of the business and labor interests they were designed to regulate.

For the reasons we just explored, both economic reasoning and decades of experience show that although regulation has the potential to improve resource allocation, it is not a magic cure.

Supply Market with Government Production

Government-operated firms — socialized firms like the U.S. Postal Service, the Tennessee Valley Authority, and many local public utilities — present an alternative to both private monopoly and regulation. Here, too, however, both economic thinking and experience indicate that socialized firms will fail to counteract fully the problems that stem from high barriers to entry. The same perverse managerial incentives that regulated firms confront— incentives to ignore efficiency and pursue personal or professional objectives at the firm's expense — also tend to plague government-operated firms. In addition. the rational ignorance effect comes into play here. Individual voters have little incentive to acquire information about the operation of government firms because. unlike the case of a personal investment, the individual's choices will not decide which investment will be purchased with the voter's money. Predictably, the "owners" of a socialized firm (voters) will be uninformed about how well the firm is run, how it might be run better, or which other firm has a better plan. This is especially true when there are no direct competitors against which the firm's performance might easily be compared.

Government-operated firms do not provide an environment that rewards managers or investors for efficient management and reductions in cost. Unlike investors in the private

sector, no small group of voters is in a position to gain a substantial amount of wealth by taking over the socialized firm and improving its management. Even more than with monopoly or oligopoly in the private sector, customers of the socialized monopoly cannot easily switch their business to other sellers. Voter-taxpayers who do not even consume the product often have to pay taxes to support its provision. When the government operates a business — particularly one with monopoly power — there is typically less investor scrutiny, less reward for efficiency, and less penalty for inefficiency or the production of inferior goods. Higher costs and lower-valued outputs are predictable results.

Government ownership, like unregulated monopoly and government regulation, is a less-than-ideal solution. Those who denounce monopoly in, for instance, the local telephone industry, seldom point to a government-operated monopoly — such as first-class mail delivery by the U.S. Postal Service — as an example of how an industry should be run.

Pulling It Together

The policy alternatives available when high barriers to entry are present are less than fully satisfying. Most of the available policy alternatives are just not very attractive. Economies of scale often make antitrust actions impractical. When larger firms have lower per-unit costs, restructuring the industry to increase the number of firms will be both costly and difficult to maintain. Rivals can also use antitrust policy against successful firms that have simply given consumers a better deal and captured a large part of the market in the process. Policy actions to stop or to punish such progress are counterproductive for customers and for economic efficiency. Lifting government restrictions such as licensing requirements, quotas, and tariffs that limit competition is perhaps the surest way to improve the performance of markets with few suppliers. But the nature of the political process, particuarly its tendency to cater to the views of well-organized interest groups, often undermines the feasibility of this option. Too often in this situation, good economics does not translate into good politics.

Regulation is also a less-than-ideal solution. Regulators do not possess the information necessary to impose an efficient outcome, and as part of the political process, they will be susceptible to manipulation by industrial and labor interests. Since public-sector managers go unchecked by private-sector consumers, they are likely to pursue political objectives at the expense of economic efficiency. This means that government operation of firms also has its problems. In short, when economies of scale are present in an industry, there is no perfect way for the government to control market power.

THE COMPETITIVE PROCESS IN THE REAL WORLD

How competitive are markets today? As George Stigler points out in the second of his chapter-opening quotations, "Competition is a tough weed, not a delicate flower." In a very real sense, every firm competes with every other firm for the consumer's additional dollar of spending. Firms in markets that appear to be unrelated often compete with one another. Sellers of compact discs, for example, compete with bookstores and local restaurants for our entertainment budgets. The suppliers of swimming pools compete with airlines, hotels, casinos, and automobile rental companies for the vacation and leisure time expenditures of consumers.

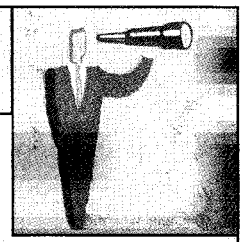
In recent decades, direct competition in almost all markets has increased. Lower transportation and communication costs have played an important role here. Lower transport costs make it more feasible for consumers to consider products offered by far-away suppliers. They also make it possible for producers to compete over a much larger geographic area, which often leads to lower per-unit costs and larger gains from economies of scale. Further, the Internet is now available to most households, and it provides even those living in small towns with quick and easy access to worldwide markets. This widening of markets reduces the market power of both buyers and sellers and, at the same time, reduces potential problems related to economies of scale.

Technological change is also a powerful force for competitive markets. Innovation and nonprice competition on product quality, design, convenience, and other factors are used

constantly by firms seeking greater market share and profitability. For a firm, innovation is the key to (temporary) profits. Successful innovation can give the innovating firm some degree of relief from competition while other firms struggle to catch up. Technical improvements, like newer and better computers or ways to make and sell them, cheaper and better inventory control methods, and MP3 players. have all made fortunes for the firms developing the information technology behind them. But the advantage is always temporary. Over time, each technical and business innovation becomes available to all firms, and the consumers of all products win, even as the profits of the innovating firms wane.

The competitiveness of markets is complex and difficult to measure directly. However, like Professor Stigler, most economists believe that competition is a tough weeed. Given the importance of quality competition, broader markets, and technological discoveries, most economists also believe that the U.S. economy is more competitive today than it was, say, fifty years ago. The pervasiveness of the competitive process helps explain why profit levels, even in manufacturing industries, are considerably lower than many people think. A national sample poll of adults conducted by Opinion Research Corporation of Princeton, New Jersey, found that the average person thought profits constituted 29 percent of every dollar of sales in manufacturing. In reality, the after-tax accounting profits of manufacturing corporations are about 4 cents to 5 cents per dollar of sales. These figures also suggest that the U.S. economy, including the manufacturing sector, is relatively competitive.

LOOKING AHEAD



The last several chapters have focused on product markets. Of course, resources are required to produce products. The following chapters will focus on resource markets and the employment decisions of both business firms and resource suppliers.

KEY POINTS

- The four major barriers to entry into a market are economies of scale (in uncontestable markets), government licensing, patents, and control of an essential resource.
- A monopoly is present when there is a single seller of a well-defined product for which there are no good substitutes and the entry barriers into the market are high. Although there are only a few markets in which the entire output is supplied by a single seller, the monopoly model also helps us better understand the operation of markets dominated by a small number of firms.
- Like other price searchers, a profit-maximizing monopolist will lower the price and expand its output as long as marginal revenue exceeds marginal cost. At the maximum-profit output, *MR* will equal *MC*. The monopolist will charge the price on its demand curve consistent with that output.
- If a monopolist is earning a profit. high barriers to entry will shield it from direct competition. thereby enabling it to earn long-run economic profits. But sometimes demand and cost conditions will be such that even a monopolist will be unable to earn economic profit.
- **An** oligopolistic market is characterized by (1) a small number of rival firms, (2) interdependence among the sellers, (3) substantial economies of scale, and (4) high entry barriers to the market.
- There is no general theory of equilibrium price and output for oligopolistic markets. If rival oligopolists acted totally independently of their competitors, they would drive the price down to the per-unit cost of production. Alternatively, if they colluded perfectly, the price would rise to the level a monopolist would charge. The actual outcome in an oligopoly will generally fall between these two extremes.

- Oligopolists have a strong incentive to collude and raise their prices. However, each firm will be able to gain if it (alone) can cut its price (or improve the quality of its product) because the demand curve confronted by the firm is more elastic than the industry demand curve. This introduces a conflict that makes it difficult to maintain collusive agreements.
- Oligopolistic firms are less likely to collude successfully against the interests of consumers if (1) the number of rival firms is large; (2) it is costly to prohibit competitors from offering secret price cuts (or quality improvements) to customers; (3) entry barriers are low; (4) market demand conditions tend to be unstable; and/or (5) the threat of antitrust action is present.
- Economists criticize high barriers to market entry because (1) the ability of consumers to discipline producers is weakened, (2) the unregulated monopolist or oligopoly group can often gain by restricting output and raising price, and (3) legal barriers to entry encourage firms to engage in wasteful rent-seeking activities.
- When entry barriers are high and competition among rival firms weak, the major policy alternatives are: (1) antitrust action designed to maintain or increase the number of firms in the industry, (2) relaxation of regulations that limit entry and trade, (3) price regulation, and (4) provision of output by government firms. When feasible, option (2) is the most attractive alternative. Under most circumstances, all the other options have shortcomings.
- Competitive forces are present and growing, even in markets with high barriers to entry. Quality competition, the widening of markets, and technological change are all important elements of the competitive process.

CRITICAL ANALYSIS QUESTIONS

*1. "Barriers to entry are crucial to the existence of long-run profits, but they cannot guarantee the existence of profits." Evaluate.

2. "Monopoly is good for producers but bad for consumers. The gains of the former offset the losses of the latter. On balance, there is no reason to think that monopoly is bad for the economy." Is this statement true? Why or why not?

*3. Do monopolists charge the highest prices for which they can sell their products? Do they maximize their average profit per sale? Are monopolistic firms always profitable? Why or why not?

4. The retail liquor industry is potentially a competitive industry. However, the liquor retailers of a southern state, with the cooperation of the state legislature, organized a trade association that sets prices for all firms. For all practical purposes, a competitive industry became a monopoly. Compare the price and output policy of a purely competitive industry with the policy that would be established by a profit-maximizing monopolist or trade association. Who benefits and who is hurt by the formation of the monopoly?

5. Does economic theory indicate that an ideal regulatory agency that forces a monopolist to charge a price equal to either marginal or average total cost will improve economic efficiency? Explain. Does economic theory suggest that a regulatory agency will in fact follow a proper regulation policy? What are some of the factors that complicate the regulatory function?

6. Is a monopolist subject to any competitive pressures? Explain. Would an unregulated monopolist have an incentive to operate and produce efficiently? Why or why not?

7. How will high entry barriers into a market influence (a) the long-run profitability of the firms, (b) the cost efficiency of the firms in the industry, (c) the likelihood that some inefficient (high-cost) firms will survive, and (d) the incentive of entrepreneurs to develop substitutes for the product supplied by the firms? Are competitive pressures present in markets with high barriers to entry? Discuss.

*8. Why is oligopolistic collusion more difficult when there is product variation than when the products of all firms are identical?

9. In large cities, taxi fares are often set above the market equilibrium rate. Sometimes the number of licenses is limited in order to maintain the above-market price. Other times licenses are automatically granted to anyone wanting to operate a taxi. When taxi fares are set above market equilibrium, compare and contrast resource allocation under the restricted license system (assume the licenses are tradable) and the free-entry system. In which case will it be easier for customers to get a taxi? In which case will

the amount of capital required to enter the taxi business be greater?

10. We have a theory to explain the equilibrium price and output for monopoly, but not for oligopoly. Why? How can game theory help us understand the decisions made by oligopolists?

***11.** Historically, the real cost of transporting both goods and people has declined substantially. What impact do lower transportation costs have on the market power of individual producers? Do you think the U.S. economy is more or less competitive today than it was 100 years ago? Explain.

***12.** "My uncle just bought 1,000 shares of Microsoft, one of the largest and most profitable companies in the United States. Given Microsoft's high profit rate, he will make a bundle from this purchase." Evaluate this statement. Is it necessarily true? Explain.

***13.** Gouge-em Cable Company is the only cable television service company licensed to operate in Backwater County. Most of its costs are access fees and maintenance expenses. These fixed costs total $640,000 monthly. The marginal cost of adding another subscriber to its system is constant at $2 per month. Gouge-em's demand curve can be determined from the data in the accompanying table.

Subscription Price	Number of Subscribers (per month)
$25	20,000
20	40,000
15	60,000
10	80,000
5	100,000
1	150,000

a. What price will Gouge-em charge for its cable services? What are its profits at this price?

b. Now suppose the Backwater County Public Utility Commission has the data and believes that cable subscription rates in the county are too expensive and that Gouge-em's profits are unfairly high. What regulated price will it set so that Gouge-em makes only a normal rate of return on its investment?

14. Suppose that you produce and sell children's tables in a local market. Past experience enables you to estimate your demand and marginal cost schedules. This information is presented in the table at the bottom of this page.

a. Fill in the missing revenue and cost schedules.

b. Assuming you are currently charging $55 per table set, what should you do if you want to maximize profits?

c. Given your demand and cost estimates, what price should you charge if you want to maximize your weekly profit? What output should you produce? What is your maximum weekly profit?

15. The diagram shows demand and long-run cost conditions in an industry.

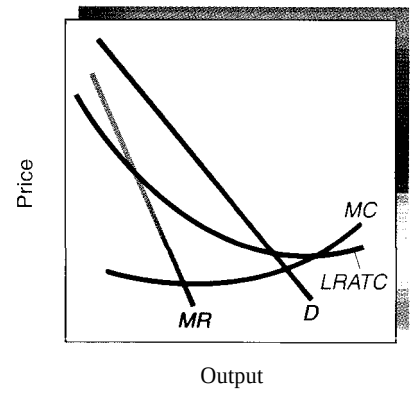
a. Explain why the industry is likely to be monopolized.

b. Indicate the price that a profit-maximizing monopolist would charge, and label it *P*.

c. Indicate the monopolist's output level, and label it *Q*.

d. Indicate the maximum profits of the monopolist.

e. Will the profits attract competitors to the industry? Why or why not? Explain.



Price	Quantity Demanded (per week)	Marginal Cost	Total Revenue	Marginal Revenue	Fixed Cost	Total Cost
$60	1	$50	____	____	$40	____
55	2	20	____	____	____	____
50	3	24	____	____	____	____
45	4	29	____	____	____	____
40	5	35	____	____	____	____
35	6	45	____	____	____	____

*Asterisk denotes questions for which answers are given in Appendix B.

CHAPTER 12

The Supply of and Demand for Productive Resources

It is . . . necessary to attach price tags to the various factors of production . . . in order to guide those who have the day-to-day decisions to make as to what is plentiful and what is scarce.

—*Professor James Meade*[1]

Chapter Focus

- **Why do business firms demand labor, machines, and other resources? Why is the demand for a productive resource inversely related to its price?**
- **How do business firms decide which resources to employ and the quantity of each that will be used?**
- **How is the quantity supplied of a resource related to its price in the short run? In the long run?**
- **What determines the market price of a resource? How do resource prices help efficiently allocate a society's resources across competing uses?**

[1]James E. Meade was a longtime professor of economics at Cambridge University.

Recent chapters have focused on product markets, markets in which consumers purchase goods and services supplied by business firms. Our analysis now shifts to **resource markets**, markets in which firms hire productive resources like machines and workers and use them to produce goods and services. *(Note:* Because resources are also referred to as *factors* or *inputs,* these markets are also known as *factor markets* or *input markets.)*

Just as in product markets, the forces of supply and demand combine to determine prices in resource markets. The buyers and sellers in resource markets are just the reverse of what they are in product markets. In resource markets, business firms are the purchasers; they demand resources that are used to produce goods and services. Households are the sellers; they (and firms they own) supply resources in exchange for income. The income from supplying productive resources, such as the wages received from the sale of labor services, is the major source of income for most of us. Prices in resource markets coordinate the choices of buyers and sellers and bring the amount of each resource demanded into harmony with the amount supplied. Resource prices also help allocate factors of production efficiently and channel them into the areas where they are most productive. This enables us to have higher incomes and a larger supply of consumer goods than would otherwise be the case.

As the circular flow diagram of **Exhibit 1** illustrates, there is a close relationship between product and resource markets. Households earn income by selling factors of production — for example, the services of their labor and capital — to business firms. Their offers to sell form the supply curve in resource markets (bottom loop). The income that households get from the sale of resources gives them the buying power they need to purchase goods and services in product markets. These expenditures by households generate revenue and motivate firms to produce goods and services (top loop). In turn, firms demand resources because they contribute to the production of goods and services that can be sold in product markets.

HUMAN AND NONHUMAN RESOURCES

Broadly speaking, there are two different types of productive inputs, nonhuman and human. **Nonhuman resources** can be broken down into physical capital, land, and natural resources. *Physical capital* consists of human-made resources, such as tools, machines, and buildings, that are uced to produce other things.

Net investment can increase the supply of nonhuman resources. However, investment involves a cost. Resources used to produce machines, upgrade the quality of land, or discover natural resources could be used to produce goods and services for current consumption instead of invested for the future. Why take the roundabout path? The answer is that sometimes indirect methods of producing goods are less costly in the long run. For example, Robinson Crusoe found he could catch more fish by taking some time off from hand-fishing to build a net. Even though his initial investment of time to make the net reduced his current catch, once the net was completed he was able to more than make up for this loss. Trade-offs like these influence what people will invest in. be it fishing nets or complex machines. An investment will be undertaken only when the decision maker expects the benefits of a larger future output to more than offset the current reduction in the production of consumption goods. Just as the supply of machines can be increased with investment, so, too, can investments in better land development and soil-conservation practices improve the quantity and quality of usable land. Similarly, the supply of natural resources like oil and gas, for example, can be increased (to some extent) by making an investment in, or dedicating more resources to, exploration and development.

Human resources consist of the skills and knowledge of workers. Investments in education, training, health, and experience can enhance the skills, abilities, and ingenuity of individuals, and thereby increase their productivity. Economists call activities like these

Resource markets
Markets in which business firms demand factors of production (for example, labor, capital, and natural resources) from household suppliers. The resources are then used to produce goods and services. These markets are sometimes called factor markets or input markets.

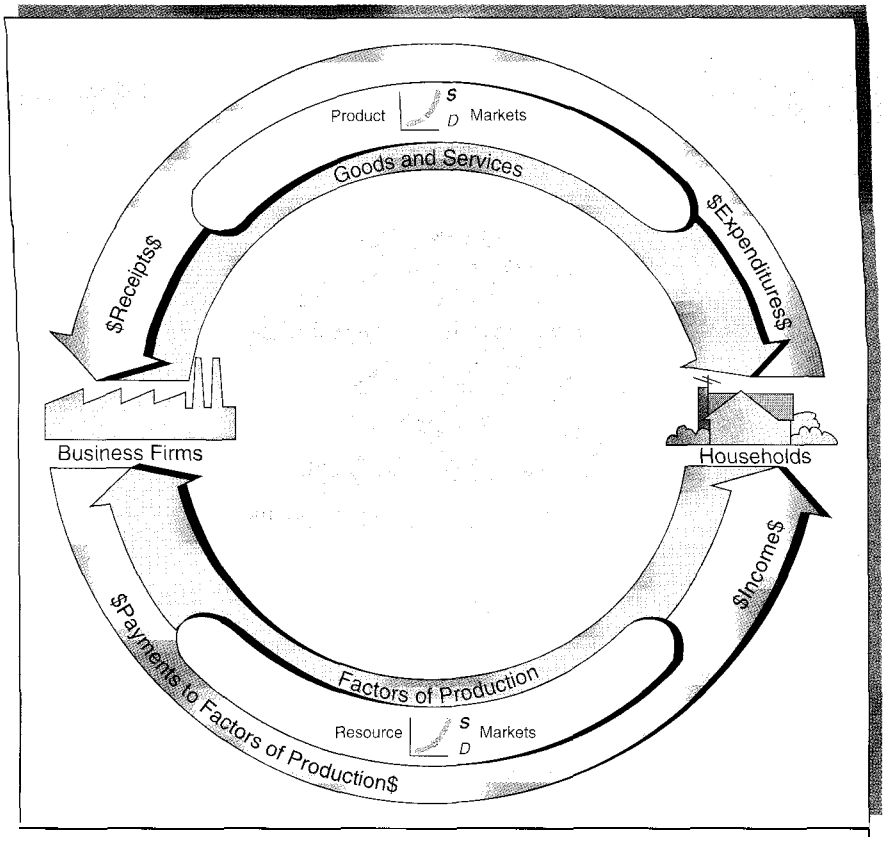
Nonhuman resources
The durable. nonhuman inputs used to produce both current and future output. Machines, buildings, land, and raw materials are examples. Investment can increase the supply of nonhuman resources. Economists often use the term physical capital when referring to nonhuman resources.

Human resources
The abilities, skills, and health of human beings that contribute to the production of both current and future output. Investment in training and education can increase the supply of human resources.

EXHIBIT 1

The Market for Resources

Until now, we have focused on product markets, in which households demand goods and services that are supplied by firms (upper loop) We now turn to resource markets, in which firms demand factors of production—human capital (like the skills and knowledge of workers) and physical capital (like machines, buildings, and land). Factors of production are supplied by households in exchange for income (bottom loop). In resource markets, firms are buyers and households are sellers—just the reverse of their roles in product markets.



Investment in human capital Expenditures on training, education, skill development, and health designed to increase human capital and people's productivity.

an **investment** in **human capital**.[2] Like physical capital, human capital also depreciates—people's skills, for example, can decline with age or lack of use. Education and training will add to the stock of human capital while depreciation detracts from it.

Decisions to invest in human capital are no different from other investment decisions we make. Consider your decision about going to college. As you know, an investment in a college education requires you to sacrifice some current earnings as well as pay for direct expenses, like tuition and books. However, you are making the investment anyway because you expect it to lead to a better job and other benefits later. A rational person will attend college only if the expected future benefits outweigh the current costs.

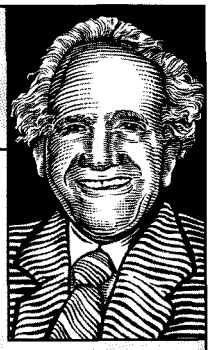
Human resources differ from nonhuman resources in two important respects. First, human capital is embodied in the individual. Individuals cannot be separated from their knowledge, skills, and health conditions in the same way that they can be separated from physical capital, like buildings or machines that they might own. As a result, in addition to money, a person's job choices are also affected by a job's working conditions, location, prestige, and other nonmonetary factors. Money, of course, influences people's human capital decisions. However, people will often choose to trade off some money income for better working conditions. Second, human resources cannot be bought and sold in nonslave societies. Workers sell only the *services* of their labor. They have the option of quitting, selling their labor services to another employer, or using them in some other way. Thus, we usually speak of the worker as selling (and the firm as buying) labor services.

In competitive markets, the price of resources, like the price of products, is determined by supply and demand. We will begin our analysis of resource markets by focusing on the demand for resources. both human and nonhuman.

[2]The contributions of T. W. Schultz and Gary Becker to the literature on human capital have been particularly significant. See Ronald G. Ehrenberg and Robert S. Smith, ***Modern Labor Economics: Theory and Public Policy***, 8th ed. (Reading, MA: Addison-Wesley, 2003), chap. 9, for additional detail on human capital theory.

OUTSTANDING ECONOMIST

Gary Becker (1930–)



This 1992 Nobel Prize recipient is best known for his role in the development of human capital theory and his innovative application of that theory to areas as diverse as employment discrimination, family development, and crime. In his widely acclaimed book *Human Capital*,* Becker developed the theoretical foundation for human investment decisions in education, on-the-job training, migration, and health. Becker is a past president of the American Economic Association and a longtime professor at the University of Chicago.

*Gary Becker, *Human Capital* (New York Columbia University Press, 1964).

THE DEMAND FOR RESOURCES

Profit-seeking producers employ laborers, machines, raw materials, and other resources because they help produce goods and services. ***The demand for a resource exists because there is a demand for goods that the resource helps to produce. The demand for each resource is thus a derived demand; it is derived from the demand of consumers for products.***

Derived demand
The demand for a resource; it stems from the demand for the final good the resource helps produce.

For example, an auto repair shop hires mechanics because customers demand repair service, not because the auto repair shop owner benefits simply from having mechanics around. If customers did not demand repair service, mechanics would not be employed for long. Similarly, the demand for inputs like carpenters, plumbers, lumber, and glass windows is derived from the demand of consumers for houses and other consumer products that these resources help to make. Most resources contribute to the production of numerous goods. For example, glass is used to produce windows, ornaments, dishes, light bulbs, and mirrors, among other things. The total demand for a resource is the sum at each resource price of the derived demands for each of its uses.

The demand curve for a resource shows the amount of the resource that will be used at different prices. As **Exhibit 2** shows, there is an inverse relationship between the price of a resource and the amount demanded of it. There are two major reasons that less of a resource will be demanded as its price increases: (1) producers will turn to substitute resources and (2) consumers will buy fewer goods that become more expensive as the result of higher resource costs. Let us take a closer look at each of these factors.

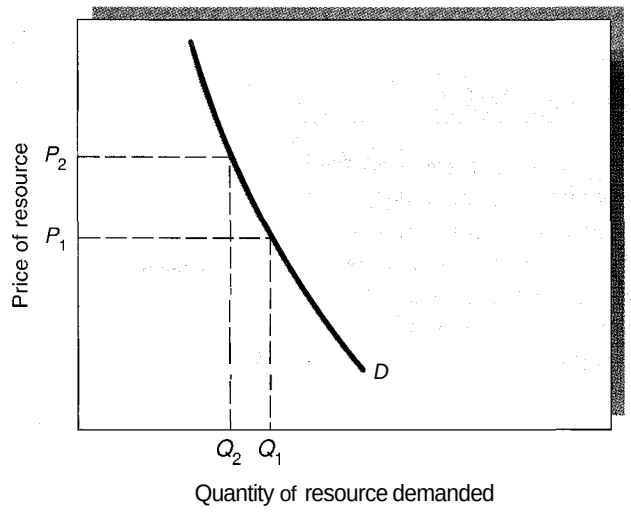
Substitution in Production

Firms will use the input combination that minimizes their costs. When the price of a resource goes up, firms will use lower-cost substitute inputs and cut back on their use of the more expensive resource.

Typically, there are many ways producers can reduce their use of a more expensive resource. For example, if the price of oak lumber increases, furniture manufacturers will use other wood varieties, metals, and plastics more intensely. Similarly, if the price of copper tubing increases, construction firms and plumbers will substitute plastic pipe for it. Sometimes producers will alter the style and dimensions of a product in order to use less of a more expensive resource. Relocation is also a substitution strategy. For example, if the price of office space and land increases in the downtown area of a large city, firms may move to the suburbs. The degree to which firms will be able to cut back on a more

EXHIBIT 2
The Demand Curve for a Resource

As the price of a resource increases, producers that use the resource intensely will (1) turn to substitute resources and (2) face higher costs. which will lead to higher product prices and lower output. At the lower rate of output, producers will use less of the resource that increased in price. Both of these factors contribute to the inverse relationship between the price and amount demanded of a resource.



expensive resource will vary. The easier it is to turn to substitutes, the more elastic the demand for a resource is. ***Other things being constant, the demand for a resource will be more elastic when more and better substitutes are available for it.***

Substitution in Consumption

An increase in the price of a resource will lead to higher costs of production and thus higher prices for the products that the input helps to produce. Faced with these higher prices, *consumers* will turn to substitute products and cut back on their purchases of the more expensive products. In turn, a smaller quantity of resources (including less of the one that rose in price) will be required to produce the smaller amount of the product demanded by consumers at the now higher price.

To illustrate the substitution-in-consumption effect, suppose the United Auto Workers negotiates a substantial wage increase for employees in the U.S. automobile industry. The large wage increase will push the costs of American automakers upward. As the price of domestically produced autos rises, many consumers will switch to substitutes, particularly automobiles produced abroad. The sales of American producers will fall, lowering the quantity of labor demanded (and employment) in the U.S. automobile industry.

Other things being constant, the more elastic the demand for a product, the more elastic the demand for the resources used to make it. This relationship stems from the derived nature of resource demand. An increase in the price of a product for which consumer demand is highly elastic will cause a sharp fall in the sales of the good. As a result, there will be a relatively sharp fall in the demand for the resources used to produce it.

In summary, the demand elasticity of a resource will vary with the ease of substitution when it comes to both production and consumption. The demand for a resource will tend to be elastic when it is easy to substitute other resources for it in production and when the demand for goods produced with it is relatively elastic. Conversely, the demand for a resource will tend to be inelastic when it is difficult to find good substitutes for it in production and the demand for the goods produced with it is more inelastic.

How Time Changes the Demand for Resources

The elasticity of resource demand is also influenced by time. It takes time for producers to fully adjust to a change in the price of a resource. Typically, a producer will be unable to

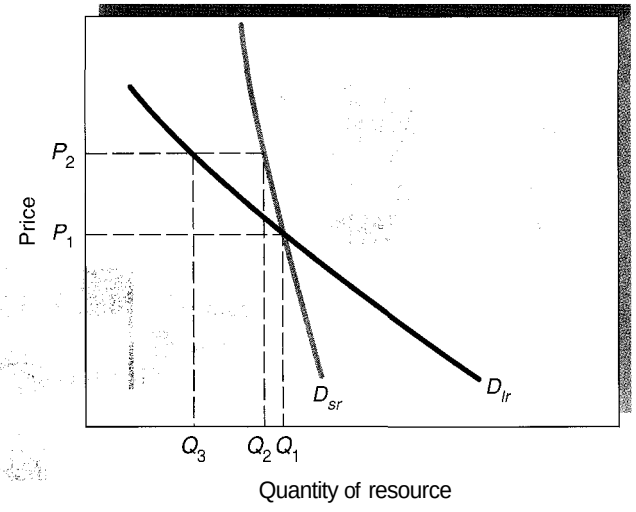


EXHIBIT 3
Time and the Demand Elasticity of Resources

The demand for a resource will be more elastic (1) the easier it is for firms to switch to substitute inputs and (2) the more elastic the consumer demand for the products the resource helps produce. As the graph here shows, demand for a resource in the long run (D_{lr}) is nearly always more elastic than demand in the short run (D_{sr}).

immediately alter a production process or redesign a product to use less of a more expensive input or more of an input that has declined in price. Consumers may also find it difficult to alter their consumption patterns quickly in response to price changes. For example, if the price of cigarettes rises due to a tax increase, cigarette smokers may find it initially hard to reduce their consumption of cigarettes very much. Over time, however, the higher price will cause more and more smokers to smoke less. Thus, the demand for a resource generally becomes more elastic with the passage of time.

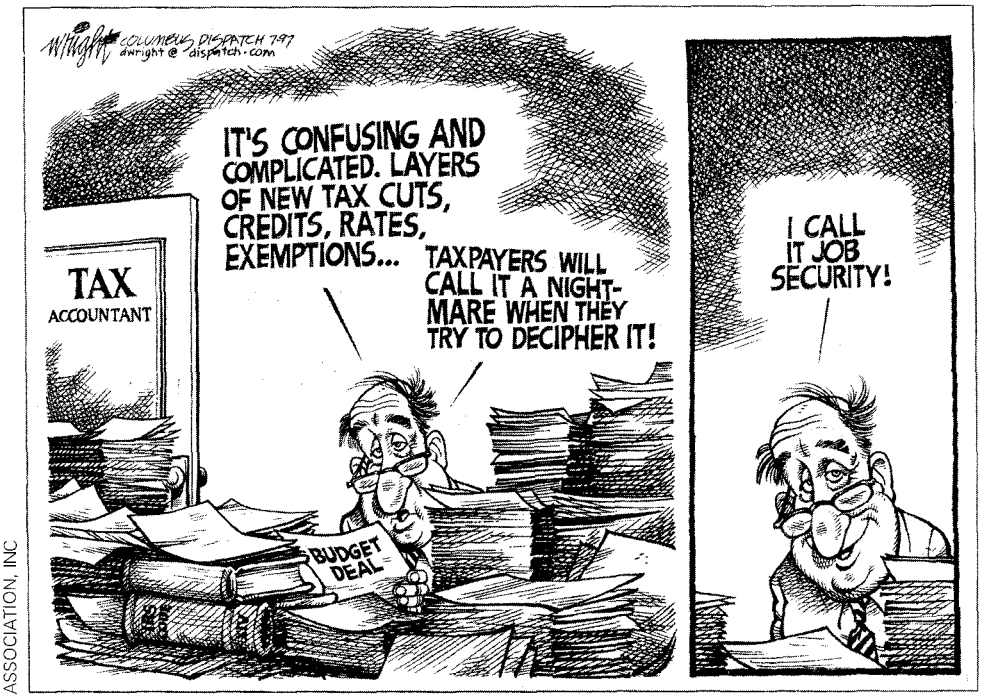
Exhibit 3 shows how time affects the elasticity of resource demand. Since it is generally difficult to substitute quickly away from a more expensive resource, demand is relatively inelastic in the short run. Notice how steep, or inelastic, the slope of the short-run demand curve (D_{sr}) is. An increase in price from P_1 to P_2 will lead to only a small fall in the quantity of the resource used (from Q, to Q_2). Given more time, however, producers will be able to make a larger substitution away from the more expensive resource. The increase in price to P_2 causes a much larger fall in the quantity demanded (to Q,) over time. The slope of the long-run demand curve (D_{lr}) is not so steep. as you can see. but is more elastic. In the long run, the demand for a resource is nearly always more elastic than in the rhort run.

Things That Change the Demand for Resources

Like the demand schedule for a product, the entire demand curve for a **resource may shift.** There are three major reasons why:

1. ***A change in the demand for a product will cause a similar change in the demand for the resources used to make the product.*** An increase in the demand for a consumer good simultaneously increases the demand for resources needed to make it Conversely, a fall in the demand for a product will lower the demand for the resources used to make it. Recent changes in the tax-preparation industry Illustrate this point. Starting in the mid-1990s, the demand for tax-preparation software increased sharply, driven by the introduction of easy-to-use software and the growing use of personal computers. This led to an increase in the demand for programmers to produce the tax-preparation software, and their employment increased rapidly as a result. In contrast, the higher consumer demand for tax-preparation software meant falling demand for tax accountants and other tax preparers. This reallocation of resources is a natural and integral part of how markets respond to changes in product demands.

The demand for resources is a derived demand. A more complex tax code would increase the demand for (and thus the wages of) accountants. whereas a simpler tax code would have the opposite effect.



DICK WRIGHT REPRINTED BY PERMISSION OF NEWSPAPER ENTERPRISE ASSOCIATION, INC

2. ***Changes in the productivity of a resource will alter demand—the higher the productivity of a resource, the greater will be the demand for it.*** As the productivity of a resource increases, so does its value to potential users. Improvements in the quality of a resource—in the case of workers, their skill levels—will increase the productivity of the resource and therefore the demand for it. For example, as workers gain valuable new knowledge and/or upgrade their skills, they enhance their productivity and essentially move into a different skill category—one where demand is greater.

The productivity of a resource will also depend on the amount of other resources used in the production process. In general, additional capital will tend to increase the productivity of labor. For example, someone with a dump truck can haul more material than the same person with a wheelbarrow. The quantity and quality of the tools with which you work will significantly affect your productivity.

Improvements in technology also tend to increase the productivity of resources, including labor. For example, technological advances in word processing equipment have enhanced the productivity of secretaries, journalists, lawyers, and writers. Similarly, computers have substantially increased the productivity of typesetters, telephone operators, scientific researchers, and workers in many other occupations. The link between technological advances and worker productivity helps explain why improvements in technology generally do not exert a large negative impact on employment, even in the occupations most directly affected. Of course, when firms substitute new technology for labor services, the demand for labor will fall. However, the new technology also makes the labor more productive, which, in turn, increases the demand for labor services. This second effect will partially, and sometimes more than completely, offset the first effect.

The productivity-demand linkage sheds light on why wage rates in the United States, Canada, Western Europe, and Japan are higher than in most other areas of the world. Given the skill level of workers, the technology, and the capital equipment with which they work, individuals in these countries produce more goods and services per hour of labor than workers in most other countries. In turn, the demand for their labor (relative to supply) is greater because of their high productivity. Essentially, the workers' greater productivity leads to their higher wage rates.

3. ***A change in the price of a related resource will affect the demand for the original resource.*** **A** rise in the price of a resource will cause the demand for substitute resources to expand. For example, when the price of lumber increases, the demand for bricks will increase as home builders switch to building more brick homes and fewer wood homes. Conversely, an increase in the price of a resource that is a *complement* to a given resource will decrease the demand for the given resource. For example, higher lumber prices will tend to lower the demand for nails.

MARGINAL PRODUCTIVITY AND THE FIRM'S HIRING DECISION

How does a producer decide whether to employ additional units of a resource? As for other decisions, the marginal benefit relative to the marginal cost provides the answer. Because firms are mostly price takers in resource markets (meaning they can hire as many units of the resource as they wish without affecting the market price of the resource), the marginal cost of hiring one more worker is simply the worker's wage, whereas the marginal cost of purchasing a machine is its price. These represent the increase in the firm's costs from employing one more unit of the resource. But what about the marginal benefit of the resource to the firm? It is measured by the increase in the firm's revenue from employing one more unit of the resource. This is called the resource's **marginal revenue product** (*MRP*). **A** profit-maximizing firm will hire an additional unit of the resource only if the marginal revenue product exceeds the cost of employing the resource.

Marginal revenue product (*MRP*)
The change in the total revenue of a firm that results from the employment of one additional unit of a resource. The marginal revenue product of an input is equal to its marginal product multiplied by the marginal revenue of the good or service produced.

Suppose that a retail store is considering hiring a security guard for $25 per hour to help reduce shoplifting. If the security guard could prevent $20 worth of shoplifting per hour, should the profit-maximizing firm hire the guard? Because the marginal cost of employing the security guard (the wage of $25) is higher than the guard's marginal revenue product (the $20 reduction in shoplifting per hour). the wise decision is for the firm not to hire the security guard. Hiring the guard will lower the firm's profit by $5 per hour. The guard should be employed only if the reduction in shoplifting exceeds the guard's wage cost. In most situations, the direct impact of hiring an additional resource on a firm's revenue is not as clear, so let's take a closer look at the firm's decision and how marginal revenue product is determined.

Using a Variable Resource with a Fixed Resource

When an additional unit of the resource is used relative to a fixed amount of other resources, the firm's output will increase by an amount equal to the resource's **marginal product** (*MP*). Because marginal product is measured in units of physical output. it **is** sometimes referred to as marginal physical product. How much additional revenue can the firm derive from the employment of the resource? Recall that **marginal revenue** (*MR*) **is** the increase in the firm's revenue that results from the sale of each additional unit of output. Thus, a resource's marginal revenue product is equal to the marginal product of the resource multiplied by the marginal revenue of the good or service produced. ***When the law of diminishing returns sets in, the marginal product of a resource will fall as employment of the resource expands. As a result, the marginal revenue product of a resource will also fall as employment expands.***

Marginal product (*MP*)
The change in total output that results from the employment of one additional unit of a resource

Marginal revenue (*MR*)
The change in a firm's total revenue that results from the production and sale of one additional unit of output.

The relationship between the marginal revenue a firm gets from selling an additional unit of output and the price for which it is sold is different for *price-taker* firms than for *price searchers*, however. Because a price-taker firm sells all units produced at the same price, the price taker's marginal revenue will be equal to the market price of the product. The price searcher, however, must reduce the price of all units in order to expand the number of units sold. Consequently, the price searcher's marginal revenue will be less than the sales price of the units. The marginal product of a resource multiplied by the selling price of the product is called the resource's **value marginal product** (*VMP*). ***For a price-taker firm, the* MRP *of a resource is equal to its* VMP *because price and marginal revenue***

Value of marginal product (*VMP*)
The marginal product of a resource multiplied by the selling price of the product it helps produce. For a price-taker firm, marginal revenue product (*MRP*) will be equal to the value marginal product (*VMP*).

EXHIBIT 4
The Short-Run Demand Schedule of a Firm

Compute-Accounting Inc. uses computer technology and data entry operators to provide accounting services in a competitive market. For each accounting statement processed, the firm receives a $200 fee (column 4). Given the firm's current fixed capital, column 2 shows how total output changes as additional data entry operators are hired. The marginal revenue product (*MRP*) schedule (column 6) indicates how hiring an additional operator affects the total revenue of the firm. Since a profit-maximizing firm will hire an additional employee if, and only if, the employee adds more to revenues than to costs, the marginal revenue product curve is the firm's short-run demand curve for the resource (see Exhibit 5).

Units of Variable Factor (Data Entry Operators) (1)	Total Output (Accounting Statements Processed per Week) (2)	Marginal Product (Change in Column 2 Divided by Change in Column 1) (3)	Sales Price per Statement (4)	Total Revenue (2) × (4) (5)	MRP (3) × (4) (6)
0	0.0	—	$200	$ 0	—
1	5.0	5.0	200	1,000	1,000
2	9.0	4.0	200	1,800	800
3	12.0	3.0	200	2,400	600
4	14.0	2.0	200	2,800	400
5	15.5	1.5	200	3,100	300
6	16.5	1.0	200	3,300	200
7	17.0	0.5	200	3,400	100

are equal. For a price-searcher firm, however, the* MRP *of a resource will be lower than its* VMP *because marginal revenue is less than price.

Using these measures, **Exhibit 4** illustrates how a firm decides how much of a resource to employ. Compute-Accounting Inc. uses computer equipment and data entry operators to supply clients with monthly accounting statements. The firm is a price taker: it sells its service in a competitive market for $200 per statement. Given the fixed quantity of computer equipment owned by Compute-Accounting, column 2 shows how much total output (quantity of accounting statements) the firm can produce with different numbers of data entry operators. One data entry operator can process five statements per week. When two operators are employed, nine statements can be completed. Column 2 indicates how total output is expected to change as additional data entry operators are employed. Column 3 presents the marginal product schedule for data entry operators. Column 6, the *MRP* schedule, shows how the employment of each additional operator affects total revenue. Due to the law of diminishing returns, both the marginal product and the marginal revenue product of workers decline as additional operators are employed.

Because Compute-Accounting is a price taker, the marginal revenue product and the value marginal product of labor are equal. Thus, the marginal revenue product of labor (column 6) can be calculated by multiplying the marginal product (column 3) times the sales price of an accounting statement (column 4).

How does Compute-Accounting decide how many operators to employ? *Answer:* It analyzes the benefits relative to the costs. As additional operators are employed, the output of processed statements (column 2) will increase, which will expand total revenue (column 5). Employing additional operators, though, will also add to production costs because the operators must be paid. Applying the profit-maximization rule, Compute-Accounting will hire additional operators as long as their employment adds more to revenues than to costs. This will be the case as long as the *MRP* (column 6) of the data entry operators exceeds their wage rate. At a weekly wage of $1,000, Compute-Accounting would hire only one operator. If the weekly wage dropped to $800, two operators would be hired. At still lower wage rates, additional operators would be hired.

Profit-maximizing firms will expand their employment of each variable resource until the* MRP *of the resource (the firm's additional revenue generated by the resource) is

just equal to the price of the resource (the firm's marginal cost of employing the resource). This profit-maximization rule applies to all firms, price takers and price searchers alike.

MRP and the Firm's Demand Curve for a Resource

Using the data in Exhibit 4, we can construct Compute-Accounting's demand curve for data entry operators. Recall that the height of a demand curve shows the maximum price (in this case, the wage) that the buyer (the firm) would be willing to pay for the unit. Because the marginal revenue product of the first data entry operator is $1,000, the firm would be willing to hire this worker only up to a maximum price of $1,000. Because of this relationship, as **Exhibit 5** shows, a firm's short-run demand curve for a resource is precisely the *MRP* curve for the resource.[3] Using this demand curve yields the identical solutions as the table. At a weekly wage of $1,000, Compute-Accounting would hire only one operator. If the weekly wage dropped to $800, two operators would be hired. At still lower wage rates, additional operators would be hired. Underlying the downward-sloping demand curve is the law of diminishing returns, which causes *MP*, and thus *MRP*, to fall as more workers are hired, when at least one factor of production is fixed.

The location of the firm's *MRP* curve depends on (1) the price of the product, (2) the productivity of the resource, and (3) the amount of other resources with which the resource is working. Changes in any one of these three factors will cause the *MRP* curve to shift. For example, if Compute-Accounting purchased additional computer equipment, making it possible for the operators to complete more statements each week, the *MRP* curve for labor would increase (shift outward). This increase in the quantity of the other resources working with labor would increase labor's productivity.

Multiple Resources and How Much to Use of Each

So far, we have analyzed the firm's hiring decision assuming that it used one variable resource (labor) and one fixed resource. Production, though, usually involves the use of many resources. How should these resources be combined to produce the product? We can answer this question by considering the conditions for either profit maximization or cost minimization.

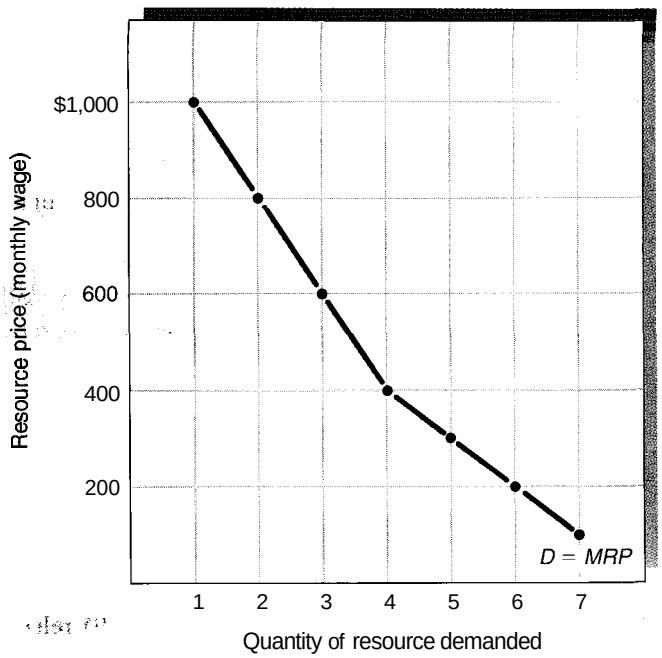


EXHIBIT 5
The Firm's Demand Curve for a Resource

The firm's demand curve for a resource will reflect the marginal revenue product (*MRP*) of the resource. In the short run, it will slope downward because the marginal product of the resource will fall as more of it is used with a fixed amount of other resources. The location of the MRP curve will depend on (1) the price of the product, (2) the productivity of the resource, and (3) the quantity of other factors working with the resource.

[3] Strictly speaking, this is true only for a variable resource that is employed with a fixed amount of another factor

Maximizing Profits When Multiple Resources Are Used

The same decision-making considerations apply when the firm employs several factors of production. The profit-maximizing firm will expand its use of a resource as long as the *MRP* of the resource exceeds its employment cost. If we assume that resources are perfectly divisible, the profit-maximizing decision rule implies that, in equilibrium, the *MRP* of each resource will be equal to the price of the resource. Therefore, the following conditions will exist for the profit-maximizing firm:

MRP of skilled labor = Price (wage rate) of skilled labor

MRP of unskilled labor = Price (wage rate) of unskilled labor

MRP of machine A = Price (explicit or implicit rental price) of machine A and so on, for all other factors

Cost Minimization When Multiple Resources Are Used

To maximize its profits, clearly the firm must produce the profit-maximizing output at the least possible cost. If the firm is minimizing costs, then the marginal dollar expenditure for each resource will have the same impact on output as an additional dollar expenditure on other resources used to produce the product. ***Factors of production will be employed such that the marginal product per last dollar spent on each factor is the same for all factors.***

To see why, consider a situation in which a $100 expenditure on labor caused output to rise by ten units, whereas an additional $100 expenditure on machines generated only a five-unit expansion in output. In this case, the firm can produce five more units of output by spending the $100 on labor instead of machines. By substituting labor for machines, it will reduce its *per-unit* cost.

If the marginal dollar spent on one resource increases output by a larger amount than a dollar expenditure on other resources, costs can always be reduced by substituting resources with a high marginal product per dollar expenditure for those with a low one. Substitution will continue to reduce unit costs (and add to profit) until the marginal product per dollar expenditure on each resource is equalized. This will occur because as additional units of a resource are hired, their marginal product will fall. Thus, the proportional relationship between the price of each resource and its marginal product will eventually be achieved.

Therefore, the following condition exists when per-unit costs are minimized:

$$\frac{\text{MP of skilled labor}}{\text{Price of skilled labor}} = \frac{\text{MP of unskilled labor}}{\text{Price of unskilled labor}} = \frac{\text{MP of machine A}}{\text{Price (rental value) of machine A}}$$

This relationship explains why workers with different skill levels earn different wages. If skilled workers are twice as productive as unskilled workers, then their wage rates will tend to be twice those of unskilled workers. For example, suppose that a construction firm hiring workers to hang doors is choosing among skilled and unskilled workers. If skilled door hangers can complete four doors per hour, while unskilled workers can hang only two doors per hour, a cost-minimizing firm would hire only skilled workers—as long as their wages are less than twice the wages of unskilled workers. On the other hand, only unskilled workers would be hired if the wages of skilled workers are more than twice that of unskilled workers. Because of competition, wages across skill categories will tend to mirror productivity differences.

Low wages do not necessarily mean low cost. In other words, it is not always cheaper to hire the lowest wage workers. It is not just wages, but rather wages *relative to productivity* that matter. If the wages of skilled workers are twice those of unskilled workers, it will still be cheaper to hire additional skilled workers if their marginal productivity (output per hour) is more than twice that of the unskilled workers.

The importance of wages relative to productivity explains why relatively few firms moved to Mexico following the passage of the North American Free Trade Agreement (NAFTA). Some forecasted there would be a "giant sucking sound" caused by the movement of both firms and jobs to Mexico. Why didn't the low wages of Mexican workers cause firms to relocate? *Answer:* Although wages are low in Mexico, so, too, is worker

productivity. Because of this, many firms are able to achieve lower production costs in the high-wage United States than in low-wage Mexico. Suppose that the average wage rate of a U.S. worker is $12 per hour and average hourly productivity is thirty-six units, while the average wage is $4 per hour in Mexico and average productivity is eight units. To maximize profits (or minimize costs), a firm should choose to operate where the *MP/P* is greatest. In the United States, the firm would get three units of output (36/12) per dollar spent on labor. In Mexico, the firm would get only two units of output (8/4) per dollar spent on labor. Thus, a cost-minimizing firm would want to locate in the United States despite the higher wages, because the productivity difference more than makes up for the wage difference. Although U.S. wages are three times higher, the productivity of workers in the United States is four and a half times higher. This more than compensates for the higher wage cost.

The Main Conclusion of the Marginal Productivity Theory of Employment

Firms minimize their per-unit costs of production when they hire additional units of each resource as long as the units' marginal productivity generates revenues in excess of their costs. Firms that minimize per-unit costs and maximize profit will never pay more for a unit of input, whether it is skilled labor, a machine, or an acre of land, than the input is worth to them. The worth of a unit of input to the firm is determined by how much additional revenue (marginal revenue product) is expected from its employment.

In the real world, it's sometimes difficult to measure the marginal product of a resource. And, as we have said, the marginal product of a resource is influenced by the other factors with which it is employed. But do decision makers really think like this—in terms of equating the marginal product/price ratio (*MP/P*) for each factor of production? Probably not. Their thought process is probably going to be something more like this: "Can we reduce costs by using more of one resource and less of another?" However, regardless of how managers think about the problem, when a firm maximizes its profits and minimizes its costs, its marginal product/price ratio will be equal for all factors of production. The results will be *as if* the firm followed the cost-minimization, decision-making rule just presented. Furthermore, competitive forces will more or less force firms to follow this rule—whether or not they're aware they're following it. Firms that fail to do so will be unable to compete successfully with rivals achieving lower per-unit costs.

The marginal productivity theory explains why the demand for some resources is higher or lower than the demand for other resources. Of course, resource prices will also be influenced by the supply of resources. We now turn to that topic.

THE SUPPLY OF RESOURCES

Just as benefits and costs direct the choices of employers, so, too, will they guide the actions of resource suppliers. Resource owners will supply their services to an employer only if they perceive that the benefits of doing so exceed their costs (the value of the other things they could do with their time or resources). Because of this. employers must offer resource owners at least as good a deal as they can get elsewhere. For example. if an employer does not offer a potential employee pay, benefits, and working conditions as good as or better than others he or she can get, the employer will be unable to hire that worker.

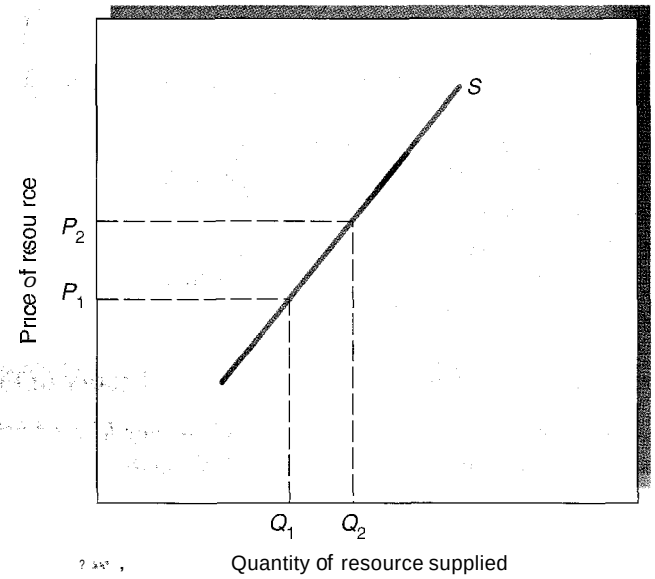
Resource owners will supply their services to those who offer them the best employment alternative. Other things being constant, as the price of a specific resource (for example, engineering services, craft labor, or wheat farmland) increases, the incentive of potential suppliers to provide the resource increases.

An increase in the price of a resource will lure resource suppliers into the market. A decrease will cause them to shift into other activities. Therefore, as **Exhibit 6** illustrates, the supply curve for a specific resource will slope upward to the right.[4]

[4]Although the supply for nonhuman resources will always slope upward, the supply of labor at very high wage rates can become backward bending. As wages rise, individuals will substitute toward more work, but simultaneously the higher income will cause them to desire more leisure. At very high wage rates the income effect might dominate, causing a negative relationship between wage rates and quantity of labor supplied in this range. For example, at a wage of $10,000 per hour, many individuals would probably supply fewer hours of work than at $500 per hour!

EXHIBIT 6
The Supply of a Resource

As the price of a resource increases, individuals have a greater incentive to supply it. Therefore, a direct relationship will exist between the price of a resource and the quantity supplied.



Short-Run Versus Long-Run Resource Supply

Like demand, supply in resource markets can be different in the short run than in the long run. If the wage rate for certified public accountants (CPAs) rose, for example, we would expect more workers supplying their services as CPAs. But where do these additional CPAs come from? In the short run, the additional supply must come from people who are already CPAs but are currently doing other things. The higher wages might induce some college accounting professors and some stay-at-home spouses with accounting credentials to move into employment as CPAs. In the short run, however, there isn't enough time to alter the availability of a resource through investment in human and physical capital. In contrast, in the long run, resource suppliers have time to adjust their investment choices in response to a change in resource prices. With time, the higher wages for CPAs will cause more students to major in accounting and other people to pursue the coursework needed to become CPAs. Higher resource prices will increase the quantity supplied in both the short run and the long run, but the response will be greater in the long run. Therefore, as **Exhibit 7** shows, the long-run supply of a resource will be more elastic than the short-run supply.

Short-Run Supply

Resource mobility
The ease with which factors of production are able to move among alternative uses. Resources that can easily be transferred to a different use or location are said to be highly mobile. Resources with few alternative uses are immobile.

The short-run supply response to a change in price is determined by how easily the resource can be transferred from one use to another—that is, **resource mobility**. The supply of resources with high mobility will be relatively elastic, even in the short run. On the other hand, resources that have few alternative uses (or are not easily transferable) are said to be immobile. The short-run supply of immobile resources will be highly inelastic.

Consider the mobility of labor. Within a skill category (for example, plumber, store manager, accountant, or secretary), labor will be highly mobile within the same geographic area. Movements between geographic areas and from one skill category to another are more costly to accomplish, though. Labor will thus be less mobile for movements of this variety. In addition, because it is easier for a high-skilled person to perform effectively in a lower-skill position than vice versa, short-run mobility will tend to decline as the skill level of the occupation rises. Thus, the short-run supply curve in high-skill occupations like that of an architect, mechanical engineer, and medical surgeon is usually quite inelastic.

What about the mobility of land? Land is highly mobile among uses when location doesn't matter. For example, the same land can often be used to raise corn, wheat, soybeans, or oats. As a result, the supply of land allocated to the production of each of these commodities will be highly responsive to changes in their relative prices. Undeveloped

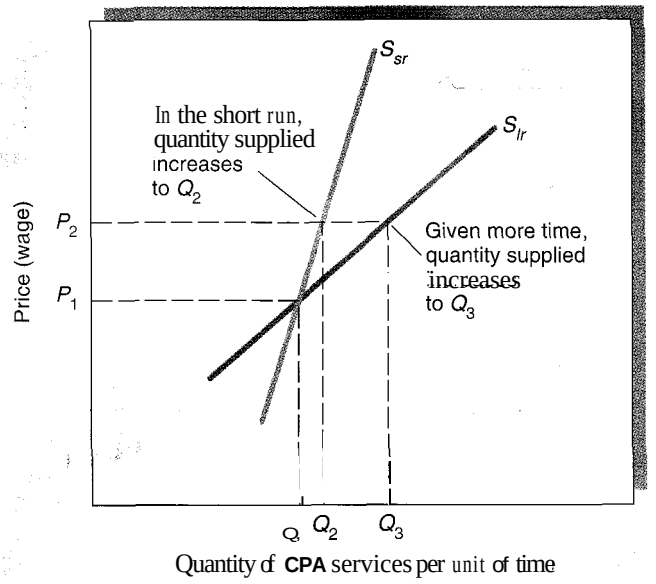


EXHIBIT 7
Time and the Elasticity of Supply for Resources

If the wage rate for certified public accountants (CPAs) rises, for example, to P_2, we would expect more workers to supply CPA services. Because it takes time to be trained as a CPA, though, the quantity of CPA services supplied in the short run (S_{sr}) won't increase by much—just to Q_2. The supply of CPA services is therefore relatively inelastic in the short run. In the long run, though, it is more elastic and the quantity supplied increases to Q_3.

land on the outskirts of cities is particularly mobile among uses. It can be used for agriculture, but it can also quickly be subdivided and used for a housing development or a shopping center. Because land is physically immobile, you might think its supply is unresponsive to changes in price when it comes to how desirable it is (or isn't) due to its location. You can't move it to make it more desirable and get a better price, obviously. You can, however, expand its usable space by constructing multiple-story buildings, for example. As the demand for a given location increases, higher and higher multilevel construction will be justified. This is why tall buildings are generally located in city centers, for example—because the demand for space is highest in these locations.

Machines are typically not very mobile among uses. A machine developed to produce airplane wings is seldom of much use when it comes to producing automobiles, appliances, or other products. Steel mills cannot easily be converted to produce aluminum. There are, of course, some exceptions. Trucks can typically be used to haul a variety of products. Building space can often be converted from one use to another. In the short run, however, immobility and inelasticity of supply characterize much of our physical capital.

Long-Run Supply

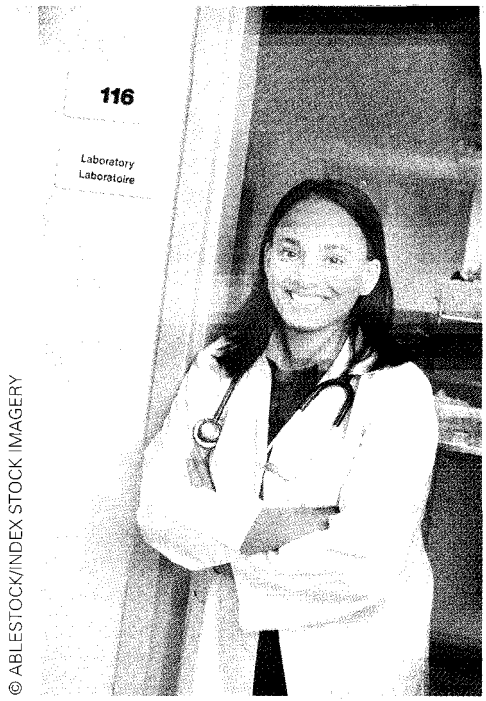
In the long run, the supply of resources can change substantially. Machines wear out. human skills depreciate, and even the fertility of land declines with use and erosion. These factors reduce the supply of resources. On the other hand. investment can expand the supply of productive resources, including machines, buildings. and durable assets. Correspondingly, investments in training and education can, over time. develop and improve the skills of future labor force participants. Thus, the supply of both physical and human resources in the long run is determined primarily by investment and depreciation.

As the price of a resource increases, more and more people will make the investments necessary to supply the resource. This will be true for human as well as physical resources. Examples abound. As the spread of the computer revolution pushed up the salaries of programmers, systems analysts, and computer technicians during the early 1990s. there was a sharp increase in the number of students training for jobs in these areas. In the late 1990s and early 2000s, attractive salaries for registered nurses led to both new programs and expanded enrollments. Higher salaries for lawyers stimulate law school enrollments. According to Harvard University economist Richard Freeman, a 1 percent increase in starting law salaries causes enrollment in the first year of law school to rise by 2 percent.

The long run, of course, is not a specified length of time. Investment can increase the availability of some resources fairly quickly. For example, it does not take very long to train

The supply curve for truck drivers will be considerably more elastic than the supply curve for doctors. Can you explain why?

GETTY IMAGES



© ABLESTOCK/INDEX STOCK IMAGERY

additional over-the-road truck drivers. However, it takes a long time to train physicians, dentists, lawyers, and pharmacists. Higher earnings in these occupations may have only a small impact on their current availability. Additional investment will flow into these areas, but typically it will be several years before there is a substantial increase in the quantity supplied.

SUPPLY, DEMAND, AND RESOURCE PRICES

In a market economy, resource prices will, of course, be determined by the forces of supply and demand and bring the choices of buyers and sellers into line with each other. **Exhibit 8** illustrates how the forces of supply and demand push the market wage rates of engineers toward equilibrium, where quantity demanded and quantity supplied are equal. Equilibrium is achieved when the price (wage rate) for engineering services is P_1. Given the market conditions illustrated by Exhibit 8, an excess supply is present if the price of engineering services exceeds P_1. Some engineers will be unable to find jobs at the above-equilibrium wage. This excess supply of engineers will cause the wage rate for engineers to fall. This will cause some of the engineers to look for jobs in other fields, and quantity supplied will fall, pushing the market toward equilibrium. In contrast, if the resource price is less than P_1, excess demand is present. Employers are unable to hire the amount of engineering services they want at a below-equilibrium resource price. Rather than doing without the resource, employers will attempt to hire engineers away from other firms by bidding the price up to P_1 and thereby eliminating the excess demand.

How will a resource market adjust to an unexpected change in market conditions? Suppose that there is a sharp increase in the demand for houses, apartments, and office buildings. The increase in demand for these products will also increase the demand for resources required for their construction. Thus, the demand for resources such as steel, lumber, brick, and the labor services of carpenters, architects, and construction engineers will increase. **Exhibit 9** shows the increase in demand for new houses and buildings (part a) and the accompanying increase in demand for construction engineers. The market demand for the services of construction engineers increases from D_1 to D_2 (part b), and initially there is a sharp rise in their wages (price increases from P_1 to P_2). The higher wages will motivate additional people to get the education and training necessary to become a construction engineer. Over time, the entry of the newly trained construction engineers will

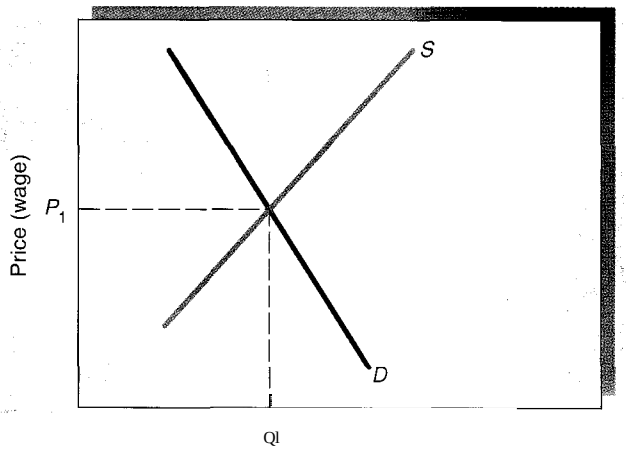
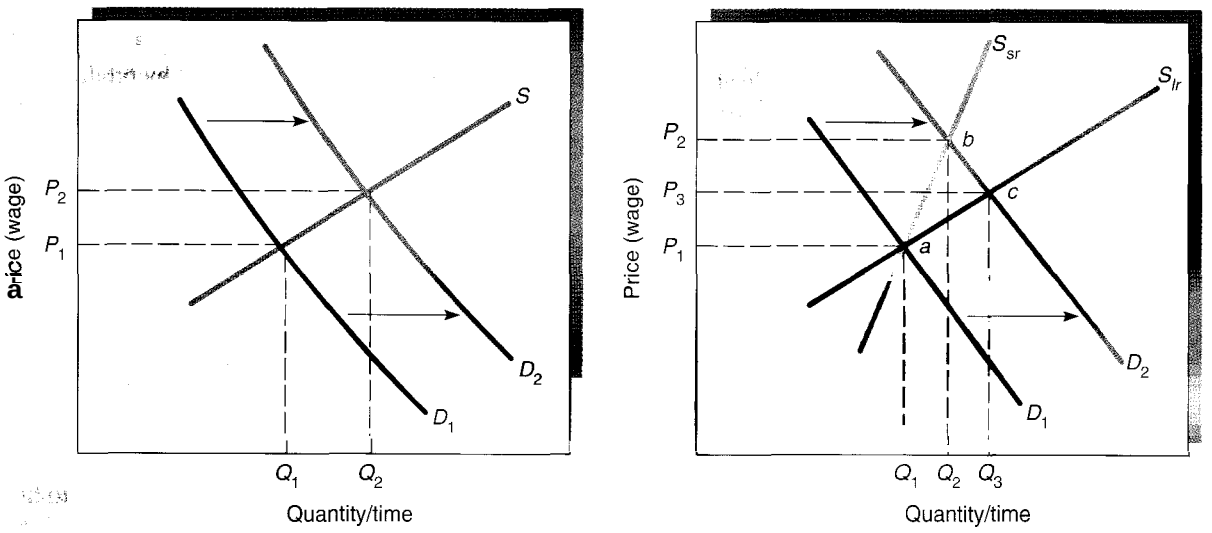


EXHIBIT 8
Equilibrium in a Resource Market

The market demand for a resource, such as engineering services, is a downward-sloping curve, reflecting the declining *MRP* of the resource. The market supply curve slopes upward because higher resource prices (wage rates) will motivate people to supply more of the resource. Market price will move toward equilibrium (P_1), where the quantity demanded and quantity supplied are in balance.

EXHIBIT 9
Adjusting to Dynamic Change

An increase in the demand for housing and commercial buildings (part a) will lead to an increase in demand for the services of construction engineers (part b) and other resources used in the construction industry. Initially, the increase in the resource price will be substantial, jumping from P_1 to P_2 (point *a* versus point *b* in the graph). The increase will be particularly sharp if the supply of the resource is highly inelastic in the short run. The higher resource price will attract additional human capital investment and, with time, the resource supply curve will become more elastic, which will moderate the price (or wage) increase to P_3 (point c).



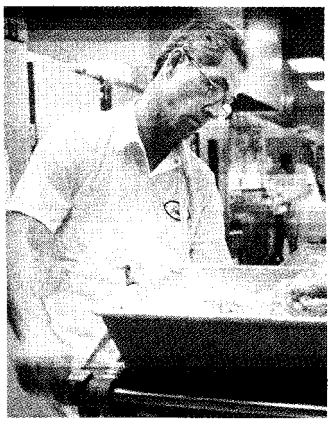
(a) Product market—new houses and commercial buildings

(b) Resource market—services of construction engineers

increase the elasticity of the resource supply curve. As these new construction engineers eventually enter the occupation, the supply curve will become more elastic (S_{lr} rather than S_{sr}), which will place downward pressure on wages in the occupation (the move from *b* to *c*). Therefore, as part (b) of Exhibit 9 illustrates, the long-run price increase (to P_3) will be less than the short-run increase (to P_2).

Workers in occupations such as nursing that require substantial skill and education will generally earn higher wages than those in occupations requiring little training or experience.

© DIGITAL VISION/PICTURE QUEST



© MARK ANDERSON/RUBBERBALL PRODUCTIONS/PICTURE QUEST

The supply response and market adjustment for other resources—physical as well as human—will be similar. For example, an unexpected construction boom will generally cause sharp initial increases in the prices of lumber, bricks, and other building materials. With time, however, additional investment will increase the availability of these resources and moderate the increase in their price, just as additional investment in human capital eventually moderated the wage increases of the construction engineers. The market adjustment to an unexpected reduction in demand for a resource is the same. The price falls further in the short run than in the long run. At the lower price, some resource suppliers will use their talents in other areas, so the incentive for potential new suppliers to offer the resource will be less. With time, the quantity of the resource supplied will decline, moderating the reduction in its price. Those with the poorest alternatives (that is, lowest opportunity cost) will continue to provide the resource at the lower prices. Those with better alternatives will look for other opportunities.

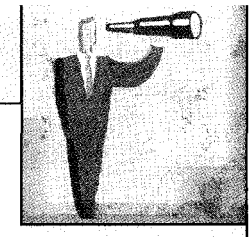
THE COORDINATING FUNCTION OF RESOURCE PRICES

Throughout this text, we have stressed that profit is a reward earned by producers who increase the value of resources, whereas loss is a punishment imposed on producers who reduce the value of resources they are using. The key links in this process are the prices of the products being sold and the prices of the resources used in production. As we have learned, a firm's profits are its revenues (which are determined by the product's sales price) minus its costs (which are determined by the prices of the resources it uses). The price of the product measures the value that consumers place on that product. The price of the resources, however, measures the value that consumers place on *other products* that could be produced with those same resources. Let us explore the importance of this link in a little more detail.

As we have shown, the price of a resource will equal the resource's marginal revenue product when the resource market is in equilibrium. The resource's marginal revenue product depends upon the price consumers are willing to pay for the output (how they value it) produced by the resource. We also learned that when a firm wishes to hire a resource, it must offer the owner of the resource a price at least as attractive as the resource could have earned elsewhere—that is, the resource's *MRP* in its next best alternative employment. Thus, the price a firm pays for a resource is equal to the resource's value (as measured by the consumer) in the alternative use. If the output the firm produces with that resource can be sold at a higher price than the price of the alternative outputs, then and only then will the firm earn a profit. Thus, profit is a reward to those entrepreneurs who are able to see and act on opportunities to put resources to higher-valued uses. Because consumer tastes and preferences continuously change, though, so do product and resource prices. As a result, opportunities for their use are created and destroyed on a daily basis in a market economy.

Pulling things together, our analysis indicates that prices in resource markets play a vitally important role. These prices coordinate the actions of the firms demanding factors of production and the households supplying them. Resource prices provide users with both information about the scarcity of the resources they're using and the incentive to conserve them in production. They also provide suppliers with an incentive to learn skills and provide resources—particularly those that are intensely demanded by users. Without the use of resource markets and the price incentives they provide, efficient use and wise conservation of resources would be extremely difficult to achieve.[5]

LOOKING AHEAD



The analysis of this chapter can be applied to a broad range of economic issues. The next chapter will focus on the labor market and earnings differences among workers. Later; we will focus on the capital market and the allocation of resources over time. The operation of these two markets plays an important role in determining the distribution of income, a topic that we will also analyze in detail in a subsequent chapter:

KEY POINTS

- Productive assets and services are bought and sold in resource markets. There are two broad classes of productive resources: (1) nonhuman capital and (2) human capital.
- The demand for resources is derived from the demand for products that the resources help produce. The quantity of a resource demanded is inversely related to its price because of substitutions made by both producers and consumers.
- The demand curve for a resource, like the demand for a product, can shift. The major factors that increase the demand for a resource are an increase in (1) demand for products that use the resource, (2) the productivity of the resource, and (3) the price of substitute resources.
- Profit-maximizing firms will hire additional units of a resource up to the point where the marginal revenue product (*MRP*) of the resource equals its price. With multiple inputs, firms will expand their use of each until the marginal product divided by the price (*MP/P*) is equal across all inputs. When real-world decision makers minimize per-unit costs, the outcome will be as if they had followed these mathematical procedures, even though they may not consciously have done so.
- The amount of a resource supplied will be directly related to its price. The supply of a resource will be more elastic in the long run than in the short run. In the long run, investment can increase the supply of both physical and human resources.
- The prices of resources are determined by supply and demand. Changes in the market prices of resources will influence the decisions of both users and suppliers. Higher resource prices give users a greater incentive to turn to substitutes and suppliers a greater incentive to provide more of the resource.
- Changes in resource prices in response to changing market conditions are essential for the efficient allocation of resources in a dynamic world. Profit is a reward to the entrepreneur who is able to see and act on opportunities to put resources to higher-valued uses.

[5] Analysis of energy consumption under central planning illustrates this point. The centrally planned economies used more energy per unit of output and their energy consumption was less responsive to changes in price than it was in economies that used markets to allocate energy. For evidence on these points, see Mikhail S. Bernstam, *The Wealth of Nations and the Environment* (London: Institute of Economic Affairs, 1991).

CRITICAL ANALYSIS QUESTIONS

1. "The demand for resources is a derived demand." What is meant by that statement? Why is the employment of a resource inversely related to its price?

2. How does a firm decide whether or not to employ an additional unit of a resource? What determines the combination of skilled and unskilled workers employed by a firm?

*3. Use the information in Exhibit 4 of this chapter to answer the following:
   a. How many employees (operators) would Compute-Accounting hire at a weekly wage of $250 if it were attempting to maximize profits?
   b. What would the firm's maximum profit be if its fixed costs were $1,500 per week?
   c. Suppose that there were a decline in demand for accounting services that reduced the market price per monthly statement to $150. At this demand level, how many employees would Compute-Accounting hire at $250 per week in the short run? Would Compute-Accounting be able to stay in business at the lower market price? Explain.

*4. Are productivity gains the major source of higher wages? If so, how does one account for the rising real wages of barbers, who, by and large, have used the same techniques for a half-century? *(Hint:* Do not forget opportunity cost and supply.)

5. Are the following statements both correct? Are they inconsistent with each other? Explain.
   a. "Firms will hire a resource only if they can make money by doing so."
   b. "In a market economy, each resource will tend to be paid according to its marginal product. Highly productive resources will command high prices, whereas less productive resources will command lower prices."

*6. Many school districts pay teachers on the basis of their highest degree earned and number of years of service (seniority). They often find it quite easy to fill the slots for English and history teachers, but very difficult to find the required number of math and science teachers. Can you explain why?

7. Suppose that you were the manager of a large retail store that was currently experiencing a shoplifting problem. Every hour, approximately $15 worth of merchandise was being stolen from your store. Suppose that a security guard would completely eliminate the shoplifting in your store. If you were interested in maximizing your profits, should you hire a security guard if the wage rate of security guards were $20 per hour? Why or why not? What does this imply about the relationship between average shoplifting per hour in the economy and the wage rates of security guards?

*8. A dressmaker uses labor and capital (sewing machines) to produce dresses in a competitive market. Suppose the last unit of labor hired cost $1,000 per month and increased output by 100 dresses. The last unit of capital hired (rented) cost $500 per month and increased output by 80 dresses. Is the dressmaker minimizing costs? If not, what changes need to be made?

9. A firm is considering moving from the United States to Mexico. The firm pays its U.S. workers $12 per hour. Current U.S. workers have a marginal product of forty, whereas the Mexican workers have a marginal product of ten. How low would the Mexican wage have to be for the firm to reduce its wage cost per unit of output by moving to Mexico?

*10. "The earnings of engineers, doctors, and lawyers are high because lots of education is necessary to practice in these fields." Evaluate this statement.

11. Other things being constant, what impact will a highly elastic demand for a product have on the elasticity of demand for the resources used to produce the product? Explain.

*12. The following chart provides information on a firm that hires labor competitively and sells its product in a competitive market.

Units of Labor	Total Output	Marginal Product	Product Price	Total Revenue	*MRP*
1	14	____	$5	____	____
2	26	____	$5	____	____
3	37	____	$5	____	____
4	46	____	$5	____	____
5	53	____	$5	____	____
6	58	____	$5	____	____
7	62	____	$5	____	____

   a. Fill in the missing figures.
   b. How many units of labor would be employed if the market wage rate were $40? Why?
   c. What would happen to employment if the wage rate rose to $50? Explain.

13. Leisure Times Inc. employs skilled workers and capital to install hot tubs. The capital includes the tools and equipment workers use to construct and install the tubs. The installation services are sold in

CHAPTER 13

Earnings, Productivity, and the Job Market

How can I be overpaid? The boss wouldn't pay me that amount if I wasn't worth it.

—Jackie Gleason[1]

Chapter Focus

- Why do some people earn more than others?
- Are earnings differences according to race and gender the result of employment discrimination?
- Why are wages higher in the United States than in India or China?
- What is the source of higher wages? Why has the growth rate of both wages and income per capita increased during the past decade?
- Does automation destroy jobs?

[1]The above statement was mdde in response to a question dbout the amount he was being paid for the popular television show *The Honeymooners,* which ran during the 1950s and 1960s.

The earnings of U.S. workers are among the highest in the world. However, they vary widely. An unskilled laborer may earn $8 per hour, or even less. Lawyers and physicians often earn $150 per hour, or more. Dentists and even economists might receive $100 per hour. How can these variations in earnings be explained? Why are the earnings of Americans so high? How have earnings changed in recent years, and what are the factors underlying these changes? This chapter will address these topics and related issues.

WHY DO EARNINGS DIFFER?

The earnings of individuals in the same occupation or with the same amount of education often differ substantially. So do the earnings of people in the same family. For example, one researcher found that the average annual earnings differential between brothers was $29,564, compared with $32,742 for men paired randomly.[2] The earnings of people with the same intelligence, level of training, or amount of experience also typically differ.

Several factors combine to determine a person's earning power. Some seem to be the result of good or bad luck. Others are clearly the result of conscious decisions people make. In the previous chapter, we analyzed how the market forces of supply and demand operate to determine resource prices. Wages are a resource price, and therefore the supply and demand model can be used to examine earnings differentials among workers.

For simplicity, we have proceeded as if employees earned only money payments. In reality, most workers receive a compensation package that includes **fringe benefits** as well as money wages. The fringe benefit component typically includes things like medical insurance, life insurance, pension benefits, and paid vacation days. When we use the terms wages and earnings in the following discussions, we are referring to the total compensation package that includes both wages and fringe benefits.

Fringe benefits
Benefits other than normal money wages that are supplied to employees in exchange for their labor services. Higher fringe benefits come at the expense of lower money wages.

The **real earnings** of all employees in a competitive market economy would be equal if: (1) all individuals were identical in preferences, skills, and background, (2) all jobs were equally attractive, and (3) workers were perfectly mobile among jobs. Given these conditions, if higher real wages existed in an area of the economy, the supply of workers in that area would expand until the wage differential were eliminated. Similarly, low wages in an area would cause workers to exit until wages there returned to normal. Of course, that's not the way things work in the real world. We all know that earnings differences are a fact of life. There are three reasons for this, and we'll discuss each next.

Real earnings
Earnings adjusted for differences in the general level of prices across time periods or geographic areas. When real earnings are equal, the same bundle of goods and services can be purchased with the earnings.

Earnings Differentials Due to Nonidentical Workers

Workers differ in several important respects that affect both the supply of and demand for their services. In turn, these factors affect their wage rates. Let's consider these differences.

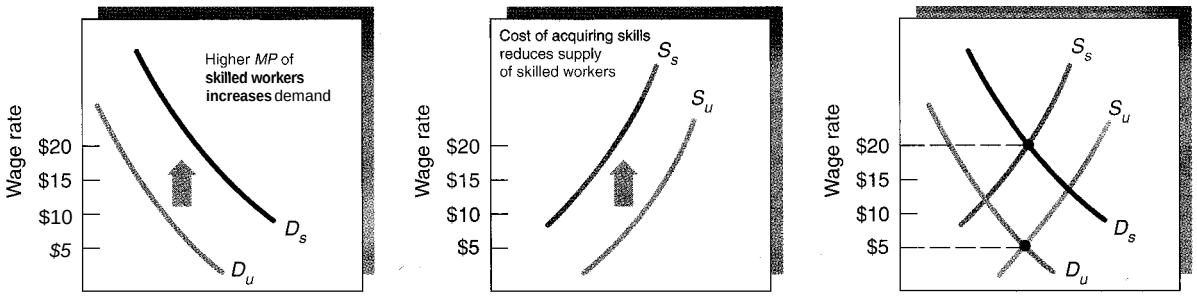
1. Worker productivity and specialized skills. The demand for employees who are highly productive is greater than the demand for those who are less productive. People who can operate a machine more skillfully, hit a baseball more consistently, or sell life insurance policies with greater regularity are more valuable to employers. Compared to their less skillful counterparts, these employees contribute more to the firm's revenue. Put another way, the marginal revenue product *(MRP)* of the more productive employees is higher than the *MRP* of less productive ones. In competitive labor markets, workers earn a wage equal to their marginal revenue product. As a result, the labor services of more productive workers will command higher wages in the marketplace.

Worker productivity is the result of a combination of factors, including native ability, parental training, hard work, and investment in human capital. The link between higher productivity and higher earnings motivates people to invest in themselves in order to

[2]Christopher Jencks, *Inequality* (New York Basic **Books,** 1972), 220 Salary figures are in 2003 dollars

EXHIBIT 1
Demand, Supply, and Wage Rates for Skilled and Unskilled Workers

The productivity—and therefore marginal product (*MP*)—of skilled workers is greater than that of unskilled workers. As a result, as part (a) illustrates, the demand for skilled workers (D_s) will exceed the demand for unskilled workers (D_u). Education and training generally enhance skills. Because upgrading one's skills through investments in human capital is costly, the supply of skilled workers (S_s) is smaller than the supply of unskilled workers at any given wage (part b). As part (c) illustrates, the wages of skilled workers are high relative to those of unskilled workers due to the strong demand and small supply of skilled workers relative to unskilled workers. (*Note:* The quantity of skilled labor employed may be far smaller, far larger, or, by accident, equal to the quantity of unskilled labor hired.)



(a) Demand for skilled and unskilled labor

(b) Supply of skilled and unskilled labor

(c) Wages of skilled and unskilled labor

upgrade their knowledge and skills. If additional worker productivity didn't lead to higher earnings, people would have little incentive to incur the costs of doing so, including paying to go to college or get additional training.

Exhibit 1 shows the demand, supply, and wage rates for skilled versus unskilled workers. The productivity of skilled workers exceeds that of unskilled workers, and so does the demand for them. The vertical distance between the demand curve for skilled workers (D_s) and the demand curve for unskilled workers (D_u), shown in part (a), reflects the higher marginal product (*MP*) of skilled workers. Because human capital investments (education, training, and so on) are costly, the supply of skilled workers (S_s) will be smaller than the supply of unskilled workers (S_u). This is shown in part (b). The vertical distance between the two supply curves is the wage differential employers will need to pay skilled workers for the costs they incurred to acquire their skills.

Wages are determined by demand relative to supply (part c). Since the demand for skilled workers is large while their supply is small, the equilibrium wage of skilled workers will be high ($20 per hour). In contrast, since the supply of unskilled workers is large relative to the demand, their wage rates will be substantially lower ($5 per hour).

The skills one person gains from a year of education, vocational school, or on-the-job training might be better or worse than those of someone else. Therefore, we should not expect a rigid relationship to exist between years of education (or training) and skill level. On average, however, there is a strong positive relationship between investment in education and earnings.

Exhibit 2 presents annual earnings data according to educational level for year-round, full-time workers in 2001. Notice that the earnings of both men and women increased consistently with additional schooling. High school graduates earned almost 50 percent more than their counterparts with less than a high school education. Male college graduates working full-time, year-round earned $70,253, compared to $37,362 for men with only a high school education. In the case of women, college graduates earned $45,290, compared to $26,660 for those who only graduated from high school. The earnings of both men and women continued to increase as they earned master's and doctoral degrees.

In addition to education, a person's intelligence, native ability, and motivation also affect his or her productivity and earnings. Research shows, however, that much of the extra income earned by high-wage earners is, in fact, the result of the knowledge and skills they acquired by making an investment in additional education. (See the accompanying Applications in Economics feature, **"A** College Degree as a Job Market Signal.") Other studies show that on-the-job training enhances the earnings of workers.

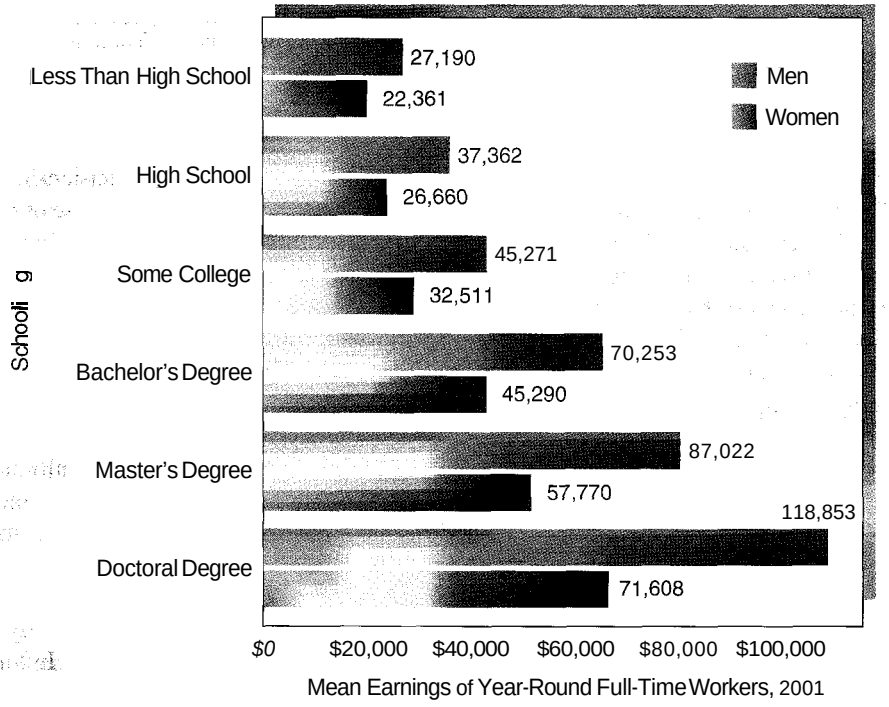


EXHIBIT 2
Education and Earnings

The accompanying graph presents data for the mean annual earnings of year-round, full-time workers based on their gender and education. Note that the earnings of both men and women increased with additional education. Even though the data are for full-time workers, the earnings of women were only about two-thirds those of men with similar education.

Source: U.S. Department of Commerce, Current Population Reports, P-60 Series. "Money Income in the United States: 2002," Table 9.

APPLICATIONS IN ECONOMICS

A College Degree as a Job Market Signal: Why You Should Take More Math

"Why should I take this difficult course?" "When will I ever use this?" Students often complain about taking courses not directly related to their future career. This complaint often reflects an incomplete understanding of exactly why college graduates do better in the Job market than those without a college degree. A college degree increases a person's earnings because of both (1) human capital—knowledge that will directly increase job productivity and (2) signaling—"signs" to employers about a person's attitude and motivational characteristics, as well as his or her general analytical skills.

Suppose that an employer is looking for an employee who is a very good analytical problem solver. Without a way to directly observe this ability, the employer will likely look for indicators that signal this attribute. For example, even if a Job doesn't directly require calculus, persons with a good calculus grade are likely to possess better problem-solving skills than those with a poor grade (or those who dodged the subject). Thus, employers may favor job applicants with good calculus grades, even if calculus isn't directly used on the job.

In other words, even if a college degree added nothing to the knowledge or skills required to do a particular Job, it could still help employers identify people with abilities that are difficult to observe. Similarly, students who are admitted to and graduate from elite universities like Harvard and Yale are likely to have superior abilities relative to those attending lower-level schools. Because of this signaling device, even mediocre graduates from top universities often have better entry-level Job market prospects than exceptional students from lesser-known schools. The signaling function also explains why students who choose majors others think are "hard" (like engineering, economics, or finance) generally do better in the Job market than those choosing "easy" majors (like physical education, marketing, or social work).

"When will I ever use this stuff?" "When will a good grade in challenging courses like math and economics matter?" The answer to these questions is, "Soon!" They will matter most when you're searching for a job at the beginning stages of your career—when it's still difficult for employers to judge your true abilities.

An investment in human capital and development of specialized skills can protect high-wage workers from the competition of others willing to offer their services at a lower price. Few persons could develop the specialized skills of a Steven Spielberg, Tiger Woods, or Oprah Winfrey. Similarly, the supply of heart surgeons, trial lawyers, engineers, business entrepreneurs, and many other specialized workers is limited in occupations in which specific skills, knowledge, and human capital investments contribute to job performance.

What about the large salaries received by CEOs of major corporations? Decisions made by CEOs can have a huge financial impact on their companies' profitability. Because of this, a good CEO can be particularly valuable when a company is in trouble, facing new competitors, or confronting a changing technological or regulatory environment. A CEO who turns losses into profits is worth millions to the stockholders of a major corporation.

The annual earnings of star athletes, entertainers, and television personalities often run into the millions of dollars. What is the marginal revenue product of a superstar entertainer like Cameron Diaz? How many more people will go see a movie starring Diaz rather than another actress less known or less talented? If an additional 2 million or 3 million people spend $8 to attend a movie, this will generate $16 million or $24 million of additional revenue. It is easy to see how hiring a "box office" star will generate a lot more money for a movie company. In turn, competitive labor markets will ensure that big stars get paid their marginal revenue product.

Tournament pay
A form of compensation in which the top performer (or performers) receives much higher rewards than other competitors, even if the others perform at only a slightly lower level.

Economic studies have found that the marginal revenue product of many sports and entertainment superstars is pretty much in line with their salaries.[3] However, there may be another factor at work here. The earnings in some markets resemble tournaments. In tournaments, only the top-ranked person receives the big payoff, while those who finish second receive much less. This type of compensation is called **tournament pay**.[4] This name refers to reward systems structured like golf tournaments, in which a slight difference in productivity (perhaps one or two shots) is associated with a difference in pay of several hundred thousand dollars.

What is the marginal revenue product of entertainment stars like Cameron Diaz? Do the earnings of star entertainers, athletes, and even business executives reflect the tournament pay nature of markets in these areas?

HUBERT BOESL/DPA/LANDOV

[3]Paul M. Sommers and Noel Quinton, "Pay and Performance in Major League Baseball: The Case of the First Family of Free Agents," *Journal of Human Resources* (Summer 1982): 426–35.

[4]Edward Lazear and Sherwin Rosen, "Rank Order Tournaments as an Optimum Labor Contract," *Journal of Political Economy* 89 (October 1981): 841–64; and Robert Frank and Phillip Cook, *The Winner-Take-All Society* (New York: The Free Press, 1995).

In this type of environment, workers basically subject themselves to a lottery where the winner (the person with the highest productivity) receives compensation higher than his or her marginal revenue product, while the losers receive less than their marginal revenue product. There are some constraints as to how low the compensation of the losers can be, however, because if their wages are depressed too low, they will shift to alternative job opportunities.

The tournament pay environment creates a strong incentive for potential superstars to expend considerable effort to become a top performer. As a result, many people spend long hours developing skills that will increase their chances of becoming a "star" in athletics, entertainment, professions, and business. The tournament system also encourages those who are unwilling to make such sacrifices to follow another path. Is this good or bad? Economics does not answer that question; it merely explains how these markets work.

2. Worker preferences. An important source of earnings differentials that is sometimes overlooked is worker preferences. People have different objectives in life. Some want to make a great deal of money. Many are willing to work two jobs or very long hours, undergo agonizing training and many years of education, or sacrifice social and family life to make money. Others might be "workaholics" because they enjoy their jobs. Still others might be satisfied with just enough money to get by, preferring to spend more time with their family, the Boy Scouts, watching television, on vacation, with a hobby, or at the local tavern.

Economics doesn't dictate that one set of worker preferences is more desirable than another, any more than it suggests that people should eat more spinach and less pastrami. It does, however, show that worker preferences in the areas of money, work, and skill development will contribute to differences in earnings. Other things being constant, persons who are more highly motivated by monetary objectives will be more likely to do the things necessary to command higher wage rates. But economics makes no judgments about whether this is good or bad.

3. Race and gender. Discrimination on the basis of race or gender also contributes to earnings differences among people. **Employment discrimination** can directly limit the earnings opportunities of minorities and women. Employment discrimination occurs when minority or women employees are treated differently than similarly productive whites or men. Of course, the earnings of minorities or women can differ from those of whites or men, respectively, for reasons other than employment discrimination. Nonemployment discrimination—including access to high-quality education or specialized training, for example—can limit the opportunities of minority groups and women to acquire human capital, enhancing their productivity and earnings. Other factors, like the limited opportunities that can result from growing up in a low-income or single-parent family, can also influence skill development and educational achievement. In a later section in this chapter, we will analyze the impact of employment discrimination in more detail.

Employment discrimination
Unequal treatment of persons on the basis of their race, gender, or religion, restricting their employment and earnings opportunities compared to others of similar productivity. Employment discrimination may stem from the prejudices of employers, customers, fellow employees, or all three.

Earnings Differentials Due to Nonidentical Jobs

When people evaluate employment alternatives, they consider working conditions as well as wage rates. Is a job dangerous? Does it offer the opportunity to acquire the experience and training that will enhance future earnings? Is the work strenuous and nerve-racking? Are the working hours, job location, and means of transportation convenient? These factors are what economists call **nonpecuniary job characteristics**. People will accept jobs with undesirable working conditions if the wages are high enough, compared to potential job alternatives with better working conditions. Because the higher wages, in essence, compensate workers for the unpleasant nonpecuniary attributes of a job, economists refer to wage differences stemming from this source as **compensating wage differentials**.

Examples abound of higher wages that compensate various workers for less-attractive working conditions. Because of the dangers involved, aerial window washers (those who hang from windows twenty stories up) earn higher wages than other window washers. Sales jobs involving a great deal of out-of-town travel typically pay more than similar jobs that aren't so inconvenient. Because the jobs are both physically demanding and sometimes dangerous, the wages of coal miners and sewer workers are generally higher than

Nonpecuniary job characteristics
Working conditions, prestige, variety, location, employee freedom and responsibilities, and other nonwage characteristics of a job that influence how employees evaluate the job.

Compensating wage differentials
Wage differences that compensate workers for risk, unpleasant working conditions, and other undesirable nonpecuniary aspects of a job.

those in other occupations available to low-skilled workers. Jobs in attractive locations pay less than similar jobs in less-attractive areas. For example, American truck drivers in post-war Iraq earned $80,000 per year tax free, which is much more than they could have earned in the United States. This high rate of compensation was the result of the dangerous conditions they faced in Iraq. Compensating factors even influence the earnings of economists. When economists work for colleges or universities, they generally enjoy a more independent and intellectually stimulating work environment than when they are employed in the business sector. Unsurprisingly, the earnings of academic economists are typically lower than those of business economists. However, it is important to remember that the academic economist with the lower earnings has chosen this job over business-sector employment. Thus, the lower earnings do not imply that the worker is worse off.

Earnings Differentials Due to the Immobility of Labor

It is costly to move to a new location or train for a new occupation in order to get a job. As a result, labor, like other resources, isn't perfectly mobile. Some wage differentials therefore result because wages have not yet adjusted fully to changes in market conditions.

For example, as we've learned, when strong demand leads to wage increases, more and more people will train for the occupation in high demand. Eventually, supply will increase and moderate the wage increases. But this process will take time. Just the opposite will happen when demand falls. When the demand for workers in a given occupation or skill category falls, wages are likely to fall sharply in the short run. With time, some workers will move into other occupations, which eventually moderates the reduction in wages. But this will also take time. Some of the wage differences reflect the fact that this adjustment process doesn't happen instantly.

Institutional barriers can also limit the mobility of labor. Licensing requirements, for example, limit the mobility of labor into many occupations—medicine, taxicab driving, architecture, and mortuary science among them. Since minimum wages raise the cost to employers of hiring workers, they may retard the ability of low-skilled workers to obtain employment in certain sectors of the economy. These restrictions on labor mobility will also influence the size of wage differentials. Labor unions often promote policies designed to increase the demand for union labor and reduce the labor supply in unionized job categories. Their actions, though, often limit the ability of nonunionized workers (and firms) to enter and compete in the unionized sectors of the economy. To the extent that unions are successful, they hamper the mobility of nonunionized workers and create higher wages for those who are unionized. For more on this topic, see the special topic "Do Labor Unions Increase the Wages of Workers?" later in this book.

THUMBNAIL SKETCH

What Are the Sources of Earnings Differentials?

Differences in Workers

1. Productivity and specialized skills that reflect native ability, parental training, and investment in human capital (education)
2. Worker preferences (the trade-off that workers are willing to make between money earnings and other factors)
3. Race and gender discrimination

Differences in Jobs

1. Location of job
2. Working conditions (such as job safety and comfort in the workplace)
3. Opportunity for training and skill-enhancing work experience

Immobility of Resources

1. Temporary disequilibrium resulting from market changes
2. Institutional restrictions (for example, occupational licensing and union-imposed restraints)

APPLICATIONS IN ECONOMICS

America's Millionaires[1]

The number of millionaires has been expanding rapidly. In 1975, there were only 350,000 households with a net worth of $1 million or more. By 1996, the figure had grown to 3.5 million. Only about one-third of the increase was due to inflation. The growth in the number of millionaires is expected to continue. Forecasts indicate that, by 2005, there will be 5.6 million millionaire households—5.2 percent of the total (1 out of every 20). Who are these millionaires?

- In terms of age, millionaires are typically in their late 50s. They tend to be older than the general population because it takes time to accumulate wealth of this magnitude. In addition, older workers have higher earnings, enabling them to increase wealth more rapidly.
- Not surprisingly, millionaires tend to be well educated. About 80 percent have a college degree. In fact, nearly two-fifths have a graduate degree. Those with more education have higher earnings, which is an important source for acquiring wealth.
- Millionaires are disproportionately self-employed entrepreneurs. While less than one-fifth of the workforce is self-employed, two-thirds of the millionaires fall into this category. To a degree, the income and wealth of millionaires are rewards that compensate them for assuming the greater financial risks that accompany self-employment and business ownership than those that accompany working for a salary.
- The vast majority of millionaires achieved their status through saving and investment. They save on average 20 percent of their income. Most of them are first-generation rich. Less than 20 percent received more than 10 percent of their wealth through an inheritance.

[1]Based on Thomas J. Stanley and William D. Danko, *The Millionaire Next Door. The Surprising Secrets of America's Wealthy* (New York: Longstreet Press, 1996).

Sources of Wage Differentials: A Summary

As the accompanying Thumbnail Sketch shows, wage differentials stem from many sources, which can be categorized in three main ways: differences in workers. differences in jobs, and immobility of resources. Many of the wage differentials in these categories play an important allocative role, compensating people for (1) human capital investments that increase their productivity or (2) unfavorable working conditions. Other wage differentials reflect, at least partially, locational preferences or the desires of individuals for higher money income rather than nonmonetary benefits. Still other differentials. like those related to discrimination and occupational restrictions, are unrelated to worker productivity or preferences and do not promote efficient production.

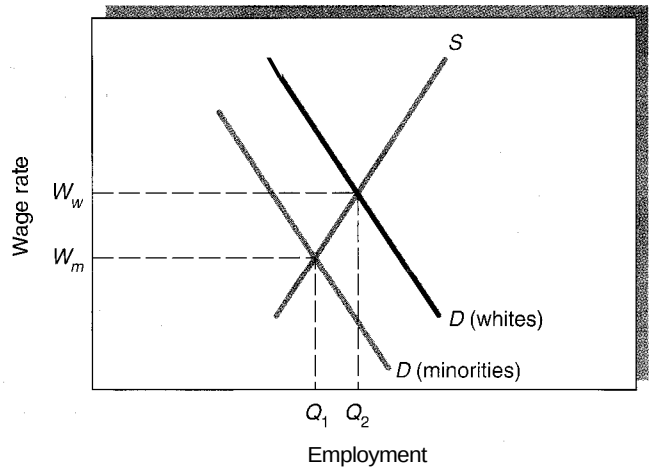
The analysis above focuses on factors that contribute to differences in real earnings. Nominal earnings will also be influenced by differences in the cost of living. In a large, geographically diverse country like the United States, the cost of living varies substantially across cities, regions, and communities. In cities like New York and San Francisco, the level of prices can be 50 percent or even 100 percent higher than in other parts of the country. Put another way, the quantity of goods and services that can be purchased with $50,000 of earnings is substantially less in New York City than in rural Georgia or Kansas. Geographical cost-of-living differences therefore also contribute to wage differences.

THE ECONOMICS OF EMPLOYMENT DISCRIMINATION

How does employment discrimination affect the job opportunities available to women and minorities? Do employers gain from discrimination? Economics sheds light on both these questions. There are two outlets for labor market discrimination: wage rates and employment restrictions. **Exhibit 3** illustrates the impact of wage discrimination. When nonminority workers are preferred to minority workers (or male to female workers), the demand for the latter groups falls and the wages of these people decline.

EXHIBIT 3
The Impact of Direct Wage Discrimination

If there is employment discrimination against minorities or women, then the demand for their services will decline, and their wage rate will fall from W_w to W_m.



Essentially, there are two labor markets — one market for the favored group and another for the group against which the discrimination is directed. The favored group, such as whites, is preferred, but the less expensive labor of minority workers is a substitute productive resource. Both white and minority employees are employed, but the whites are paid a higher wage rate.

Exclusionary practices are another form of employment discrimination. Either in response to outside pressure or because of their own views, employers might primarily hire whites and men for certain types of jobs. When minority and female workers are excluded from a large number of occupations, they are crowded into a smaller number of remaining jobs and occupations. If entry restraints prevent people from becoming supervisors, plumbers, electricians, and airline pilots, they will be forced to accept alternatives. Thus, the supply of labor in the unrestricted occupations increases, causing wage rates in these occupations to fall. In turn, the exclusionary practices reduce supply and push wages up in occupations and industries dominated by white men.

Discrimination is costly to employers when they are merely reflecting their own prejudices. ***If employers can hire equally productive minority employees (or women) at a lower wage than whites (or men), then the profit motive gives them a strong incentive to do so. Hiring the higher-wage whites when similar minority employees are available will increase the costs of firms that discriminate.*** Employers who hire employees regardless of their race or gender will have lower costs and higher profits than rival firms that try to fill positions with (mostly) white men. Thus, competitive forces tend to reduce the profitability of firms that discriminate.

Discriminatory hiring practices can stem from factors other than the employer prejudice, however. If either the firm's employees or its customers have a preference for or against various groups, this may lead to discriminatory hiring, even if the employer is totally unbiased. When discrimination is customer based, a worker from a favored group will be able to bring in more revenue for the firm. For example, adult nightclubs that hire attractive young women as dancers will generate more revenue than those that hire dancers from all age and gender groups. Similarly, Chinese restaurants that hire all (or almost all) Chinese servers are likely to do better than if the ethnic and racial composition of their employees mirrors that of the labor force. Historically, customer-based discrimination has often gone unchallenged. This appears to be changing. In a highly publicized case, the Hooters restaurant chain was charged with discriminating against male servers. Although the Equal Employment Opportunity Commission dropped its four-year investigation, Hooters agreed to pay $3.75 million in damages and begin hiring male servers as the result of privately filed lawsuits in Illinois and Maryland.

How Much Impact Does Employment Discrimination Have on Earnings?

If we want to isolate the impact of employment discrimination, we must (1) adjust for differences between groups in education, experience, and other productivity-related factors and (2) then make comparisons between similarly qualified groups of employees who differ only with regard to race (or gender).

How do the earnings of minorities compare with those of similarly productive whites? **Exhibit 4** presents the actual wages of minorities relative to whites and the "productivity-adjusted" minority/white wage ratio. The adjusted ratio is an estimate of how the wages of minorities would compare with those of whites if the two groups had the same productivity characteristics (schooling, work experience, marital status, regional location, and union and industry status). In 2000–2003, the actual wages of black men were 79 percent of the wages of white men. When the workforce characteristics of black men were taken into account, however, the adjusted hourly earnings of black men rose to 86 percent of the earnings of white men. This implies that productivity-related factors accounted for a little more than one-quarter of the wage differential between the two groups. A 14 percent differential, which may well be the result of employment discrimination, remained after adjustment for the productivity characteristics.[5]

Mexican-Americans constitute the second-largest minority group in the United States. Even though the actual wages of Mexican-American men were only 67 percent of the wages of white men, their "adjusted" earnings were 93 percent of those for whites. These data were not adjusted for the ability to speak English. Adjustment for this factor would almost certainly further narrow the differential. This suggests that when Mexican-American men possess the same worker characteristics as white men, their earnings are nearly on par with those of their white counterparts. Interestingly, adjustment for productivity factors reduces the wages of Asian-Americans relative to whites. Asian-Americans have greater productivity characteristics, particularly educational levels, than whites. As a result, the "adjusted" earnings of Asian-Americans are lower than their actual earnings figures.

Turning to the data for women, the productivity-adjusted wage rates of minority women relative to white women are between 94 percent and 98 percent for each of the

EXHIBIT 4
The Actual and Productivity-Adjusted Wages of Minorities Compared to Whites: 2000–2003.

	Men		Women	
	Actual	Adjusted	Actual	Adjusted
White	100	100	100	100
African-American	79	86	90	94
American Indian	82	92	87	95
Asian-American[a]	100	94	103	97
Mexican-American	67	93	75	98
Other Hispanic	79	92	85	95

[a]Primarily Chinese-Americans and Japanese-Americans.
Source: These data were supplied by David Macpherson. They were derived from the 2000–2003 *Current Population Surveys.* The data were adjusted for years of schooling, work experience, region, industry, sector of employment, union status, and marital status.

[5]The figures presented in Exhibit 4 do not control for the lower average quality of schooling received by blacks. Other researchers using more refined data have found that productivity factors account for a larger share of the earnings differential between whites and blacks. For evidence on this point, see Francine D. Blau and Lawrence M. Kahn, "Race and Gender Pay Differentials," in *Research Frontiers in Industrial Relations and Human Resources,* ed. David Lewin, Olivia S. Mitchell, and Peter D. Scherer, 381416 (Madison, Wis.: Industrial Relations Research Association, 1992); and Derek A. Neal and William R. Johnson, "The Role of Premarket Factors in Black-White Wage Differences," *Journal of Political Economy* (October 1996): 869–95. For a detailed explanation of how the adjusted ratios of Exhibit 4 were derived and information on the significance of productivity factors and employment discrimination on the basis of gender, see David A. Macpherson and Barry T. Hirsch, "Wages and Gender Composition: Why Do Women's Jobs Pay Less?" *Journal of Labor Economics* (July 1995).

groups included in Exhibit 4. This implies that employment discrimination on the basis of race adds little to earnings differentials that may reflect discrimination on the basis of gender. For an analysis of earnings differences according to gender, see the special topic "Is Discrimination Responsible for the Earnings Differences between Men and Women?" later in this book.

LINK BETWEEN PRODUCTIVITY AND EARNINGS

Link between Productivity and Earnings

In a market economy, productivity and earnings are closely linked. To earn a large income, one must provide large benefits to others.

In a competitive market setting, productivity—that is, output per worker—and earnings are closely linked. When workers are more productive, the demand for their services will be higher, and therefore they will be able to command higher wages. ***High productivity is the source of high wages. When the outputper hour of workers is high, the real wages of the workers will also be high.***

In turn, the linkage between productivity and earnings provides individuals with a strong incentive to develop their talents and utilize their resources in ways that are helpful to others. As the value of the goods and services supplied to others increases, there will also be a tendency for one's earnings to increase. If you want to make a large income, you had better figure out how to provide services that are highly valued. Self-interest is a powerful motivator and, as Adam Smith noted long ago, competitive markets can bring it into harmony with economic progress.

Productivity differences are an important source of differences in earnings among individuals. They are also an important source of earnings differences across countries. For example, the earnings per worker are vastly greater in the United States than they are in India or China because the output of U.S. workers is much greater than the output of their counterparts in those countries. The average worker in the United States is better educated, works with more productive machines, and benefits from more efficient economic organization than the average person in India or China. Thus, the value of the output produced by the average U.S. worker is approximately fifteen times that produced by the average worker in India or China. American workers earn more because they produce more. If they did not produce more, they would not be able to earn more.

GETTY IMAGES

Labor-saving equipment can improve productivity. But profit-maximizing firms will adopt automated production methods and high-tech equipment only when they reduce costs. Cost-effective automation releases labor and other resources so that they can be used to expand production in other areas. In turn, the higher worker productivity and expansion in production makes higher income levels and living standards possible.

Productivity differences also affect earnings across time. Today, American workers are substantially more productive than they were fifty years ago.[6] The output of goods and services per hour of U.S. workers in *2003* was approximately twice the level of the mid-1950s. Similarly, average real earnings (total compensation) per hour in *2003* were approximately double those of fifty years ago. Earnings rose because productivity increased. If productivity had not increased, the increase in earnings would not have been possible.

Increasedphysical capital, improvements in the skill level of the laborforce, and advances in technology drive productivity and earnings growth. For several decades, both the educational level of American workers and the capital equipment per worker have steadily increased. Technological advances have also enhanced productivity and contributed to the growth of output and income. Some people argue that technology and **automation** adversely affect workers (see the accompanying Myth of Economics feature). In fact, just the opposite is true. ***Once you recognize that higher output is the source of higher earnings, the value of new production technology is apparent: better technology makes it possible for workers to produce more and earn more.*** For example, accountants can handle more business accounts using microcomputers than they can with pencils and calculators. **A** secretary can prepare more letters with a word processor than a typewriter.

Automation
A production technique that reduces the amount of labor required to produce a good or service. It is beneficial to adopt the new labor-saving technology only if it reduces the cost of

Productivity, Wages, and the Computer Revolution

What has been happening to the growth of productivity, and how has this influenced the wages of workers? **Exhibit 5** presents data on the change in both productivity (output per hour) and real hourly compensation of the United States for the period 1948–1973, 1974–1995, and 1996–2003. Predictably, productivity growth and increases in real compensation per hour have moved together. During the period 1948–1973, both productivity and real hourly compensation grew at a rapid rate—approximately *3* percent annually.

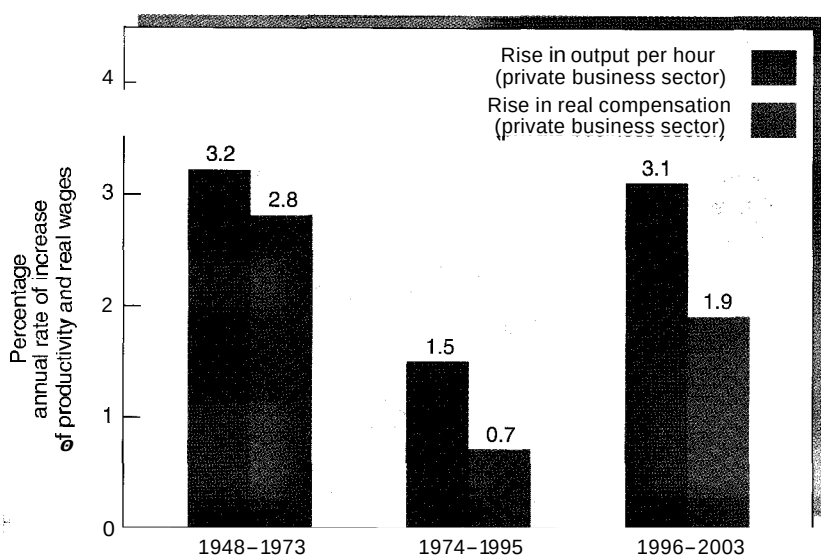


EXHIBIT 5
Productivity and Employee Compensation in the United States, 1948–2003

As shown in the graph, worker productivity and compensation per hour are closely linked. Between 1974 and 1995, the growth of both productivity and real compensation per hour slowed substantially compared to the growth figures achieved during the 1948–1973 period. However, worker productivity and real hourly compensation rebounded during the 1996–2003 period.

Source: Bureau of Labor Statistics, http://www.bls.gov/.

[6]For more on productivity growth, see Kevin J. Stiroh, "What Drives Productivity Growth?" *Economic Policy Review* (Federal Reserve Bank of New York) (March 2001) 37–59

MYTHS IN ECONOMICS

"Automation Is the Major Cause of Unemployment. If We Keep Allowing Machines to Replace People, We Are Going to Run Out of Jobs."

Machines are substituted for people if, and only if, the machines reduce costs of production. Why did the automatic elevator replace human elevator operators or the power shovel virtually eliminate human ditchdiggers? Because each is a cheaper method of accomplishing a task.

When automation and technological improvements reduce the cost of producing a good, they allow us to obtain each unit of the product with fewer resources. If the demand for the product is inelastic, consumers will spend less on the good and therefore have more of their income available for spending on other things.

Consider the following example. Suppose that someone develops a new toothpaste that really prevents cavities and sells it for half the price of other brands. If the demand for toothpaste is inelastic, at the lower price consumers will spend less on this product than they did before. Furthermore, the decline in cavities will lower the demand for and spending on dental service?. Will the lower toothpaste and dental care expenditures reduce employment? The lower level of spending on dental care will mean that households now have more income to spend on other goods and services. As a result, they will spend more on clothes, recreation, vacations, personal computers, education, and other items. This additional spending, which would not have taken place without lower dental costs, will generate additional demand and employment in these sectors, and employment will decline in the dental care sector. Although jobs will be reshuffled, there's no reason to expect that total employment will decline.

What would happen to employment if a technological improvement reduced production costs and the demand for the product were elastic? Under these circumstances, a cost-saving invention can lead to higher employment, even in the industry affected by the invention. This was essentially what happened in the automobile industry when Henry Ford's mass-production techniques reduced the cost (and price) of cars. When the price of automobiles fell 50 percent, consumers bought three times as many cars. Even though the worker-hours per car fell by 25 percent between 1920 and 1930, employment in the industry increased by approximately 50 percent during the decade.

More recently, the same thing happened in the computer-manufacturing industry. As technological improvements reduced the cost of producing various types of computer equipment, the lower costs and lower prices for computer products generated such a large increase in sales that employment in the computer-manufacturing sector actually increased.

Of course, technological advances can diminish the earnings of individual people or groups. Home appliances, like automatic washers and dryers, dishwashers, and microwave ovens, have reduced the job opportunities of maids. Voice recognition technology has lowered the demand for telephone operators. In the future, videotaped lectures may reduce the earnings and opportunities of college professors. It is understandable why groups directly affected in this manner often fear and oppose automation.

Focusing on the loss of specific jobs, however, can be misleading. Clearly, running out of jobs is not a problem. Jobs represent obstacles—tasks that must be accomplished to loosen the bonds of scarcity. As long as our ability to produce goods and services falls short of our consumption desires, there will be jobs. A society running out of jobs would be in an enviable position: It would be nearing the impossible goal—victory over scarcity!

In contrast, both sagged badly during the period 1974–1995. Productivity growth averaged only 1.5 percent annually and hourly real compensation rose at an annual rate of only 0.7 percent over the two decades (1974–1995). Although the reasons for slow productivity growth during this period are not entirely clear, most economists believe that slower improvement in the educational quality of the workforce, environmental regulations, and the inflation of the 1970s were contributing factors. Beginning in the mid-1990s, both productivity and real compensation rebounded. During the period 1996–2003, productivity and hourly compensation grew at annual rates of 3.1 percent and 1.9 percent, respectively.[7]

[7]One explanation for the slower growth in compensation than productivity is that the compensation measure does not account for the improvements in working conditions, such as job comfort and safety, that have occurred over the past few decades

What accounts for the recent rebound in the growth of productivity? Will productivity continue to grow rapidly over the next decade or so? Most economists believe that the recent acceleration in productivity growth is largely the result of the computer revolution and related technological innovations.[8] Information technology and investment in computer equipment soared during the 1990s, and productivity began to grow more rapidly. Approximately three-fifths of the productivity growth during the decade was due to this increase in investment, one study found.[9]

Furthermore, some believe that the pace of technological change has quickened and new technologies now spread throughout the economy more rapidly. As measured by price reductions, the speed of innovations in computer technology has certainly increased in recent years. For example, the price of computers in the second half of the 1990s fell at nearly twice the rate as that of the first half of the decade. A recent study estimated that about one-third of the recent productivity growth was due to a single factor: increased efficiency in the production of computers and semiconductors.[10]

Is the computer revolution going to continue, or have we already witnessed its major impact? Will the speedup of technological change continue at a rapid rate, or perhaps even accelerate, in the decades immediately ahead? It is too early to determine the answers to these questions. If the development and dissemination of technological improvements do continue at a rapid rate, higher rates of productivity growth can be expected. This will help us achieve higher income levels and living standards more quickly. However, the rapid growth of technology may also mean changing work requirements, more job switching, and less career stability. (See the accompanying feature "Are Lifetime Jobs Disappearing?")

APPLICATIONS IN ECONOMICS

Are Lifetime Jobs Disappearing?[1]

Workers in earlier generations could typically expect to have a long-term career with one firm. This is no longer the case. Today's workers believe they will have to switch employers more often than their parents did. In 1983, the average full-time male worker expected to remain with his current employer for an additional 18.6 years. By 1998, this figure had fallen to 14.7 years. A similar pattern exists for women. Their expected time of employment with their current employer dropped from 15.9 years in 1983 to 12.8 years in 1998.

Why might there be more job switching and shorter periods of employment with the same firm in the future? Two trends seem to be pushing things in that direction. First, employment is expanding in the service sector but declining in manufacturing. On average, workers in manufacturing have been with their current employer more than two years longer than workers in the service sector. Thus, movement of employment toward the service sector tends to reduce the average length of job tenure.

Second, technological advances and rapid growth of productivity can also reduce Job tenure. Workers in industries with larger productivity increases had shorter periods of job tenure during the last two decades. Technological advances tend to both replace some workers with machines and change the nature of jobs, which often means that current employees do not have the skills required to handle them. Thus, to the extent that productivity grows more rapidly in the future, it may also mean more job switching and shorter periods of job tenure with the same employer.

[1]Leora Friedberg and Michael Owyang 'Explaining the Evolution of Pension Structure and Job Tenure." Working Paper 2002-022b, Federal Reserve Bank of St Louis (October 2002).

[8]For a discussion of the new–economy perspective, see Kevin J. Stiroh, "Is There a New Economy?" *Challenge* (July/August 1999): 82–101.

[9]Stephen D. Oliner and Daniel E. Sichel, "Information Technology and Productivity: Where Are We Now and Where Are We Going?' *Economic* ***Review*** (Federal Reserve Bank of Atlanta) (Third Quarter 2002): 15–44.

[10]Ibid.

LOOKING AHEAD

As we have learned, productivity and earnings are influenced by both human capital and physical capital investments. This chapter focused on the labor market. The following chapter will analyze investment choices and the operation of the capital market.

KEY POINTS

- The real earnings of individuals would be equal if (1) all individuals were identical in preferences, skills, and background, (2) all jobs were equally attractive, and (3) workers were perfectly mobile among jobs. Earnings differences among individuals result from the absence of these conditions.
- Wage differences play an important allocative role. They generally compensate people for (1) investments in education and training that enhance productivity and the development of highly specialized skills and (2) unfavorable working conditions and/or job locations. Wage differences can also result from differences in workers' preferences, employment discrimination, and institutional factors that restrict worker mobility.
- Employment discrimination reduces the wages of people being discriminated against by either lowering the demand for their services or restricting their entry into various job categories. Productivity differences also contribute to earnings differentials between groups. Research indicates that the earnings of African-American and Mexican-American men are, respectively, approximately 86 percent and 93 percent those of white men of similar productivity.
- Productivity is the ultimate source of high wages and earnings. Workers in the United States (and other high-income industrial countries) earn high wages because their output per hour is high as the result of (1) greater worker knowledge and skills (human capital) and (2) the use of modern machinery (physical capital).
- Automated methods of production will be adopted only if they reduce costs. Although automation might reduce employment in a specific industry, it also releases resources that can be employed in other areas. Improved technology permits us to achieve larger output and income levels than would otherwise be possible.
- During the last seven years, the growth of productivity in the United States has increased well above the growth rate achieved during the prior twenty years. If the acceleration in productivity continues, increases in wages in the years ahead will be more rapid than during the 1974–1995 period.

CRITICAL ANALYSIS QUESTIONS

1. Why do some people earn higher wages than others? Why are wages in some occupations higher than in others? How do wage differentials influence the allocation of resources? Explain.

*2. Why are real wages in the United States higher than in other countries? Is the labor force itself responsible for the higher wages of American workers? Explain.

3. What explains the earnings differences between (a) a lawyer and a minister, (b) an accountant and an elementary school teacher, (c) a business executive and a social worker, (d) a country lawyer and a Wall Street lawyer, (e) an experienced, skilled craftsperson and a twenty-year-old high school dropout, and (f) an upper-story and a ground-floor window washer?

4. a. If minority employees are discriminated against, how will this affect their earnings? Use supply and demand analysis to explain your answer.

b. If the average earnings differ between two groups of employees (for example, whites and blacks), does this mean that the group with the lower earnings is experiencing employment discrimination? Why or why not?

5. Is there a relationship between the growth of productivity and changes in wage rates? Can higher earnings be achieved without higher productivity? Why or why not? Discuss.

***6.** "Jobs are the key to economic progress. Unless we create more jobs, our standard of living will fall." Is this statement true or false? Explain.

7. "If Jones has a skill that is highly valued, she will be able to achieve high market earnings. In contrast, Smith may work just as hard or even harder, and still earn only a low income."

a. Does hard work necessarily lead to a high income?
b. Why are the incomes of some workers high and others low?
c. Do you think the market system of wage determination is fair? Why or why not?
d. Can you think of a more equitable system? If so, explain why it is more equitable.

***8.** People who have invested heavily in human capital (for example, lawyers, doctors, and even college professors) generally have higher wages, but they also generally work more hours than other workers. Can you explain why?

***9.** "If individuals had identical abilities and opportunities, earnings would be equal." Is this statement true or false?

***10.** Other things being constant, how will the following influence the hourly earnings of employees? Explain your answer.

a. The employee must work the midnight to 8:00 A.M. shift.
b. The job involves split shifts (work three hours, off two hours, work three additional hours, and so on).
c. The employer provides low-cost child care services on the premises.
d. The job is widely viewed as prestigious.
e. The job requires employees to move often from city to city.
f. The job requires substantial amounts of out-of-town travel.

***11.** Consider two occupations (**A** and B) that employ people with the same skills and abilities. When employed, workers in the two occupations work the same number of hours per day. In occupation **A,** employment is stable throughout the year, while employment in B is characterized by seasonal layoffs. In which occupation will the hourly wage rate be highest? Why? In which occupation will the annual wage be highest? Why?

1 "Technological change eliminates thousands of jobs every year. Unless something is done to slow the growth of technology, ordinary workers will face a bleak future of low wages and high unemployment." Explain why you either agree or disagree with this statement.

13. If an individual is motivated primarily by the desire to make money, will he or she have an incentive to be helpful to others? Will he or she have an incentive to develop skills that others value highly? Why or why not?

*Asterisk denotes questions **for** which answers are given in Appendix B.

CHAPTER

14

Investment, the Capital Market, and the Wealth of Nations

To produce capital, people must forgo the opportunity to produce goods for current consumption. People can choose whether to spend their time picking apples or planting apple trees. In the first case there are more apples today; in the second, more apples tomorrow.

—*Steven Landsburg*[1]

Chapter Focus

- **Why do people invest? Why are capital resources often used to produce consumer goods?**
- **What is the interest rate? Why are investors willing to pay interest to get loanable funds? Why are lenders willing to loan funds?**
- **Why is the interest rate so important when costs and revenues are evaluated across time periods?**
- **When is an investment profitable? How do profitable and unprofitable investments influence the wealth of nations?**
- **How does the capital market influence growth and prosperity?**

[1]Steven E. Landsburg, *Price Theory and Applications* (Fort Worth, Tex.: Dryden Press, 1992), 581.

In the previous chapter, we noted that there is a close relationship between productivity and earnings. In turn. productivity is influenced by investment choices: consider choices about whether to construct an office building, purchase a harvecting machine, or go to law school. The returns derived from investments like these are usually spread over several years (or even decades). Some investment costs, such as maintenance expenses, can also be incurred over a length! time period. Why should we expect profit-seeking individuals and corporate decision makers to pay now to create benefits later — sometimes much later? How can people compare the benefits and costs of an activity when both are spread across lengthy periods of time? Why is the method of allocating investment a vitally important determinant of economic progress? As we explain the investment process and capital markets, this chapter will help you answer these questions.

WHY PEOPLE INVEST

Capital is a term used by economists to describe long-lasting resources that are valued because they can help us produce goods and services in the future. As we previously discussed, there are two broad categories of capital: (1) *physical capital:* nonhuman resources, like buildings, machines, tools, and natural resources, and *(2) human capital:* human resources, that is, the knowledge and skills of people. **Investment** is the purchase, construction, or development of a capital resource. Thus, investment expands the availability of capital resources.

Capital
Resources that enhance our ability to produce output in the future.

Investment
The purchase, construction, or development of capital resources, including both nonhuman capital and human capital. Investments increase the supply of capital.

Saving is income not spent on current consumption. ***Investment and saving are closely linked. In fact, the two words describe different aspects of the capital formation process. Saving refers to the nonconsumption of income, whereas investment refers to the use of the unconsumed income to produce a capital resource.*** Sometimes saving and investment are conducted by the same person, as when a farmer saves current income (refrains from spending it on consumption goods) in order to purchase a new tractor (an investment good).

Saving
Current income that is not spent on consumption goods. Savings deposited by some individuals allow resources to be devoted to investments (such as the making of tractors or other equipment used in production).

It is important to recognize that saving is required for investment. Someone must save — refrain from consumption — in order to provide the resources for investment. When investors finance a project with their own funds, they are also saving (refraining from current consumption). Investors, however, do not always use their own funds to finance investments. Sometimes they will borrow funds from others. When this is the case, it is the lender rather than the investor who is doing the saving.

Considering the alternative use of resources also highlights the link between investment and saving. Resources used to produce capital will be unavailable for the direct production of consumption goods. The opportunity cost of investing more and using more of our resources to produce capital resources today is that fewer current resources will be available to produce consumption goods.

Why would anyone want to delay consumption in order to undertake an investment? After all, consumption is the ultimate objective of all production. However, we can sometimes produce more consumption goods by first using resources to produce capital resources and then using these resources to produce the desired consumer goods. Using capital to produce consumption goods makes sense only when it allows us to produce more consumption goods than we otherwise could.

Perhaps a simple illustration can highlight the potential gains from using capital to produce consumption goods. Suppose that Robinson Crusoe can catch fish by either (1) combining his labor with natural resources (direct production) or (2) constructing a net and eventually combining his labor with this capital resource (indirect production). Let's assume that Crusoe can catch 2 fish per day by hand-fishing, but 3 fish per day if he constructs and uses a net that will last for 310 days. Now suppose it will take Crusoe 55 days to build the net. The opportunity cost of constructing the net will be 110 fish (2 fish per day not caught for each of the 55 days he spends building the net). As the accompanying chart shows, if Crusoe invests in the capital resource (the net), his output during the next year (including the 55 days required to build the net) would be 930 fish (3 per day for 310 days). Alternatively, hand-fishing during the year would lead to an output of only 730 fish (2 fish per day for 365 days).

	Number of Fish Caught	
	Without Net	With Net
Per day	2	3
Annually	730	930

In other words, Crusoe's investment in the net will enhance his productivity by 200 fish annually. In the short term, however, investing in the net will require a sacrifice. During the fifty-five days it takes to construct the net, Crusoe's production of consumption goods will decline.

How can Crusoe or any other investor know if the value of the larger future output is worth the short-term cost? Most of us have a preference for goods now rather than later. For example, if you are a typical person, you would prefer a sleek new sports car now rather than the same car ten years from now. On average, individuals possess a **positive rate of time preference**. That is, other things being the same, people subjectively value more highly goods obtained sooner than goods obtained later.

Positive rate of time preference
The desire of consumers for goods now rather than in the future.

When only Crusoe is involved, the attractiveness of the investment in the fishing net depends upon his time preference. If he places a high value on a couple of fish per day during the next fifty-five days, as indeed he may if he is on the verge of starvation, the cost of the investment may well exceed the value of the larger future output. If Crusoe could find someone who would loan him fish while he built the net, however, he could consume the borrowed fish now while building the net, and pay later with the extra fish made possible by the net. If such a loan is available, the attractiveness of the investment (building the net instead of hand-fishing now) will be influenced by the price of borrowing fish. Is the cost of borrowing fish in order to maintain his consumption while he constructs the net worth the extra cost? To answer this question, Crusoe must consider the cost of paying for earlier availability—he must consider, in effect, the interest rate.

ECONOMICS AT THE MOVIES

COLUMBIA/THE KOBAL COLLECTION

***A Knight's Tale* (2001)**

After winning a jousting competition, Heath Ledger tries to persuade two other peasants to invest their winnings in preparing for an even larger, upcoming competition to win an even bigger prize, rather than spending the winnings right away (as the two peasants want to do). The three clearly differ when it comes to their rate of time preference. This scene illustrates the trade-off between consuming now versus investing for the future.

INTEREST RATES

The interest rate links the future to the present. It allows individuals to evaluate the value today—the present value—of future income and costs. In essence, it is the market price of earlier availability. From the viewpoint of a potential borrower, the interest rate is the premium that must be paid in order to acquire goods sooner and pay for them later. From the lender's viewpoint, it is a reward for waiting—a payment for supplying others with current purchasing power. The interest rate allows the lender to calculate the future benefit (future payments earned) of extending a loan or saving funds today.

In a modern economy, people often borrow funds to finance current investments and consumption. Because of this, the interest rate is often defined as the price of loanable

funds. This definition is correct. But we should remember that it is the earlier availability of goods and services purchased, not the money itself, that is desired by the borrower.

How Interest Rates Are Determined

Interest rates are determined by the demand for and supply of loanable funds. Investors demand funds in order to finance capital assets they believe will increase output and generate profit. Simultaneously, consumers demand loanable funds because they have a positive rate of time preference: they prefer earlier availability.

The demand of investors for loanable funds stems from the productivity of capital. Investors are willing to borrow in order to finance the use of capital in production because they expect that expanding future output will provide them with more than enough resources to repay the amount borrowed—the principal—and interest on the loan. Our prior example of Robinson Crusoe illustrates this point. Remember, Crusoe could increase his output by 200 fish this year if he could take off fifty-five days from hand-fishing to build a net. But doing so would reduce Crusoe's fish production by 2 fish per day while he was constructing the net. Suppose a fishing crew from a neighboring island visited Crusoe and offered to lend him 110 fish so he could undertake the capital investment project (building the net). If Crusoe could borrow the 110 fish (the principal) in exchange for, say, 165 fish one year later (110 fish to repay the principal and 55 as interest on the loan), the investment project would be highly profitable. Crusoe could repay the funds borrowed, plus the 50 percent interest rate, and still have 145 additional fish (the 200 additional fish caught minus the 55 fish paid in interest).

Crusoe's demand for loanable fish—and, more generally, the demand of investors for loanable funds—stems directly from the productivity of the capital investment. Crusoe can gain by borrowing to finance the construction of a fishing net only because the net enables him to expand his total output during the year. Similarly, investors can gain by borrowing funds to undertake investment projects only when the capital assets they purchase permit them to expand output (or reduce costs) by enough to make the interest payments and still have more output than they would have without the investment.

As **Exhibit 1** illustrates, the interest rate brings the choices of investors and consumers wanting to borrow funds into harmony with the choices of lenders willing to supply funds. Higher interest rates make it more costly for investors to undertake capital spending projects and for consumers to buy now rather than later. Both investors and consumers will therefore curtail their borrowing as the interest rate rises. Investors will borrow less because some investment projects that would be profitable at a low interest rate

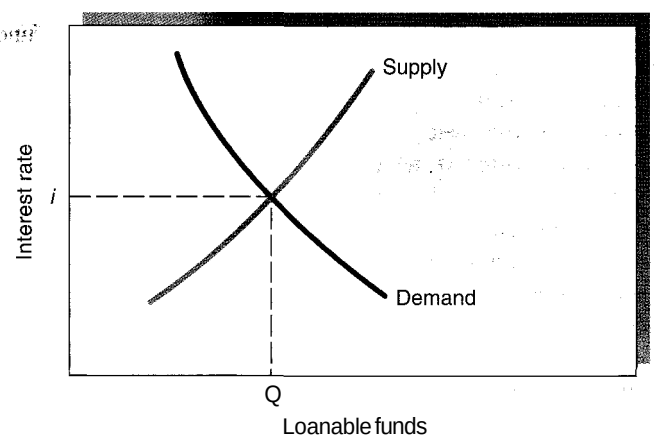


EXHIBIT 1
Determining the Interest Rate

The demand for loanable funds stems from the consumer's desire for earlier availability and the productivity of capital. As the interest rate rises, current goods become more expensive in comparison with future goods. Therefore, borrowers will demand fewer loanable funds. On the other hand, higher interest rates will stimulate lenders to supply additional funds to the market.

will be unprofitable at higher rates. Some consumers will reduce their current consumption rather than pay the high interest premium when the interest rate increases. Therefore, the amount of funds demanded by borrowers is inversely related to the interest rate.

The interest rate also rewards people (lenders) willing to reduce their current consumption in order to provide loanable funds to others. If some people are going to borrow in order to undertake an investment project (or consume more than their current income), others must curtail their current consumption by an equal amount. In essence, the interest rate provides lenders with the incentive to reduce their current consumption so that borrowers can either invest or consume beyond their current income. Higher interest rates give people who are willing to save (willing to supply loanable funds) the ability to purchase more goods in the future in exchange for sacrificing current consumption. Even though people have a positive rate of time preference, they will give up current consumption to supply funds to the loanable funds market if the price is right. Higher interest rates will induce people to save more. Therefore, as the interest rate rises, the quantity of funds supplied to the loanable funds market will increase.

As Exhibit 1 illustrates, the interest rate will bring the quantity of funds demanded into balance with the quantity supplied. ***At the equilibrium interest rate, the quantity of funds borrowers demand for investment and consumption now (rather than later) will just equal the quantity of funds lenders save. So the interest rate brings the choices of borrowers and lenders into harmony.***

The Money Rate Versus the Real Rate of Interest

We have emphasized that the interest rate is a premium paid by borrowers for earlier availability and a reward received by lenders for delaying consumption. However, during a period of inflation—a general increase in prices—the nominal interest rate, or **money rate of interest,** is a misleading indicator of how much borrowers are paying and lenders are receiving. Inflation reduces the purchasing power of a loan's principal. Rising prices mean that when the borrower repays the principal in the future, the repayment amount will not purchase as much as it would have when the funds were initially loaned.

Money rate of interest
The rate of interest in monetary terms that borrowers pay for borrowed funds. During periods when borrowers and lenders expect inflation, the money rate of interest exceeds the real rate of interest.

When inflation is common, lenders will recognize that they are being repaid with dollars of less purchasing power. Unless they are compensated for the anticipated inflation by an upward adjustment in the interest rate, they will supply fewer funds to the loanable funds market. At the same time, when borrowers anticipate inflation, they will want to purchase goods and services now before they become even more expensive in the future. Thus, they are willing to pay an **inflationary premium**, an additional amount of interest that reflects the expected rate of future price increases. For example, if borrowers and lenders fully anticipate a 5 percent rate of inflation, they will be just as willing to agree on a 9 percent interest rate as they were earlier to agree on a 4 percent interest rate when both anticipated price stability.

Inflationary premium
A component of the money interest rate that reflects compensation to the lender for the expected decrease, due to inflation, in the purchasing power of the principal and interest during the course of the loan. It is determined by the expected rate of future inflation.

Unlike when the general price level is stable, the supply of loanable funds will decline (the supply curve will shift to the left) and the demand will increase (the demand curve will shift to the right) once decision makers anticipate future inflation. The money interest rate thus rises, overstating the "true" cost of borrowing and the yield from lending. This true cost is the **real rate of interest,** which is equal to the money rate of interest minus the inflationary premium. It reflects the real burden to borrowers and payoff to lenders in terms of their being able to buy goods and services.

Real rate of interest
The money rate of interest minus the expected rate of inflation. The real rate of interest indicates the interest premium, in terms of real goods and services, that one must pay for earlier availability.

Our analysis indicates that high rates of inflation will raise the money rate of interest. The real world is consistent with this view. Money interest rates rose to historical highs in the United States as inflation soared to double-digit rates during the 1970s. These same nominal rates fell to the 5 percent range as inflation fell below 2 percent in the 1990s. Cross-country comparisons also illustrate the link between inflation and high interest rates. The lowest money interest rates in the 1990s and early 2000s were found in nations such as Germany, Switzerland, and Japan, all with low rates of inflation. In contrast, the highest money interest rates were observed in Russia, Brazil, Turkey, and other countries with high rates of inflation during the period.

Interest Rates and Risk

So far, we've assumed that there is only a single interest rate present in the loanable funds market. In the real world, of course, there are many interest rates. There is the mortgage rate, the prime interest rate (the rate charged to business firms with strong credit ratings), the consumer loan rate, and the credit card rate, to name only a few.

Interest rates in the loanable funds market will differ mainly because of differences in the risks associated with the loans. It is riskier, for example, to loan money to an unemployed worker than to a well-established business with substantial assets. Similarly, credit card loans are riskier than loans secured by an asset. An example of a secured loan is a mortgage loan on a house. If the borrower defaults, the lender can repossess the house. The risk also increases with the duration of the loan. The longer the time period, the more likely that the ability of the borrower to repay the loan or market conditions will change—perhaps unfavorably.

As **Exhibit 2** shows, the money rate of interest on a loan has three components. The pure-interest component is the real price one must pay for earlier availability. The inflationary-premium component reflects the expectation that the loan will be repaid with dollars of less purchasing power as the result of inflation. The risk-premium component reflects the probability of default—the risk imposed on the lender by the possibility that the borrower may be unable to repay the loan.

THE PRESENT VALUE OF FUTURE INCOME AND COSTS

If you deposited $100 today in a savings account earning 6 percent interest, you would have $106 one year from now. To put it another way, the present value of $106 one year from now is equal to the amount ($100) you would have to invest today in order to get $106 a year later. ***The interest rate allows us to make this calculation. The interest rate connects the value of dollars (and capital assets) today with the value of dollars (expected income and receipts) in the future. It is used to "discount" the value of a dollar received in the future so that its present worth can be determined today.***

The **present value *(PV)*** of a payment received one year from now can be expressed as follows:

Present value (*PV*)
The current worth of future income after it is discounted to reflect the fact that revenues in the future are valued less highly than revenues now.

$$PV = \frac{\text{Receipts One Year from Now}}{1 + \text{Interest Rate}}$$

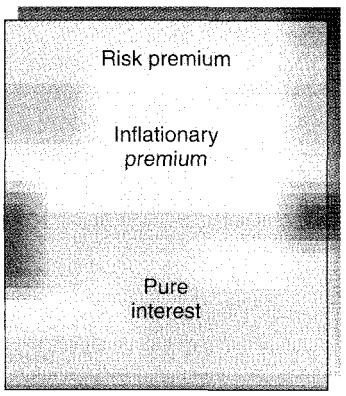


EXHIBIT 2
The Three Components of Money Interest

The money interest rate reflects three components: pure interest, inflationary premium, and risk premium. When decision makers expect a high rate of inflation while the loan is outstanding, the inflationary premium will be substantial. Similarly, the risk premium will be large when the probability of default by the borrower is substantial.

If the interest rate is 6 percent, the current value of $100 to be received one year from now is

$$PV = \frac{100}{1.06} = 94.34$$

If you put $94.34 in a savings account yielding 6 percent interest, during the year the account would earn $5.66 interest (6 percent of $94.34) and grow to $100 one year from now. Thus, the present value of $100 to be received a year from now is $94.34.

Discounting
The procedure used to calculate the present value of future income, which is inversely related to both the interest rate and the amount of time that passes before the funds are received.

Economists use the term **discounting** to describe this procedure of reducing the value of a dollar to be received in the future to its present worth. Clearly, the value of a dollar in the future is inversely related to the interest rate. For example, if the interest rate is 10 percent, the present value of $100 received one year from now would be only $90.91 ($100 divided by 1.10). At still higher interest rates, the present value of $100 a year from now would be even lower.

The present value of $100 received two years from now is

$$PV = \frac{100}{(1 + \textit{Interest Rate})^2}$$

If the interest rate is 6 percent, $100 received two years from now would be equal to $89 today ($100 divided by 1.06^2). In other words, $89 invested today would yield $100 two years from now.

The present-value procedure can be used to determine the current value of any future income (or cost) stream. If R represents receipts received at the end of various years in the future (indicated by the subscripts) and i represents the interest rate, the present value of the future income stream is

$$PV = \frac{R}{(1 + i)} + \frac{R_2}{(1 + i)^2} + \cdots + \frac{R_n}{(1 + i)^n}$$

Exhibit 3 shows the present value of $100 received at various times in the future at several different discount rates. The chart clearly illustrates two points. First, the present value of income received in the future declines with the interest rate. The present value of the $100 received a year from now, when discounted at a 4 percent interest rate, is $96.15, compared to $98.04 when a 2 percent discount rate is used. Second, the present value of the $100 also declines as the date of its receipt is set further into the future. If the discount rate is 6 per-

EXHIBIT 3
The Present Value of $100 to be Received in the Future

The columns indicate the present value of $100 to be received a designated number of years in the future at different discount (interest) rates. For example, at a discount rate of 2 percent, the present value of $100 to be received five years from now is $90.57. Note that the present value of the $100 declines as either the interest rate or the number of years in the future increases.

	PRESENT VALUE OF $100 TO BE RECEIVED A DESIGNATED NUMBER OF YEARS IN THE FUTURE AT ALTERNATIVE DISCOUNT RATES					
YEARS IN THE FUTURE	2 PERCENT	4 PERCENT	6 PERCENT	8 PERCENT	12 PERCENT	20 PERCENT
1	98.04	96.15	94.34	92.59	89.29	83.33
2	96.12	92.46	89.00	85.73	79.72	69.44
3	94.23	88.90	83.96	79.38	71.18	57.87
4	92.39	85.48	79.21	73.50	63.55	48.23
5	90.57	82.19	74.73	68.06	56.74	40.19
6	88.80	79.03	70.50	63.02	50.66	33.49
7	87.06	75.99	66.51	58.35	45.23	27.08
8	85.35	73.07	62.74	54.03	40.39	23.26
9	83.68	70.26	59.19	50.02	36.06	19.38
10	82.03	67.56	55.84	46.32	32.20	16.15
15	74.30	55.53	41.73	31.52	18.27	6.49
20	67.30	45.64	31.18	21.45	10.37	2.61
30	55.21	30.83	17.41	9.94	3.34	0.42
50	37.15	14.07	5.43	2.13	0.35	0.01

cent, the present value of $100 received a year from now is $94.34, compared to $89 if the $100 is received two years from now. If the $100 is received five years from now, its current worth is only $14.73. ***So the present value of a future dollar payment is inversely related to both the interest rate and how far in the future the payment will be received.***

To see in a more personal way the importance of interest and the value of saving, consider what you could gain by saving just $1,000 per year ($83.33 per month) for ten years in a tax-free individual retirement account. If you begin at age twenty-five and continue only until age thirty-five (putting nothing into the account after that), and the account returns 8 percent annually, it will be worth $168,627 when you reach age sixty-five. In contrast, if you wait until age thirty-five and then save $1,000 per year for thirty years (not ten years of savings as before) with the same 8 percent annual return, by age sixty-five your account will be worth only $125,228. Ten years of saving, starting at age twenty-five, yields far more than thirty years of saving at the same amount each year but waiting to start until age thirty-five. Which savings plan is more attractive to you?

PRESENT VALUE. PROFITABILITY. AND INVESTMENT

Investment decisions, like saving decisions, require us to compare costs and benefits over time. Investment involves an up-front cost of acquiring a machine, skill, or other asset expected to generate additional output and revenue in the future. How can an investor know if the expected future revenues will be sufficient to cover the costs? The discounting procedure helps provide the answer. It permits the investor to place both the costs and the expected future revenues of an investment project in present-value terms. If the present value of the revenue derived from the investment exceeds the present value of the cost, it makes sense to undertake the investment. If the revenues and costs of such an investment turn out as expected, the investor will earn a profit, and the project will create wealth.

On the other hand, if the cost of the project exceeds the discounted value of the future receipts, the project is a loser. The losses reflect that the project diminished the value of the resources. Investments like these are counterproductive.

Let's pencil in the figures for a hypothetical investment option and determine whether it is a good investment. Suppose that a truck rental firm is contemplating the purchase of a new $40,000 truck. Past experience indicates that after figuring in the truck's operational and maintenance expenses, the firm can rent it out for a net revenue of $12,000 annually (received at the end of each year). The firm will be able to do this every year for four years—the expected life of the vehicle.[2] Since the firm can borrow and lend funds at an interest rate of 8 percent, we will discount the future expected income at an 8 percent rate. **Exhibit 4** illustrates the calculation. Column 4 shows how much $12,000, available at year-end for each of the next four years, is worth today. In total, the present value of the expected rental receipts is $39,744—less than the purchase price of the truck. Therefore, the project should not be undertaken.

EXHIBIT 4
The Discounted Present Value of $12,000 of Truck Rental Receipts for Four Years (Interest Rate = 8 Percent)

YEAR (1)	EXPECTED FUTURE INCOME RECEIVED AT YEAR-END (2)	DISCOUNTED VALUE PER DOLLAR (8 PERCENT RATE) (3)	PRESENT VALUE OF INCOME (4)
1	$12,000	$0.926	$11,112
2	12,000	0.857	10,284
3	12,000	0.794	9,528
4	12,000	0.735	8,820
Total			$39,744

[2]For the *sake* of simplicity, we assume that the **truck has** no scrap value at the end of four years

APPLICATIONS IN ECONOMICS

Would You Like to Become a Millionaire?

In the previous chapter, we indicated that approximately one out of every twenty American households has assets worth a million dollars or more. The proportion of millionaires among those fifty-five to sixty-four years of age is even higher. Many millionaires earned large incomes in business, professional, and other activities, but some achieved this status even though their incomes were not very much different from those of other Americans. They did so by saving and investing regularly throughout their lives.

How could a person become a millionaire without making a huge income? Saving and compound interest provide the answer. The real rate of return on a broad portfolio of stock has averaged more than 7 percent during the last 200 years. Moreover, stock investments over 35- to 40-year periods have consistently earned a real return in this range. A real return of 7 percent is highly realistic for even a novice investor. Those with more entrepreneurial talent might well do better. If a 25-year-old began putting money into a tax-free retirement account, pension fund, or similar investment earning a 7 percent real return, how much would the person have to save annually in order for the funds to be worth a million dollars (measured in the purchasing power of today's dollar) when he or she reaches age 65? The answer is $5,009. Put another way, a person (or household) with an income of $50,000 (which is only slightly above the U.S. average) could achieve millionaire status by saving just 10 percent of it annually. A 5 percent savings rate would do the job with $100,000 of income annually. The power of compound interest illustrates that millionaire status is well within the reach of many Americans willing to save and invest for many years.

If the interest rate in our example had been 6 percent, the results would have been different. The present value of the future rental income would have been $41,580.[3] It pays to purchase a capital good whenever the present value of the income generated by it exceeds its purchase price. At the lower interest rate, the investment project would have been productive (and profitable).

Expected Future Earnings and Asset Values

The present value of the expected revenue from a project minus the cost of an investment will tell us whether the project should be undertaken. However, *once an investment project has been completed,* the present value of the expected future net earnings will determine the market value of the asset. If the present value of the expected net earnings rises (or falls), so too will the value of the asset.

The value of an asset is equal to the present value of the expected net revenues that can be earned by the asset. If the asset is expected to generate a constant annual net income each year in the future, its value would be equal to:

$$\textit{Asset Value} = \frac{\textit{Annual Net Income from the Asset}}{\textit{Interest Rate}}$$

What is the market value of a tract of land that is expected to generate $1,000 of rental income net of costs each year indefinitely into the future? If the market interest rate is 10 percent, investors would be willing to pay $10,000 for the land ($1,000 divided by .10).

[3]The denvation of **this** figure is shown in the following tabulation

Year	Expected Future Income (dollars)	Discounted Value per Dollar (6% rate)	Present Value of Income (dollars)
1	12,000	0.943	11,316
2	12,000	0.890	10,680
3	12,000	0.840	10,080
4	12.000	0.792	9,504
Total			41,580

When purchased at this price, the land would provide an investor with the 10 percent market rate of return. Similarly, if an asset generates $2,500 of net earnings annually and the market interest rate is 10 percent, the asset would be worth $25,000 ($2,500 divided by .10). There is a direct relationship between the expected future earnings generated by an asset and the asset's market value. ***As the present value of the future earnings of an asset increases, so, too, does the market value of the asset.***

This link between expected future earnings and the price of an asset motivates asset owners to make sure that the assets are being used wisely. Some entrepreneurial investors are particularly good at (1) identifying a business that is poorly operated, (2) purchasing the business at a depressed price, (3) improving the operational efficiency of the firm, and then (4) reselling the business at a handsome profit. Suppose that a poorly run business currently has net earnings of $1 million per year. What is the market value of the business? If the firm is expected to continue earning $1 million per year, the market value of the firm would be $10 million if the interest rate is 10 percent. Suppose that an alert entrepreneur buys the business for $10 million and improves the operational efficiency of the firm. As the result of these changes, the annual net earnings of the firm increase to $2 million per year. Now how much is the firm worth? If the $2 million annual earnings are expected to continue into the future, the net present value of the firm would rise to $20 million. Thus, the entrepreneur would be able to sell the firm for a very substantial profit. You can see from this example that business managers and asset owners have a strong incentive to use the resources under their control efficiently. If they don't, the value of the assets will decline, and the business will be vulnerable to a takeover by those who do.

Investors and Corporate Investments

In modern market economies, investors typically are not entrepreneurs who personally decide which factories to expand, which machine tools to build, and which research investments to undertake. Instead corporate officers, under the scrutiny of corporate boards of directors, make the entrepreneurial capital investment choices. Nevertheless, individual investors (buyers and sellers of stock) influence that process through the stock market. Stock market investors who believe that a corporation is currently making decisions likely to increase its future profits will buy more of the corporation's stock, driving up its price. Similarly, stockholders who believe that the corporation's current investment decisions will reduce future profits have an incentive to "bail out" by selling their stock holdings, reducing the market value of the stock. Either way, the price of a corporation's shares gives corporate officers fast feedback on what investors think of their decisions.

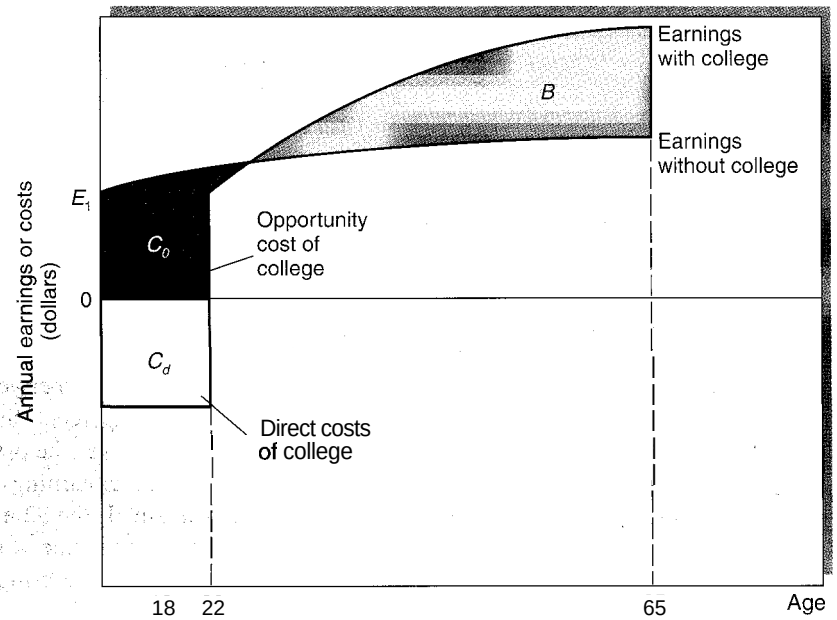
Do corporate officers respond to stock price changes? Normally they do. Often they own stock themselves, and the value of their pay package generally depends on the stock price. Also, the members of the corporate board (which hires and fires the officers) are typically large shareholders. Thus, the corporate officers have a strong incentive to act on the feedback from the stock market. The choices of individual buyers (and nonbuyers) of the firm's products are the ultimate judge of its performance. However, the choices of investors and their fund managers provide early information about the expected success of business ventures.

INVESTING IN HUMAN CAPITAL

Investments in human capital (like the decision to go to college, for example) are the same as other investment decisions. As with physical capital, discounting can also help us gauge whether or not to make a human capital investment.

Exhibit 5 shows the human capital decision confronting Juanita, an eighteen-year-old high school graduate thinking about pursuing a bachelor's degree in business administration. Just as an investment in a truck involves a cost in order to generate a future income, so does a degree in business administration. If Juanita does not go to college, she will be

EXHIBIT 5
Investing in Human Capital



Juanita, an eighteen-year-old who has just finished high school, is trying to decide whether or not to go to college and get a business degree. If Juanita goes to college and majors in business administration, she will incur the direct cost (C_d) of the college education—tuition, books, transportation, and so on—plus the opportunity cost (C_0) of earnings she forgoes while in college. However, with a business degree, she can expect to earn an additional amount of income (B) during her career. If the discounted present value of the additional future earnings exceeds the discounted value of the direct and indirect costs of a college education, the business degree will be a profitable investment for Juanita.

able to begin work immediately, starting at annual earnings of E_1. But if she goes to college, she will incur direct costs (C_d) in the form of tuition, books, transportation, and related expenses. She will also bear the opportunity cost (C_0) of lower earnings while she's in college. However, studying business will expand Juanita's knowledge and skills, and enable her to earn more money later (see the areas in the exhibit labeled "Earnings with college" versus "Earnings without college"). Will the higher future income be worth the cost? To answer this question, Juanita must project the additional income she would earn by getting a business degree, discount each year's additional income, and compare the total with the discounted value of the cost of getting the degree—including the opportunity cost of the earnings she forgoes while attending college. If the discounted present value of Juanita's additional future income exceeds the discounted present value of the cost, acquiring the degree is a worthwhile investment.

Of course, nonmonetary considerations might also be important to Juanita. For example, Juanita might prefer having a degree and working in the business world (rather than in the jobs available to high school graduates) even if she did not make more money. Thus, the nonmonetary attractiveness of business may induce her to pursue the business degree, even if the monetary rate of return is low (or even negative).

Like choosing to purchase a new machine, choosing a human capital investment project such as Juanita's degree involves cost, the possibility of profit, and uncertainty. The same principles apply to both types of decisions. Giving due consideration to nonmonetary factors as a potential "benefit," human capital investors, like physical capital investors, will undertake only those projects that they think will yield benefits greater than their costs.

UNCERTAINTY, ENTREPRENEURSHIP, AND PROFIT

As we previously discussed, firms and individual investors can sometimes earn persistent economic profits—returns in excess of the opportunity cost of funds—if they can restrict

entry into various industries and occupations. Economic profits, however, can also be earned in competitive markets with no entry restrictions. ***In competitive markets, there are two sources of economic profit: uncertainty and entrepreneurship.***

First, let's talk about how uncertainty enables people to earn an economic profit. Investment opportunities put people in a position to earn a handsome return, but they also expose them to additional uncertainty. We live in a world of uncertainty, imperfect information, and dynamic change. Unanticipated changes, changes that no one could have foreseen, create winners and losers. If people didn't care about the uncertainty related to different investment projects, they wouldn't demand higher returns for higher-risk projects. Most people, though, dislike uncertainty. They would prefer to know that they'll collect $1,000 rather than take a 50-50 chance of collecting either nothing or $2,000. To accept the uncertainty associated with riskier investments, they will demand a premium—an economic profit. For example, compared to government bonds, the returns on stocks are considerably more uncertain. This uncertainty of return is one reason that stocks have historically yielded a higher average return than bonds.

Entrepreneurship is a second source of economic profit. Some people are better than others at identifying entrepreneurial opportunities. At any given time, there are infinite numbers of potential investment projects to choose from in the economy. Investing in some of them will increase the value of resources and lead to handsome returns. Investing in others will actually reduce the value of resources, generating losses. Entrepreneurship involves the ability to recognize and undertake economically beneficial projects that have gone unnoticed by others. Creativity, foresight, and leadership are important aspects of entrepreneurship. Naturally, risk is involved. Unfortunately, all too frequently, the entrepreneur's vision turns out to have been a mirage. What appeared to be a profitable opportunity turns out to be an expensive mistake.

The great Harvard economist Joseph Schumpeter believed that entrepreneurship and innovation were the moving forces behind capitalism. According to Schumpeter, the entrepreneurial discovery of new, improved ways of doing things drove economic progress and improved living standards. As Schumpeter put it,

> The fundamental impulse that sets the capitalist engine in motion comes from the new consumer's goods, the new methods of production or transportation, and new markets, and the new forms of industrial organization that capitalist enterprise creates.[4]

OUTSTANDING ECONOMIST

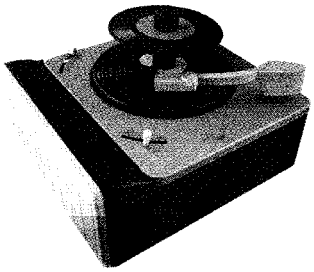
Joseph Schumpeter (1883–1950)

Born in Austria, Joseph Schumpeter was a longtime professor of economics at Harvard University. Generally recognized as one of the top five economists of this century, Schumpeter is perhaps best known for his views on entrepreneurship and the future of capitalism. He believed that the discovery and introduction of new and better products and organizational methods were constantly making the old ways of doing things obsolete. He referred to this process as "creative destruction" and argued that it was the primary fuel of economic

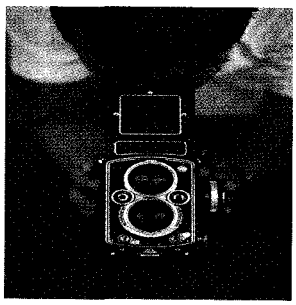
[4]Joseph A. Schumpeter, *Capitalism, Socialism, and Democracy* (New York: Harper Torchbooks, 1950), 83

The introduction of new products often makes older ones obsolete. Examples abound: the personal computer replaced the typewriter, digital music players replaced the phonograph, and digital cameras are replacing film models. Joseph Schumpeter referred to outcomes of this type as "creative destruction." How does creative destruction influence the living standards of people like you?

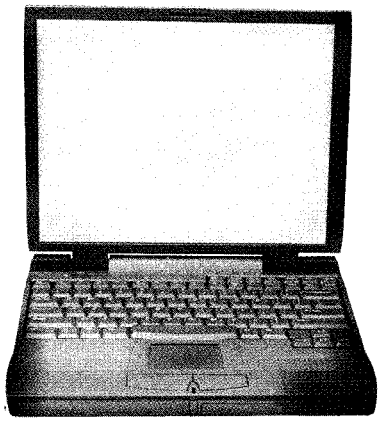
GETTY IMAGES



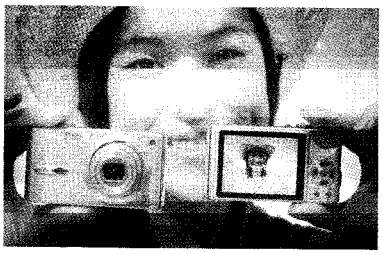
©ROYALTY-FREE/CORBIS



GETTY IMAGES



GETTY IMAGES



To finance their investments, would-be entrepreneurs usually have to use their own money and the money of co-venturers, in addition to whatever they can borrow. Entrepreneurs with successful track records find it easier to get capital for their investment projects. As a result, in a market economy, successful entrepreneurs will have a greater say about which projects will be undertaken and which will not.

Returns to Physical and Human Capital

Both interest and profit perform important allocative functions. Interest induces people to give up current consumption, a sacrifice that is necessary for capital formation. Economic profit provides both human and physical capital decision makers with the incentive to (1) undertake investments yielding an uncertain return and (2) discover and undertake beneficial (wealth-creating) investment opportunities.

Income reflects the returns from both human capital and physical capital. As **Exhibit 6** shows, approximately four-fifths of the national income in the United States is earned by employees and self-employed workers. The earnings in these two categories primarily reflect a return on human capital. The other one-fifth—income in the form of interest, corporate profits, and rents—reflects mostly returns to physical capital.[5] These shares of human and physical capital have been relatively constant for several decades.

[5]As used here, rent is income derived by owners who lease (rent) assets like buildings and equipment to another party for a penod of time

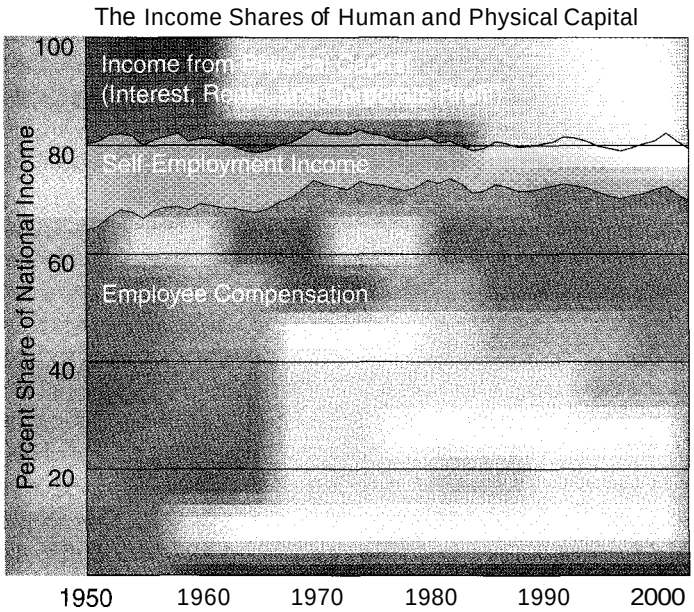


EXHIBIT 6
The Income Shares of Human and Physical Capital

The share of national income earned by employees, self-employed proprietors, and owners of physical capital (interest, corporate profits, and rents) is shown here. Employee compensation and self-employment income represent primarily returns to human capital. These two components have constituted approximately 80 percent of total national income in the United States for several decades.

Source: http://www.bea.doc.gov/.

THE CAPITAL MARKET AND THE WEALTH OF NATIONS

If the potential gains from innovative ideas and human ingenuity are going to be fully realized, it must be relatively easy for individuals to try their innovative and potentially ingenious ideas, but difficult to continue if the idea is a bad one.

Innovation and the Capital Market

Improved knowledge about how to transform resources into desired goods and services is the main reason that our output today is far greater than our ancestors' a century or two ago. Predictably, further improvements in our knowledge and and the innovative applications of those improvements will propel growth and prosperity in the future, as well. The human ingenuity and entrepreneurial talent capable of pushing the frontier of economic progress outward are found everywhere and often come from unexpected sources. For example, college dropout Ted Turner, the founder of CNN, transformed the way people around the world received news during the last two decades of the twentieth century. A relatively unsuccessful marketer of milk shake machines (Ray Kroc) perfected the franchising business and developed McDonald's into the world's largest restaurant chain. A plainspoken owner of a small retail store in one of the poorest states (Sam Walton, founder in Arkansas of Wal-Mart) became America's largest retailer.

It is impossible to determine who will come up with the next innovative idea that will promote human progress. This makes it vitally important that individuals from all backgrounds and walks of life be able to try out their ideas at a relatively low cost. Of course, counterproductive projects can waste substantial amounts of resources. Thus, if a nation is going to get the most out of its resources, it must also have a mechanism that will bring unsound projects to a halt.

Competitive capital markets perform these two functions. In a market economy, it is relatively easy to try out a new idea. Entrepreneurs do not have to convince a majority of the populace or key political leaders that their idea is sound. They can undertake the project if they can get the required financing. This can be achieved by either putting up one's own money or convincing financiers of the project's profit potential merits. Once the project is undertaken, in order to continue, it will have to pass the market test. If the revenues derived from sale of the output produced are insufficient to cover the cost, it will be difficult to keep

the project afloat. Losses will soon drive unprofitable business ventures from the market. Conversely, if a project is profitable, this will encourage operation on a larger scale.

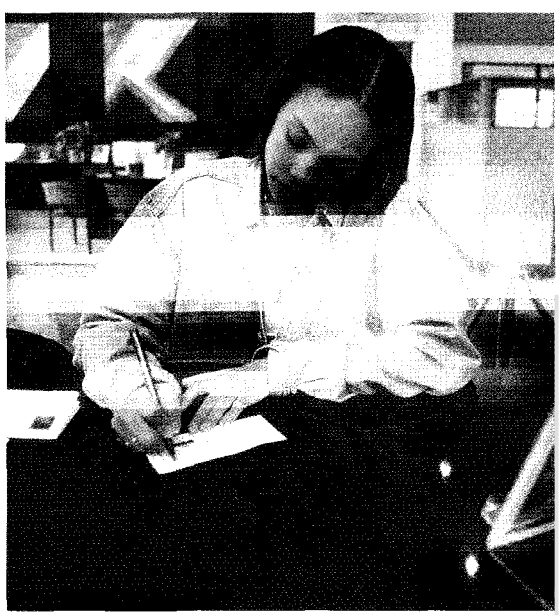
The modern capital market also accommodates lenders with widely varying preferences and goals. For example, a person wanting to save for retirement that will begin only ten years from now may nonetheless be willing to buy a new twenty-year bond from General Motors or some other private firm because the bond market makes it relatively easy to sell the bond to another person if the bondholder desires to do so. Varying preferences for risk are also accommodated. Some people want to supply savings in exchange for a fixed rate of return. People who purchase bonds and maintain savings deposits are examples. Others are willing to supply funds in exchange for an uncertain return linked to the success or failure of a business or investment project. Stockholders and partnership investors fall into this category. Still others supply funds to the capital market when they use their own funds to purchase a business or acquire additional schooling.

Private investors, ranging from small-business owners to corporate stockholders to venture capitalists, place their own funds at risk in the capital market. This provides them with a strong incentive to search for, discover, and undertake profitable projects — those that generate revenues that exceed costs. In turn, profit indicates that people value the output more than the resources utilized. Thus, profitable investments tend to increase not only the wealth of the investor, but also the wealth of the nation.

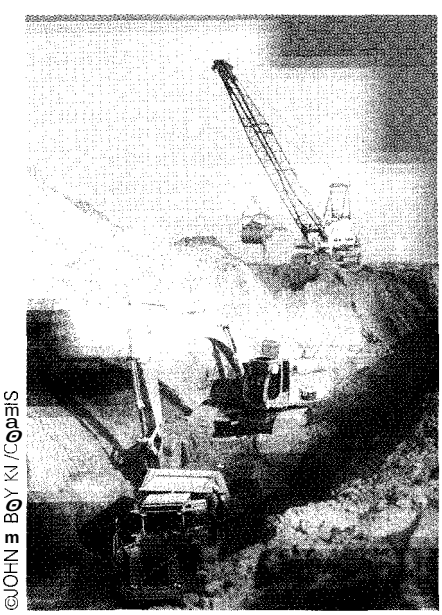
In an uncertain world, of course, mistakes will occur. Sometimes projects will be undertaken that prove unprofitable. If investors were unwilling to take chances, many innovative ideas and worthwhile but risky projects would never be tested. In a world of uncertainty, mistaken investments are a necessary price that must be paid for fruitful innovations to occur. The capital market will at least ensure that the mistakes are self-correcting. Unexpected losses will signal investors to terminate unprofitable and unproductive projects, and investors whose own wealth is at stake will be very sensitive and responsive to these signals.

In summary, the capital market will tend to direct savings toward entrepreneurial activities that are profitable. When property rights are clearly defined and enforced, profitable projects will also be productive — they will create wealth. Properly guided investment in both physical and human capital is an important source of productivity and income growth. Other things being constant, countries that invest more and channel more of their investments into productive projects today will tend to have a higher income tomorrow.

Without a private capital market, it is hard to see how investment funds could be consistently channeled into wealth-creating projects. In today's dynamic world, economic



©BILL ARON/PHOTO EDIT



©JOHN m BoY KI /CoaeIS

The capital market will channel the savings of individuals into the finance of investment goods like trucks and earth-moving equipment. Does it make any difference whether these investment projects are profitable? Why or why not?

growth is driven by discovery and innovation. Given the pace of dynamic change and the diversity of entrepreneurial talent that exists in the world, the knowledge required for making sound decisions about how capital should be allocated is far beyond the scope of any single gifted leader, industrial planning committee, or government agency. Perhaps more important, when investment funds are allocated by the government rather than by the market, an entirely different set of criteria comes into play. The individual voter, unlike a partner or an owner of stock, has little incentive to be alert for new opportunities or wary of investment mistakes made by government officials. Furthermore, the citizen-voter can neither (1) invest more upon discovery of an attractive project nor *(2)* opt out (sell) when problems are detected. Consequently, voters in economies like these are in a weak position to monitor and direct government investment effectively. Most simply remain on the sidelines. As they do so, the political clout of special interests replaces market returns as the basis for allocating funds. Predictably, investment projects that reduce wealth rather than enhance it become more likely.

The experience of Eastern Europe and the former Soviet Union provides a vivid illustration of what happens when political factors replace economic considerations in the allocation of investment. At one time, the investment rates of these countries were among the highest in the world. Their central planners, in fact, channeled approximately one-third of the national output into capital formation. Living standards were kept low to finance these investments. But even these high rates of investment did little to improve living standards over time. Government planners and state enterprises were slow to incorporate scientific breakthroughs. Available consumer goods generally reflected outdated technologies. Funds were often wasted on political boondoggles and high-visibility projects favored by important political leaders. Misdirection of investment and failure to keep up with dynamic change eventually led to the collapse of the system.

LOOKING AHEAD

As the preceding two chapters have stressed, investment in physical and human capital will influence the wealth and income of both individuals and nations. Differences among individuals with regard to these factors will also contribute to income inequality. In the next chapter; we will consider the issue of income inequality in some detail.

KEY POINTS

- We can often produce more consumption goods by first using our resources to produce physical and human capital resources and then using these capital resources to produce the desired consumption goods. Because resources used to produce capital goods will be unavailable for the direct production of consumption goods, saving is necessary for investment.
- The interest rate is the price of earlier availability. It is the premium that borrowers must pay to lenders to acquire goods now rather than later.
- The demand for loanable funds reflects the productivity of capital resources and the positive rate of time preference. The market interest rate will bring the quantity of funds demanded by borrowers into balance with the quantity supplied by lenders.
- During inflationary times, the money rate of interest incorporates an inflationary premium reflecting the expected future increase in the price level. When inflation is expected, the money rate of interest exceeds the real rate of interest.
- The money rate of interest on a specific loan reflects three basic factors—the pure interest rate, an inflationary premium, and a risk premium that is directly related to the probability of default by the borrower.

- The interest rate allows individuals to put a current value on future revenues and costs. The discounting procedure can be used to calculate the present value of an expected net income stream from a potential investment project. If the present value of the expected revenues exceeds the present value of the expected costs—and if things turn out as expected—the project will be profitable.
- The present value of expected future net earnings will determine the market value of existing assets. An increase in the expected future earnings of an asset will increase the market value of the asset. Conversely, a decrease in the expected future earnings of an asset will reduce the market value of the asset.
- Economic profit plays a central role in allocating capital and determining which investment projects will be undertaken. In a competitive environment, economic profit reflects uncertainty and entrepreneurship—the ability to recognize and undertake profitable projects that have gone unnoticed by others.
- To grow and prosper, a nation must have a mechanism that will attract savings and channel them into investments that create wealth. The capital market performs this function in a market economy. When property rights are clearly defined and enforced, profits and losses will channel investment into projects that promote economic progress.

CRITICAL ANALYSIS QUESTIONS

*1. How would the following changes influence the rate of interest in the United States?
   a. An increase in the positive time preference of lenders
   b. An increase in the positive time preference of borrowers
   c. An increase in domestic inflation
   d. Increased uncertainty about a nuclear war
   e. Improved investment opportunities in Europe

2. "Any return to capital above the pure interest yield is unnecessary. The pure interest yield is sufficient to provide capitalists with the earnings necessary to replace their assets and to compensate for their sacrifice of current consumption. Any return above that is pure gravy; it is excess profit." Do you agree with this view? Why or why not?

3. How are human and physical capital investment decisions similar? How do they differ? What determines the profitability of a physical capital investment? Do human capital investors make profits? If so, what is the source of the profit? Explain.

*4. A lender made the following statement to a borrower: "You are borrowing $1,000, which is to be repaid in twelve monthly installments of $100 each. Your total interest charge is $200, which means your interest rate is 20 percent." Is the effective interest rate on the loan really 20 percent? Explain.

5. In a market economy, investors have a strong incentive to undertake profitable investments. What makes an investment profitable? Do profitable investments create wealth? Why or why not? Do all investments create wealth? Discuss.

*6. Over long periods of time, the rate of return of an average investment in the stock market has exceeded the return on high-quality bonds. Is the higher return on stocks surprising? Why or why not?

7. The interest rates charged on outstanding credit card balances are generally higher than the interest rate that banks charge customers with a good credit rating. Why do you think the credit card rate is so high? Should the government impose an interest rate ceiling of, say, 10 percent? If it did, who would be hurt and who would be helped? Discuss.

*8. If the money rate of interest on a low-risk government bond is 10 percent and the inflation rate for the last several years has been steady at 4 percent, what is the estimated real rate of interest?

9. Suppose that you are contemplating the purchase of a commercial lawn mower at a cost of $10,000. The expected lifetime of the machine is three years. You can lease the asset to a local business for $4,000 annually (payable at the end of each year) for three years. The lessee is responsible for the upkeep and maintenance of the machine during the three-year period. If you can borrow (and lend) money at an interest rate of 8 percent, will the investment be a profitable undertaking? Is the project profitable at an interest rate of 12 percent? Provide calculations in support of your answer.

*10. Alicia's philosophy of life is summed up by the proverb, "A penny saved is a penny earned." She plans and saves for the future. In contrast, Mike's view is, "Life is uncertain; eat dessert first." Mike wants as much as possible now.
   a. Who has the higher rate of time preference?

b. Do people like Alicia benefit from the presence of people like Mike? Explain.

c. Do people like Mike benefit from the presence of people like Alicia? Explain.

***11.** Some countries with very low incomes per capita are unable to save very much. Are people in these countries helped or hurt by people in high-income countries with much higher rates of saving?

***12.** According to a news item, the owner of a lottery ticket paying $3 million over twenty years is offering to sell the ticket for $1.2 million cash now. "Who knows?" the ticket owner explained. "We might not even be here in twenty years, and I do not want to leave it to the dinosaurs."

a. If the ticket pays $150,000 per year at the end of each year for the next twenty years, what is the present value of the ticket when the appropriate rate for discounting the future income is thought to be 10 percent?

b. If the discount rate is in the 10 percent range, is the sale price of $1.2 million reasonable?

c. Can you think of any disadvantages of buying the lottery earnings rather than a bond?

13. Suppose that you are moving into a new apartment you expect to rent for five years. The owner of the apartment offers to provide you with a used refrigerator for free and promises to maintain and repair the refrigerator during the next five years. You also have the option of buying a new, energy-efficient refrigerator (with a free, five-year maintenance agreement) for $700. The new refrigerator will reduce your electric bill by $150 per year and will have a market value of $200 after five years. If necessary, you can borrow money from the bank at an 8 percent rate of interest. Which option should you choose?

***14.** Suppose that you are considering whether to enroll in a summer computer training program that costs $2,500. If you take the program, you will have to give up $1,500 of earnings from your summer job. You figure that the program will increase your earnings by $500 per year for each of the next ten years. Beyond that, it is not expected to affect your earnings. If you take the program, you will have to borrow the funds at an 8 percent rate of interest. From a strictly monetary viewpoint, should you enroll in the program?

15. "In a world of uncertainty, it is important that entrepreneurs be able to introduce new products and try out innovative ideas. But it is also important that unproductive projects be brought to a halt." Evaluate this statement.

***16.** Will political officials be more likely than private investors and entrepreneurs to channel funds into wealth-creating projects? Why or why not? Discuss.

***Asterisk** denotes questions for **which** answers **appear** in **Appendix** B

CHAPTER

15 Income Inequality and Poverty

All animals are equal, but some animals are more equal than others.

—*George Orwell*[1]

Chapter Focus

- **How do resource prices and income differences influence the incentive of people to develop resources and use them productively?**
- **How much income inequality is there in the United States? Has the degree of inequality changed in recent years?**
- **How much income mobility exists? Do the rich stay rich while the poor stay poor?**
- **How widespread is poverty? How has the poverty rate changed in recent decades? Did the War on Poverty help reduce the poverty rate?**
- **Is there too much income inequality? Should the government try to reduce inequality? Why or why not?**

[1]George Orwell, *Animal Farm* (New York: Harcourt and Brace Company, 1946), p. 112.

Differences in resource prices and the productivity of individuals will cause incomes to vary. Of course. economic inequality is present in all societies. Neither politics nor central planning will eliminate it. Nonetheless, most of us are troubled by the extremes of inequality—the extravagant luxury of some and the grinding poverty of others. How much inequality is there in the United States? Do wealthy families continually enjoy high incomes? Are those in poverty unable to escape it? What effect have income transfer programs had on the welfare of the poor? Do these income transfer programs lessen the incentive for both the poor and nonpoor to earn income? This chapter focuses on these questions and related issues.

HOW MUCH INCOME INEQUALITY EXISTS IN THE UNITED STATES?

Money income is only one component of economic well-being. Factors such as leisure, noncash transfer benefits, the nonmonetary advantages and disadvantages of a job, and the expected stability of future income also affect people's economic welfare. Money income is quite important, however, because it makes it possible to buy market goods and services. It is also easy to measure and is therefore widely used as a yardstick of economic well-being and inequality prevailing in society.

Exhibit 1 presents data on money income in the United States. First, look at the share of *before-tax* annual money income by quintile—that is, by each fifth of the total population of families—ranked from the lowest to the highest. If there were total income equality, each quintile (20 percent) of families would generate 20 percent of the aggregate income. Given the differences in education, skills, ability, work effort, ages, family sizes, and other factors among families, clearly we would not expect this to be the case. Some

EXHIBIT 1
Share of Money Income by Quintile During Selected Years, 1950–2001

	Lowest 20 Percent of Recipients	Second Quintile	Third Quintile	Fourth Quintile	Top 20 Percent of Recipients
Family Income Before Taxes					
1950	4.5	12.0	17.4	23.4	42.7
1960	4.8	12.2	17.8	24.0	41.3
1970	5.4	12.2	17.6	23.8	40.9
1980	5.1	11.6	17.5	24.3	41.6
1990	4.6	10.8	16.6	23.8	44.3
2001	4.2	9.7	15.4	22.9	47.7
Impact of Taxes and Transfers on 2001 Household Income					
Before	3.5	8.7	14.6	23.0	50.1
After	4.4	10.4	16.3	24.0	44.9

Source. Bureau of the Census, *Current Population Survey,* Series P-60; *Statistical Abstract of* the *United States, 1995,* table *733;* Congressional Budget Office; 1994, Green Book; and Bureau of the Census, *Money Income in the United States: 2002.*

families will be able to generate more income than others. It is, however, informative to look at the shares of income by quintile.

The figures show that before-tax income inequality fell throughout the 1950s and 1960s. In other words, the income gap between the top and bottom earning quintiles closed. For example, in 1970, the top quintile earned 40.9 percent of the aggregate money income, down from 42.7 percent in 1950. During the same period, the share of income earned by the lowest quintile rose from 4.5 percent in 1950 to 5.4 percent in 1970.

Beginning in the 1970s, however, this trend reversed. During the last three decades, the share of income earned by the top quintile has steadily risen, while that earned by the bottom group has fallen.[2] By 2001, the share of before-tax money income of the top group had risen to 47.7 percent, while that earned by the bottom 20 percent of families had fallen to 4.2 percent. Thus, in 2001, the top quintile of families earned approximately 11 times as much before-tax money income as the bottom quintile of families.

Low-income families are the primary beneficiaries of noncash transfer programs that provide people with food (food stamps), health care, and housing. Correspondingly, under a system of progressive taxation, which the United States has, taxes take a larger share of income as one's earnings increase. As a result, we would expect *after-tax and transfer* incomes to be more equal than before-tax incomes. Using household data, Exhibit 1 shows that this is indeed the case. After taxes and transfers, the bottom quintile of households received 4.4 percent of the total income in 2001 (compared to 3.5 percent of the before-tax income), and the top quintile of families received 44.9 percent (compared to 50.1 percent before taxes). Notice that taxes and transfers increased the income share of every quintile except the top group. As you can see, taxes and transfers do, in fact, reduce income inequality. *[Note:* Household income is used here because after-tax and family transfer-income data are unavailable. In addition to families, households include individuals living by themselves and unrelated persons living together.]

A market economy does not have a central distributing agency that carves up the economic pie and allocates slices to various individuals. Rather, each individual produces his or her own slice of the pie. If one person's income grows, it does not mean that another's must shrink. The distribution of income is actually a result of many individual efforts and decisions. We now turn to consider what factors underlie the distribution of income in the United States.

The Factors Affecting Income Distribution

How meaningful are the data of Exhibit 1? If all families (and households) were similar except with regard to the amount of income they earned, the use of annual income data as an index of inequality would be quite reasonable. However, the fact is that the aggregate data lump together (1) small and large families, (2) prime-age earners and elderly retirees, (3) multi-earner families and families without any current earners, and (4) husband-wife families and single-parent families.

Consider just one factor: the impact of age and the pattern of lifetime income. Typically, the annual income of young people is low, particularly if they are going to school or are getting training. Many people under 25 years of age studying to be lawyers, doctors, engineers, and economists will have low annual incomes during this phase of their lives. But this does not mean they are poor—at least not in the usual sense. After completing their formal education and acquiring work experience, these people move into their prime working years. At this point, their annual incomes are generally quite high, particularly in families in which both spouses are employed outside the home. Remember, however, that this is also a time when families are purchasing homes and providing for their children, which is costly. Consequently, the economic well-being of these people as measured by their earnings may well be overstated. In other words, although these people are earning

[2] A reduction in the share of income earned by the bottom quintile income group does not imply that their income level fell. It merely indicates that their income did not grow as rapidly as that of other groups, particularly the top quintile. The inflation-adjusted income of the bottom quintile of earners has increased in recent decades.

more at this stage in their lives, they also have more responsibilities. Finally, people reach the retirement phase, characterized by less work, more leisure, and smaller family sizes. Even families that are quite well-off tend to experience incomes well below average during the retirement phase. ***Given the life cycle of income, lumping together families of different ages (phases of their life-cycle earnings) results in substantial inequality in the annual income figures. This would be so even if the incomes of families over a lifetime were approximately equal.***

Exhibit 2 highlights major differences between high- and low-income families that underlie the distributional data of Exhibit 1. The typical high-income family (top 20 percent) is headed by a well-educated person in the prime working-age phase of life. His or her income is supplemented by the earnings of other family members, particularly working spouses. In contrast, people with little education, nonworking retirees, younger workers (under age 35), and single-parent families are substantially overrepresented among low-income families (the bottom 20 percent of income recipients). In 2001, 34 percent of the householders in the lowest income quintile failed to complete high school, compared to only 3 percent for the highest income quintile. Although only 9 percent of the household heads in the bottom quintile completed college, 58 percent of the householders in the top group did so. Seventy-nine percent of the high-income families had household heads in the prime working-age category (age 35 to 64), compared with only 44 percent of the low-income families. Only one parent was present in 49 percent of the low-income families, whereas 93 percent of the high-income group were dual-parent families. High-income families are also larger than low-income families, which might surprise you. In 2001, there were 3.4 persons per family in the top income quintile, compared to only 2.9 persons per family in the bottom quintile.

There was a striking difference between low- and high-income families in the percentage of hours each group worked. No doubt, much of this difference reflected factors like family size, the age of the wage earners, whether or not both spouses worked, and whether two spouses are present in the home. In high-income families, the average number of workers per family was 2.2, compared with 0.80 for low-income families. Among married-couple families, a wife working full time was present only 13 percent of the time in low-income families, compared to 63 percent of the time in high-income families. As

EXHIBIT 2
Differing Characteristics of High- and Low-income Families, 2001

	Bottom 20 Percent of Income Recipients	Top 20 Percent of Income Recipients
Education of Householder		
Percent with less than high school	34	3
Percent with college degree or more	9	58
Age of Householder (percent distribution)		
Under 35	32	13
35-64	44	79
65 and over	24	8
Family Status		
Married-couple family (percent of total)	51	93
Single-parent family (percent of total)	49	7
Persons per Family	2.9	3.4
Earners per Family	0.80	2.2
Percent of married-couple families in which wife works full time	13	63
Percent of Total Hours Worked Supplied by Group	10	25

Source. U S. Department of Commerce, Money Income in the United States: 2001 (Washington, DC: Government Printing Office, 2002).

we would expect, couples who decide not to have the wife work full-time pay for this choice by falling down the income distribution ladder.

In terms of their work effort, the top 20 percent of income recipients contributed 25 percent of the total number of hours worked in the economy, whereas the bottom 20 contributed only 10 percent of the total work time. In other words, high-income families worked 2.5 times as many hours as low-income families. As we noted on page 590, they also earned approximately 11 times as much before-tax income. This implies that the earnings per hour worked by the top income recipients were only about 4.4 times the earnings per hour worked by the low-income recipients. Clearly, differences in the amount of time worked were a major factor contributing to the income inequality of Exhibit 1.

In summary, Exhibit 2 sheds a great deal of light on the distributional data presented in Exhibit 1. ***Those with high incomes are far more likely to be well-educated, dual-parent families with both parents working outside the home in their prime earning years. In contrast, those with low incomes are often single-parent families headed by a poorly educated adult who is either youthful or elderly. The household heads of those families with little income are often either out of the labor force or working only part-time.*** Given these factors, it is not surprising that the top 20 percent of recipients have substantially higher incomes, both before and after taxes, than the bottom quintile.

Why Has Income Inequality Increased?

Exhibit 1 indicates that there has been an increase in income inequality in the United States during the last couple of decades. Why has the gap between the rich and the poor been growing? Research in this area shows that at least four factors have contributed to the rise in inequality.

1. The Increasing Proportion of Single-parent and Dual-earner Families The nature of the family and the allocation of work responsibilities within the family have changed dramatically in recent decades. In 2002, more than one-fourth (28 percent) of all families with children were headed by a single parent, double the figure of the mid-1960s. At the same time, the labor force participation rate of married women increased from 40 percent in 1970 to 62 percent in 2002.

As a result, today we now have more single-parent families and more dual-earner families. Both of these changes tend to promote income inequality. Consider two hypothetical families, the Smiths and the Browns. In 1970, both were middle-income families with two children and one parent working earning $40,000 (in 2005 dollars). Now consider their 2005 counterparts. The Smiths of 2005 are divorced, and one of them, probably Mrs. Smith, is trying to work part-time and take care of the two children. The probability is very high that the single-parent Smith family of 2005 will be in the low- rather than the middle-income category. The Smiths may well be in the bottom quintile of the income distribution. In contrast, the Browns of 2005 both work outside the home, and each earns $40,000 annually. Given their dual incomes, the Browns are now a high- rather than a middle-income family. Along with many other dual-income families (see Exhibit 2), the Browns' 2005 family income will probably place them in the top quintile of income recipients.

Even if there were no changes in earnings between skilled and less-skilled workers, the recent changes within the family would increase the income inequality among families and households. Today, more single-parent families like the Smiths fall into the low-income quintile, while more dual-earner families like the Browns fall into the high-income quintile. The gap between the earnings of these two groups has therefore increased the overall income inequality in the United States.

2. Earnings Differentials between Skilled and Less-skilled Workers In 1970, workers with little education who were willing to work hard, often in a hot and sweaty environment, received high wages. This is less true today. In 1974, the annual earnings of men who graduated from college were only 27 percent higher than the earnings of male high school graduates—hardly a huge payoff for the time and cost of a college degree.

Since 1974, however, things have changed dramatically. By the mid-1980s, the earnings premium of male college graduates relative to male high school grads rose to the 50–60 percent range, approximately twice the premium of 1974. By 2002, the income premium of male college graduates relative to high school graduates rose to 85 percent. Similarly, during the last two decades, the earnings of women college graduates have increased sharply relative to women with only a high school education.

Why have the earnings of persons with more education (and skill) risen relative to those with less education (and skill)? Deregulation of the transportation industry and the declining power of unions may have reduced the number of high-wage, blue-collar jobs available to workers with little education. International competition has also played a role because, increasingly, American workers must compete in a global economy. Furthermore, innovations and cost reductions in both communications and transportation have made it easier for firms to hire workers in different locations. Firms producing goods that require substantial amounts of low-skilled labor can easily move to places like Korea, Taiwan, and Mexico, where low-skilled labor is cheaper. In contrast, the United States is more attractive than most other countries to firms requiring substantial amounts of high-skilled, well-educated workers. Globalization has, therefore, reduced the demand for American workers with few skills and little education but increased the demand for high-skilled workers with college degrees. This has made for a wider earnings gap among workers in high-income countries like the United States. However, it is worth noting that this process enhances income levels in low-income countries and thereby reduces *worldwide* income inequality.[3]

3. The Increasing Number of Markets Characterized by a Few People at the Top with Very High Earnings As we discussed in the chapter on wage differentials, the market compensation of star entertainers and athletes, the most talented professionals, and top business executives is often like that of winner-take-all tournaments. At any point in time, a few people at the top have huge earnings, while most others in these areas have modest or even low incomes. As transportation and communication costs have declined, markets have increasingly become national and even global, rather than local. This increases the incomes of a few people at the top, but it also increases the degree of income inequality.

4. Sharply Lower Marginal Tax Rates than in the 1970s Prior to 1981, high-income Americans confronted top marginal tax rates of up to 70 percent (SO percent of their earnings). This encouraged high-income earners to structure their business affairs and invest in ways that sheltered much of their income from the Internal Revenue Service. As we indicated in Chapter 4, the taxable incomes of the top 10 percent of earners expanded sharply when the top marginal tax rates were reduced to the 30 percent range during the 1980s. Some of this increase in income reflected greater work effort due to the increased incentive to earn. Much of it, however, merely reflected the fact that people were engaging in fewer activities to shelter their earnings from taxes. The flip side of fewer tax shelter activities accompanying lower marginal tax rates led to an increase in the visible income of the rich. As more of the income of the rich is observable, money income statistics such as those of Exhibit 1 will register an increase in income inequality.

INCOME MOBILITY AND INEQUALITY IN ECONOMIC STATUS

Income mobility
Movement of individuals and families either up or down income-distribution rankings when comparisons are made at two different points in time. When substantial income mobility is present, a person's current position in the rankings will not be a very good indicator of what his or her position will be a few years in the future.

Statistics on the distribution of annual income fail to reveal income mobility — movement up and down the income ladder. Therefore, they may be misleading. Consider two countries with identical distributions of annual income. In both cases, the annual income of the top quintile of income recipients is ten times greater than the bottom quintile. Now, suppose that in the first country — we will refer to it as Static — the same people are at the top

[3]For additional information on the relationship between international trade and income inequality, see Gary Burtless, "International Trade and the Rise in Earnings Inequality," *Journal of Economic Literature* (June 1995): 800–816; and Symposium on "Income Inequality and Trade," *Journal of Economic Perspectives* (Summer 1995).

of the income distribution year after year. Similarly, the poor people of Static remain poor year after year. Static is characterized by an absence of income mobility. In contrast, earners in the second country, which we will call Dynamic, are constantly changing places. Indeed, during every five-year period, each family spends one year in the upper-income quintile, one year in each of the three middle-income quintiles, and one year in the bottom-income quintile. In Dynamic, no one is rich for more than one year (out of each five), and no one is poor for more than a year. Obviously, the nature of economic inequality in Static is vastly different from that in Dynamic. You would not know it, though, by looking at their identical annual income distributions.

The contrast between Static and Dynamic highlights an important point: income mobility is an important facet of economic inequality. Until recently, detailed data on income mobility were sparse. This is now beginning to change.[4] **Exhibit 3** presents data on the mobility of family income between 1988 and 1998 in the United States. The data compare the relative income positions of the ***same families*** at two different points in time. Based on their 1988 income, each family was placed into income quintiles ranked from highest to lowest. Later, the 1998 real-income level of the ***same families*** was used once again to group the income of each by quintiles.

The first row of Exhibit 3 shows the relative income position in 1998 of people in the top quintile of income recipients in 1988. Approximately one-half (53 percent) of those with incomes in the top quintile in 1988 were able to maintain this lofty position 10 years later. Slightly less than one-half (47 percent) of the top income group in 1988 had fallen to a lower income quintile by 1998. However, less than 1 in 10 of the high-income individuals fell to one of the bottom two quintiles of the 1998 income distribution. This suggests that once families are able to achieve high-income status, they rarely fall back to a very low level of income.

The bottom row of Exhibit 3 tracks the experience of those in the lowest-income quintile in 1988. Slightly more than half (53 percent) of the families in the lowest-income quintile remained there in 1998. Nearly one-quarter (23 percent) moved up to one of the top three income quintiles in 1998. Among those in the next-to-lowest income quintile in 1988, nearly two-fifths (38 percent) had moved up to one of these higher-income groupings by 1998.

The income mobility data reveal something concealed by the annual figures: there is considerable movement up and down the economic ladder. Relative income positions often change over time. A sizable portion of those with a high relative income during one year subsequently find themselves in a lower income position. At the same time, many of

EXHIBIT 3
Income Mobility—Income Ranking, 1988 and 1998

	PERCENTAGE DISTRIBUTION BY INCOME STATUS OF FAMILY IN 1998				
INCOME STATUS OF FAMILY, 1988	HIGHEST QUINTILE	NEXT-HIGHEST QUINTILE	MIDDLE QUINTILE	NEXT-LOWEST QUINTILE	LOWEST QUINTILE
Highest Quintile	53.2	23.2	14.9	5.7	3.0
Next-Highest Quintile	25.8	31.1	23.7	12.9	6.5
Middle Qurntiie	12.6	27.5	28.3	20.7	10.9
Next-Lowest Quintile	4.3	11.0	22.6	36.3	25.7
Lowest Quintile	4.3	6.4	12.4	23.6	53.3

Source. Katherine Bradbury and Jane Katz, "Women's Labor Market involvement and Family income Mobility When Marriages End," *New England Economic Review,* Fourth Quarter 2002, pages 41–74.

[4]For a review of the literature on income mobility, see Isabel V Sawhill and Daniel P McMurrer, ***Income Mobility in the United States*** (Washington, DC Urban Institute, 1996) For an early classic work on thir topic, see Greg J Duncan et al , ***Years of Poverty, Years of Plenty The Changing Fortunes of American Workers and Families*** (Ann Arbor Institute for Social Research, University of Michigan, 1984)

GETTY IMAGES

Generational Mobility: There is a weak positive correlation between the earnings of fathers and sons. If a father has lifetime earnings 20 percent above the average of his generation, a son can expect to earn about 8 percent more. There is virtually no correlation between the earnings of grandparents and their grandchildren. Apparently, there is some truth in the old saying, "From shirtsleeves to shirtsleeves in three generations."

those with low relative incomes during a given year move up to higher-income quintiles in subsequent years.

Many economists argue that differences in household expenditures are a more accurate indicator of economic status than annual income precisely because of this. If your income or earnings prospects in the future are good, you're likely to spend more than your current income. On the other hand, if your future income is expected to be lower, you're likely to save more, causing spending to be less than your current income. To a large degree, current expenditures reflect long-term economic status.

What do expenditures reveal about the degree of inequality in the United States? The expenditure statistics indicate that the degree of inequality is significantly smaller than the corresponding figure for annual income. For example, in recent years, the household expenditures of the bottom 20 percent of households have summed to approximately 7 percent of the total. This is substantially greater than their income share both before and after taxes. Correspondingly, the expenditures of the top 20 percent of households have constituted approximately 38 percent of the total, well below their income share. Moreover, the share of household expenditures by quintile has been relatively constant during the last four decades. This suggests that the increase in inequality as measured by the annual income data is a reflection of dynamic change, temporary fluctuations in annual income, and measurement issues rather than a true increase in economic inequality.

POVERTY IN THE UNITED STATES

In an affluent country like the United States, income inequality and poverty are related issues. Poverty could be defined in strictly relative terms — the bottom one-fifth of all income recipients, for example. However, this definition would not be very helpful because it would mean that poverty would never fall below what these 20 percent earn.

The official definition of poverty in the United States is based on the perceived minimum income necessary to provide food, clothing, shelter, and basic necessities economically for a family to survive. This **poverty threshold income level** varies with family size and composition, and it is adjusted annually for changes in prices. The official poverty threshold is based only on money income. (See the Measures of Economic Activity feature for additional details on how the poverty rate is measured.)

Poverty threshold income level
The level of money income below which a family is considered to be poor. It differs according to family characteristics (for example, number of family members) and is adjusted when consumer prices change.

Who are the poor? **Exhibit 4** presents data on both the number of poor families and the poverty rate for 1959, 1976, and 2003. In 2003, 7.6 million families, 10.0 percent of

MEASURES OF ECONOMIC ACTIVITY

How Is the Poverty Rate Calculated?

Families and individuals are classified as poor or nonpoor based on the poverty threshold income level originally developed by the Social Security Administration (SSA) in 1964. Consumption survey data showed that low- and median-income families of three or more persons spent approximately one-third of their income on food. Because of this, the SSA established the poverty threshold income level at three times the cost of an economical, nutritionally adequate food plan. A slightly larger multiple was used for smaller families and individuals living alone. The poverty threshold figure varies according to family size, because the food costs vary by family size and composition. It is adjusted annually to account for rising prices. The following chart illustrates how the poverty threshold for a family of four has increased as prices have risen from 1959 to 2004:

1959	$2,973
1970	3,968
1980	8,414
1990	13,359
2004	18,850

Even though the poverty threshold income level is adjusted for prices, it is actually an absolute measure of economic status. As real income increases, the poverty threshold declines relative to the income of the general populace.

The official poverty rate is the number of persons or families living in households with a money income below the poverty threshold as a percentage of the total. Only money income is considered. Income received in the form of noncash benefits, like food stamps, medical care, and housing subsidies, is completely ignored in the calculation of the official poverty rate. Because of this omission, the official rate tends to overstate the degree of poverty. To remedy this deficiency, the Bureau of the Census has developed several alternative measures of poverty that count the estimated value of noncash benefits as income. In addition to the official poverty rate, the bureau now publishes annual data for the "adjusted" poverty rates that include a valuation for various noncash benefits. Of course, including these benefits reduces the poverty rate. For example, the official poverty rate for families was 10.0 percent in 2003, but the adjusted poverty rate that included the value of the noncash food, housing, and medical benefits was only 7.4 percent.

The poverty rate is calculated each year based on a survey of about 60,000 households designed to reflect the population of the United States. The two major sources for comprehensive data on this topic are the Bureau of the Census annual publications, *Money Income in the United States* and *Poverty in the United States.*

EXHIBIT 4
The Changing Composition of the Poor and Poverty Rates, Selected Groups: 1959, 1976, and 2003

	1959	1976	2003
Number of Poor Families (in millions)	8.3	5.3	7.6
Percent of Poor Families Headed by a:			
Female	23	48	51
Black	26	30	27
Elderly person (age 65 and over)	22	14	10
Person who worked at least some during the year	70	55	48
Poverty Rate			
All families	18.5	10.1	10.0
Married couple families	15.8	7.2	5.4
Female-headed families	42.6	32.5	28.0
All individuals	22.4	11.7	12.5
Whites	18.1	9.1	10.5
Blacks	55.1	31.1	24.4
Children (under age 18)	27.3	16.0	17.6

Sources. U.S. *Department* of *Commerce, Characteristics* of *the Population Below the Poverty Level 1982, Table 5; and Poverty* in *the United States. 2003, (P60-226)*

the total, were classified as poor. Half of these families were headed by a woman. As **Exhibit** 4 shows, there has been a substantial change in the composition of the poverty population during the last four decades. In 1959, elderly persons and the working poor formed the core of the poverty population. Twenty-two percent of the poor families were headed by an elderly person in 1959 versus only 10 percent by 2003. In 1959, 70 percent of the heads of poor families worked at least part of the year, but in 2003, only 48 percent of them worked at least part of the year.

During the last several decades, the proportion of families headed by women in the United States has increased while the proportion of husband-wife families has decreased. Since the poverty rate of families headed by women is several times higher than the rate for husband-wife families (28.0 percent compared with 5.4 percent in 2003), an increase in family instability tends to push the poverty rate upward. It also tends to increase the number of households headed by women among poor families. In 2003, 51 percent of the poor families were headed by a woman, compared to only 23 percent in 1959.

Rather than calculating the poverty rate of families, one could measure the poverty rate of individuals. The poverty rate of individuals is, in fact, somewhat higher than the poverty rate of families. In 2003, 12.5 percent of people in the United States were classified as poor, compared to 10.0 percent for families. The poverty rate of blacks in 2003 was 24.4 percent, compared to 10.5 percent for whites. Nonetheless, 70 percent of the people in poverty were white. Perhaps the most tragic consequence of poverty is its impact on children. Little progress has been made in this area during the last two decades. In 2003, 17.6 percent of the children in the United States lived in poverty, up from 16.0 percent in 1976.

Just as people move up and down the income distribution ladder, there is movement into and out of poverty. **A** large proportion of poor families remain poor for only a brief period of time. For example, 34.2 percent of the U.S. population was poor for two consective months between 1996 and 1999. However, only 2.0 percent of the population were poor during the entire period. In fact, the median duration of poverty spells during 1996–1999 was only 4 months. Many of these short-term spells of poverty were the result of factors like medical problems or job changes. In fact, poverty is a long-term problem for only a small number of families in the United States.

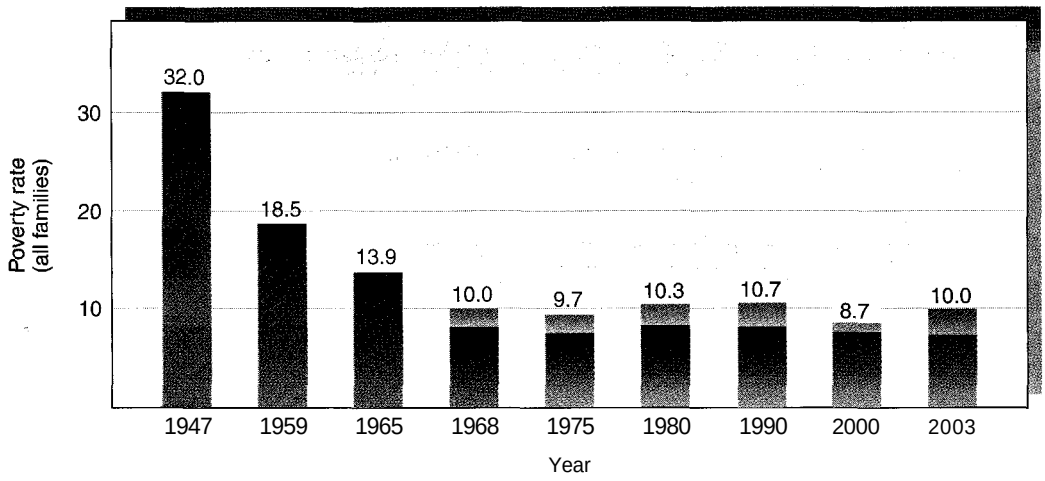
Transfer Payments and the Poverty Rate

In the mid-1960s, it was widely believed that an increase in income transfers directed toward the poor would substantially reduce, if not eliminate, the incidence of poverty. The 1964 *Economic Report of the President* argued that poverty could be virtually eliminated if the federal government increased its expenditures on transfer programs by approximately 2 percent of aggregate income. Following the declaration of the "War on Poverty" by President Lyndon Johnson's administration, transfer expenditures increased rapidly. Measured in 1982–1984 dollars, means-tested income transfers—those limited to people with incomes below a certain cutoff point—tripled, expanding from $24 billion in 1965 to $70 billion in 1975. After 1975, transfer expenditures continued to expand, but the rate of increase was slower. Measured as a share of aggregate income, means-tested transfers jumped from 1.5 percent in 1965 to 5.2 percent in 2000.

Means-tested income transfers
Transfers that are limited to persons or families with an income below a certain cutoff point. Eligibility is thus dependent on low-income status.

Did the expansion in government income transfers reduce the poverty rate as the 1964 *Economic Report of the President* anticipated? Exhibit **5** shows the poverty rate from 1947 through 2003. Interestingly, the poverty rate declined substantially during the period prior to the War on Poverty. It fell from 32 percent in 1947 to 13.9 percent in 1965. The downward trend continued for a few more years, reaching 10 percent in 1968. However, shortly after the War on Poverty was initiated, the downward trend in the poverty rate came to a halt. Since 1968, the poverty rate has fluctuated within a narrow band around the 10 percent level. In 2003, it was 10.0 percent—the same as it was in 1968. When noncash benefits like Medicaid and food stamps were counted as income, the poverty rate was lower, but the overall pattern is still much the same. This is particularly remarkable when you consider that over this same time span, real income per person in the United States approximately doubled.

EXHIBIT 5
The Poverty Rate, 1947-2003



The official poverty rate of families fell sharply during the 1950s and 1960s, but it has remained near 10 percent since 1968. The shaded area of the bars show the additional decline in the poverty rate when noncash benefits are counted as income. When the value of these noncash transfer benefits is added, the poverty rate ranges from 7 percent to 8 percent. This statistic remained unchanged throughout most of the 1968–2003 period.

Sources: U.S. Department of Commerce, *Characteristics of the Population Below the Poverty Level: 1982, Table 5;* and *Poverty in the United States: 2002,* (P60-226).

Why wasn't the War on Poverty more effective? One reason is because means-tested income transfers generate two major side effects that reduce their effectiveness. ***First, transfer programs that significantly reduce the hardship of poverty also reduce the opportunity cost of choices that often lead to poverty.*** This factor is sometimes called the **Samaritan's dilemma**. To the extent that antipoverty programs reduce the negative effects of, for example, births by unmarried mothers, abandonment of children by fathers, dependence on drugs or alcohol, or dropping out of school, they inadvertently encourage people to make choices that result in these conditions. Of course, this is not the intent of the transfers, but nonetheless it is one of their side effects. In the short run, these secondary effects are probably not very important. Over the longer term, however, their negative consequences can be substantial.

Samaritan's dilemma
The dilemma that occurs when assisting low-Income citizens with transfers reduces the opportunity cost of choices that lead to poverty. Providing income transfers to the poor and discouraging behavior that leads to poverty are conflicting goals.

Second, income-linked transfers reduce the incentive of low-income individuals to help themselves. When the size of transfers is linked to income, larger transfers tend to increase the **implicit marginal tax rate** imposed on the poor. The implicit marginal rate is the amount of additional (marginal) earnings that must be paid explicitly in taxes or—in the case of the poor—implicitly in the form of lower income supplements. The marginal tax rate determines the fraction of an additional dollar earned that an individual is permitted to keep, so it is an important determinant of the incentive to work. Participants in the food stamp program, for example, have their food stamp benefits reduced by $30 for every $100 of income they earn. Consequently, every $100 of additional income leads to only a $70 increase in well-being (spendable net income) after the reduction in food stamp benefits is taken into account. When a person qualifies for several programs like food stamps, Medicaid, and housing benefits, the problem is compounded, and the combined implicit marginal tax rate can frequently exceed 50 or 60 percent. In some cases, the rate may exceed 100 percent. In other words, if recipients work, the money they make can actually reduce the family's economic status. Because the high implicit marginal tax rates reduce the incentive of poor families to work and earn, transfers often merely replace income that otherwise would have been earned. When this is the case, they add little or nothing to the ***net*** income of the poor.

Implicit marginal tax rate
The amount of additional (marginal) earnings that must be paid explicitly in taxes or implicitly in the form of lower income supplements. The marginal tax rate establishes the fraction of an additional dollar earned that an individual is permitted to keep, so it is an important determinant of the incentive to work.

Estimating the Costs of Redistribution

The U.S. government has numerous income transfer programs, including many that are unrelated to helping the poor. In 2003, income transfers and subsidy programs intended to help various groups purchase food, housing, and health care summed to 22 percent of aggregate income. As we discussed in Chapter 6, politicians will often be able to gain by proposing programs that provide concentrated benefits to readily identifiable groups (for example, farmers, senior citizens, or unemployed workers) while imposing a small personal cost on disparate groups like taxpayers and consumers in general. Because of this, the presence and size of transfer programs is not surprising.

Income transfers are also costly in terms of the economy's overall output level. The higher marginal tax rates needed to finance transfers and subsidies mean that taxpayers get to keep less of what they earn. Furthermore, recipients often qualify for larger transfers

APPLICATIONS IN ECONOMICS

Competition for Transfers and the Net Gain of Recipients

As the data of this chapter show, the achievements of antipoverty transfer programs fell short of initial expectations. Indeed, the economic way of thinking would lead a person to believe that it's difficult to transfer income to a group of recipients in a way that will significantly improve their long-term well-being. The unintended consequences that alter people's incentives explain why this is true.

Governments must establish criteria to determine who should get the transfers, subsidies, and other political favors and who should not. Without these criteria, benefit programs would exhaust their budgets. Typically, governments require potential recipients to do something, own something, buy something, or be something in order to qualify for the transfer. But, as people take the steps necessary to qualify, they erode much of the benefit.

Consider the case in which the recipients have to do something (for example, fill out forms, take an exam, submit a detailed proposal, lobby government officials, endure delays, or contribute to selected political campaigns) in order to obtain the transfer. These actions are costly and will reduce the recipient's net gain. In some cases, almost all of the gain may be eroded. Perhaps a simple example will help illustrate this important point. Suppose the U.S. government decided to give away a $50 bill between 9 a.m. and 5 p.m. each week day to anyone willing to show up at one of six teller windows at the U.S. Department of Treasury. Certainly, a line would emerge. How much time would people be willing to wait in the line? A person whose time was worth $5 per hour would be willing to spend up to 10 hours waiting in line for the $50. But it might take longer than 10 hours if there were enough other people, whose time was worth less, say $3 or $4 per hour. And everyone would find that the waiting consumed much of the value of the $50 transfer. If the government's objective is to make the recipients $50 richer, it will fail. In fact, some of those waiting in line for the $50 bill will gain very little. For example, if the wait were 10 hours, the net benefit of a person with an opportunity cost of $4.90 would be only $1. So it is with real-world transfers. When you consider the cost of what has to be done for someone to qualify for the transfer, the net gain of the recipient is often small.

In other cases, the government will require recipients to own something (for example, land with an acreage allotment to grow wheat, a license to operate a taxicab, or a permit to sell in a restricted market) in order to qualify for a transfer or subsidy. Recipients will gain when programs like this are initiated or unexpectedly expanded. But once programs of this type are established, people will bid up the price of the asset needed to acquire the subsidy. The higher price of the asset, such as the taxicab license or the land with a wheat allotment, will capture the value of the expected future subsidy. Once this happens, the rate of return in the business or occupation (for example, farming or operating a taxicab) will be driven down to the normal level. The current recipients don't benefit from the subsidy program continuing at the expected benefit level. Essentially, they paid for the expected benefit of the subsidy when they purchased the asset needed to get it.

Interestingly, even though they don't benefit, current recipients will fight to keep the program going if it is threatened. Repealing (or reducing) the subsidies will be harmful because it will reduce the value of the asset they currently hold. This explains why eliminating farm subsidies, for example, is difficult even though the subsidies do not make farming more profitable for the current operators.

(continued)

APPLICATIONS IN ECONOMICS

(continued)

Economist Gordon Tullock refers to this as the *transitional gains* trap, because, once programs like these get started, they will be difficult to eliminate even though current recipients aren't benefiting from them.[1]

Sometimes recipients are required to purchase a good or service in order to qualify for a subsidy. For example, the Medicare and Medicaid programs subsidize health care purchases made by senior citizens and the poor. Several states provide educational subsidies for college tuition or the purchase of textbooks. As we discussed in Chapter 4, subsidizing the purchase of a good will increase its demand and lead to higher prices. As a result, some of the benefits of such subsidies will accrue to the suppliers rather than the purchasers. Again, the subsidized recipients will not gain as much as it might appear.

Finally, sometimes the government requires that recipients be something in order to qualify for the subsidy. Programs of this type are often intended to help protect people against the adverse effects of certain events such as unemployment or catastrophic storms. Unfortunately, the programs also encourage choices that increase the adverse effects of such events. For example, consider the impact of subsidized insurance premiums for those living in hurricane-prone areas. These subsidies reduce the owners' personal cost of rebuilding their properties after a hurricane. But the lower personal costs also encourage people to build in hurricane-prone areas. As more of them do so, the overall cost of damage from hurricanes increases. The impact of unemployment insurance is much the same. The benefits paid to the unemployed will reduce the opportunity cost of continuing a job search while unemployed. As a result, periods of unemployment will be longer and the rate of unemployment higher than would have otherwise been the case.

The ability of transfers to improve the well-being of a designated group is more limited than is generally recognized. Transfer programs alter incentives and generate unintended consequences that erode some, if not most, of the benefits derived by the recipients. Thus, the net gains of the intended beneficiaries are often much less than they appear to be.

[1]See Gordon Tullock, "The Transitional Gains Trap," *Bell Journal of Economics* 6 (Autumn 1975), 671–678.

if their income is lower, which motivates them to earn even less. Thus, transfer programs typically weaken the incentive of both taxpayer-donors and the subsidy recipients to work and earn. As a result, the cost of the transfers will be greater than the amount of income transferred. A study by economists Sam Allgood and Arthur Snow sought to measure the loss of output due to the lower supply of labor associated with higher marginal tax rates accompanying income transfers.[5] Allgood and Snow estimated that it costs between $1.26 and $3.22 in terms of lost output for every additional dollar redistributed from the top 60 percent of income recipients to the bottom 40 percent.

Helping transfer recipients achieve long-term well-being is also more difficult than it first appears. Competition for the transfers and subsidies will often erode much of the recipient's gain. See the Application in Economics "Competition for Transfers and the Net Gain of Recipients" for more about how this works.

INCOME INEQUALITY: SOME CONCLUDING THOUGHTS

Is there too much inequality in the United States? Should transfer programs be expanded or contracted? These are normative questions that positive economics cannot definitively answer. Economics can help identify and quantify the cost (in terms of lost output) and the expected effectiveness of alternative programs, but it can't make a judgment about whether that cost or amount is good or bad. Many people, including a good number of

[5]See Sam Allgood and Arthur Snow, "The Marginal Cost of Raising Tax Revenue and Redistributing Income," *Journal of Political Economy* (December 1998). This study builds on earlier work by Edgar K. Browning and William Johnson, "The Trade-off between Equality and Efficiency," *Journal of Political Economy* (April 1984); and Charles L. Ballard, "The Marginal Efficiency Cost of Redistribution," *American Economic Review* (December 1988).

economists, would like to see income inequality reduced. Some dimensions of U.S. public policy are directed toward this. The progressive income tax system, the **Earned Income Tax Credit** (a tax credit that provides supplementary payments to workers with low incomes), and means-tested programs like food stamps, Medicaid, and housing subsidies are examples.

Earned Income Tax Credit
A feature of the personal income tax system that provides supplementary payments to workers with low incomes.

The problem is that the political process is multidimensional. It represents a diverse set of forces, including powerful special-interest groups. Therefore, it shouldn't surprise you that the political record in this area, like others, is mixed. Social Security, the largest income transfer program, redirects income toward the elderly, today a group with above-average levels of both income and wealth. Agriculture subsidies constitute another large transfer program; the bulk of these benefits go to large farmers with incomes well above the average. Another sizable share of government transfers is allocated to business interests, including many that are large and highly profitable. Thus, many of the income transfers are from those with modest incomes to those with incomes well above the average.

Finally, it is important to recognize that income in a market economy is not like manna from heaven. It is something people produce and earn by providing others with goods and services they are willing to pay for. The allocation of income reflects the choices of individuals with differing preferences, talents, educational levels, entrepreneurial skills, rate of saving, and so on. Wage differentials, profits, losses, and interest rates coordinate the choices of these vastly different individuals and bring them into harmony with one another. Income inequality is a natural outgrowth of this process. Indeed, policies that modify the process can generate perverse, counterproductive incentives.

When it comes to issues of fairness, some people argue that the process—the system, in other words—that generates the outcomes is more important than the actual result. Do all people in the economy have an opportunity to acquire education and training? Are people from all segments of society free to compete in business and labor markets? Do incomes reflect their choices, voluntary exchanges, and productive efforts? These questions are about opportunity, economic freedom, and how income is acquired. Many people believe that these things, rather than the income distribution patterns that result from them, are the key elements of economic fairness.

Perhaps the following example will illustrate the difference between the process and pattern view. Suppose a million people purchase a $10 lottery ticket, and the proceeds are used to finance a $10 million jackpot for one person. This activity will clearly increase income inequality. But is it unfair? Those who adhere to the process view would stress that the outcome merely reflects the voluntary choices of participants who were well aware of the rules of the game prior to their purchasing a ticket. According to this view, it is the process rather than the outcome that is the primary determinant of fairness.

Economics provides insight into both the allocative role and sources of differences in income. It also indicates that it will be costly, in terms of lost output, to redistribute income through taxes and transfers. Of course, this does not reveal whether there is too much inequality or whether the government should play a larger or smaller role in the allocation of income. It does, however, enable us to address these normative issues in a more thoughtful way.

LOOKING AHEAD

Income transfers are the focal point of many current policy issues, including Social Security and health care. The special topics in the following section analyze this issue and other topics.

KEY POINTS

- In 2001, the bottom 20 percent of families earned 4.2 percent of aggregate income; the top 20 percent of families earned approximately 11 times that amount (47.7 percent). After taxes and transfers are taken into account, the top quintile of households earn nearly 10 times the income of the bottom quintile.
- A substantial percentage of the inequality in annual income reflects differences in age, education, family size, marital status, number of earners in the family, and time worked. Young inexperienced workers, students, single-parent families, and retirees are overrepresented among those with low current incomes.
- Income inequality has risen during the last 25 years. The following four factors contributed to this increase: (1) an increase in the proportion of both single-parent *and* dual-earner families, (2) an increase in earnings differentials on the basis of skill and education, (3) more "winner-take-all" markets, and (4) increases in the reported income of those in the top tax brackets due to lower marginal tax rates.
- The tracking of household income over time indicates that there is considerable movement of individuals both up and down the income spectrum. The data on the distribution of income at a particular point in time can be misleading because they do not reflect this movement up and down the income ladder.
- One-tenth of American families were officially classified as poor in 2003. Those living in poverty were generally younger, less educated, less likely to be working, and more likely to be living in families headed by a woman than those who were not poor. There is considerable movement both into and out of poverty. A relatively small proportion of families constitute the long-term poor.
- During the last several decades, income transfers—including means-tested transfers—have expanded rapidly both in real dollars and as a share of personal income. As a weapon against poverty, transfers have been largely ineffective. Even though per capita income increased by more than 100 percent between 1965 and 2002, there was little reduction in the overall poverty rate during this period.
- Income transfers large enough to improve the economic status of the poor will (1) encourage behavior that increases the risk of poverty and/or (2) create high implicit marginal tax rates that reduce the recipient's incentive to earn.
- Positive economics cannot determine how much inequality should be present. The nature of the income-generating process as well as the pattern of income distribution is relevant to the issue of fairness.

CRITICAL ANALYSIS QUESTIONS

1. Do you think the current distribution of income in the United States is too unequal? Why or why not? What criteria do you think should be used to judge the fairness of the distribution of income? Is the final outcome more important than the process that generates the income?

*2. Is annual money income a good measure of economic status? Is a family with an $80,000 annual income able to purchase twice the quantity of goods and services as a family with $40,000 of annual income? Is the standard of living of the $80,000 family twice as high as that of the $40,000 family? Discuss.

3. What is income mobility? If there is substantial income mobility in a society, how does this influence the importance of income distribution data?

*4. Consider a table such as Exhibit 3 in which the family income of parents is grouped by quintiles down the first column, and that of their offspring is grouped by quintiles across the other columns. If there were no intergenerational mobility in this country, what pattern of numbers would appear in the table? If the nation had attained complete equality of opportunity, what pattern of numbers would emerge? Explain.

5. Do individuals have a property right to income they acquire from market transactions? Is it a proper function of government to tax some people in order to provide benefits to others? Why or why not? Discuss.

*6. Because income transfers to the poor typically increase the implicit marginal tax rate they confront,

does a $1,000 transfer payment necessarily increase the income of poor recipients by $1,000? Why or why not?

"7. Sue is a single parent with two children. She is considering a job that pays $800 per month. She is currently drawing monthly cash benefits of $300, food stamp benefits of $100, and Medicaid benefits valued at $80. If she accepts the job, she will be liable for employment taxes of $56 per month and lose all transfer benefits. What is Sue's implicit marginal tax rate for this job?

8. What groups are overrepresented among those with relatively low incomes? Do the poor in the United States generally stay poor? Why or why not?

9. Some people argue that taxes exert little effect on people's incentive to earn income. Suppose you were required to pay a tax rate of 50 percent on all money income you earn while in school. Would this affect your employment? How might you minimize the personal effects of this tax?

10. Large income transfers are targeted toward the elderly, farmers, and the unemployed, regardless of their economic condition. Why do you think this is so? Do you think there would be less income inequality if the government levied higher taxes in order to make larger income transfers? Why or why not?

11. The outcome of a state lottery game is certainly a very unequal distribution of the prize income. Some players are made very rich, while others lose their money. Using this example, discuss whether the fairness of the process or the fairness of the outcome is more important, and how they differ.

12. "Means-tested transfer payments reduce the current poverty rate. However, they also create an incentive structure that discourages self-sufficiency and self-improvement. Thus, they tend to increase the future poverty rate. Welfare programs essentially purchase a lower poverty rate today in exchange for a higher poverty rate in the future." Evaluate this statement.

13. Was the poverty rate increasing or decreasing prior to the War on Poverty initiated by the Johnson administration? As income transfer programs accompanying the War on Poverty increased beginning in the latter half of the 1960s, what happened to the poverty rate?

14. Suppose one family has $100,000 while another has only $20,000. Is this outcome fair? What is your initial reaction? Compare and contrast your views depending upon the following:
   a. The family with the higher income has both a husband and wife working, while the other family has chosen for the wife to remain home with the children rather than work in the labor force.
   b. The family with the higher income is headed by a person who completed a college degree, while the other family is headed by someone who dropped out of high school.
   c. The family with the higher income derived most of its income from the farm subsidy program.
   d. The family with the higher income received it as an inheritance from parents who just died.

*Asterisk denotes questions for which answers are given in Appendix B

Gaining from International Trade

CHAPTER 16

Chapter Focus

- How has the volume of international trade changed in recent decades?
- Under what conditions can a nation gain from international trade?
- What effects do trade restrictions have on an economy?
- How have open economies performed relative to those that are more closed?
- What accounts for the political popularity of trade restraints?
- Do trade restrictions create jobs? Does trade with low-wage countries depress wage rates in high-wage countries like the United States?

The evidence is overwhelmingly persuasive that the massive increase in world competition—a consequence of broadening trade flows—has fostered markedly higher standards of living for almost all countries who have participated in cross-border trade. I include most especially the United States.

—Alan Greenspan[1]

[1] Alan Greenspan, speech before the Alliance for the Commonwealth Conference on International Business (Boston, Massachusetts, June 2, 1999).

We live in a shrinking world. The breakfast of many Americans includes bananas from Honduras, coffee from Brazil, or hot chocolate made from Nigerian cocoa beans. Americans often drive a car produced by a Japanese or European manufacturer that consumes gasoline refined from petroleum extracted in Saudi Arabia or Venezuela. Similarly, many Americans work for companies that sell a substantial number of their products to foreigners.

Spurred by cost reductions in transportation and communications, the volume of international trade has grown rapidly in recent decades. It may surprise some people that most international trade is not between the governments of different nations but rather between people and firms located in different countries. Why do people engage in international trade? The expectation of gain provides the answer. Domestic producers are often able to sell their products to foreigners at attractive prices, and domestic consumers sometimes find that the best deals are available from foreign suppliers. Like other voluntary exchanges, international trade occurs because both the buyer and the seller expect to gain and generally do. If both parties did not expect to gain, they would not agree to the exchange.

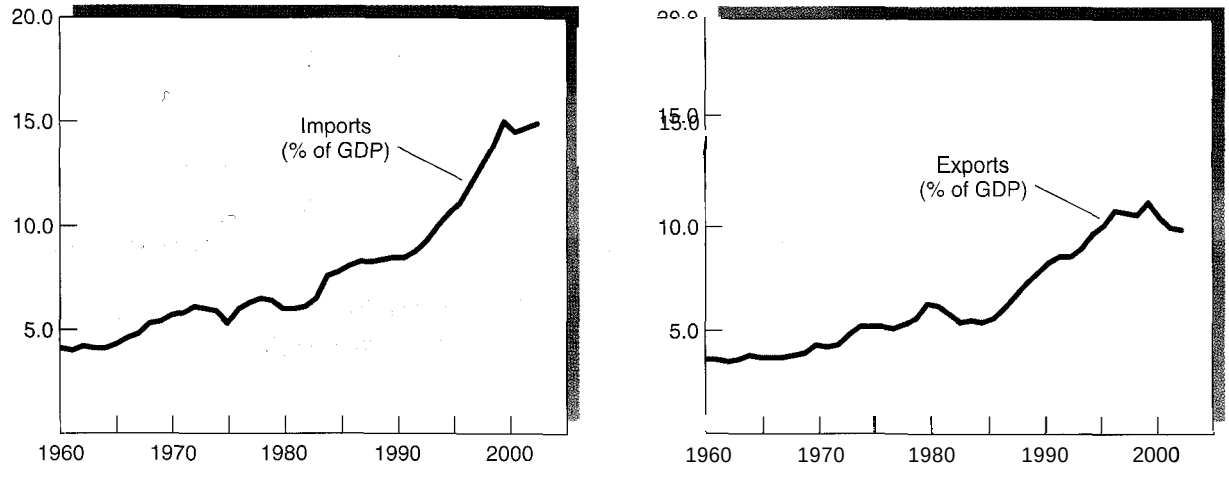
THE TRADE SECTOR OF THE UNITED STATES

As **Exhibit 1** illustrates, the size of the trade sector of the United States has grown rapidly during the last several decades. In 1960, total exports of goods and services accounted for 3.6 percent of the U.S. economy, while imports summed to 4.1 percent. By 1980, both exports and imports were approximately 6 percent of the economy. In 2003, exports accounted for 10 percent of total output, while imports summed to 15 percent. Thus, U.S. international trade (exports + imports) in goods and services has approximately doubled as a share of the economy since 1980 and tripled since 1960.

Who are the major trading partners of Americans? **Exhibit 2** shows the share of U.S. trade (exports + imports) with each of its ten leading trading partners. These ten countries

EXHIBIT 1
The Growth of the Trade Sector in the United States: 1960–2003

During the past several decades, international trade has persistently risen as a share of GDP. Imports of goods and services as a share of GDP rose from 4 percent in 1960 to 6 percent in 1980 and almost 15 percent in 2003. Similarly, exports increased from 4 percent of GDP in 1960 to 6 percent in 1980 and 10 percent in 2003.



Source: http://www.economagic.com. The figures are based on data for real imports, exports, and GDP.

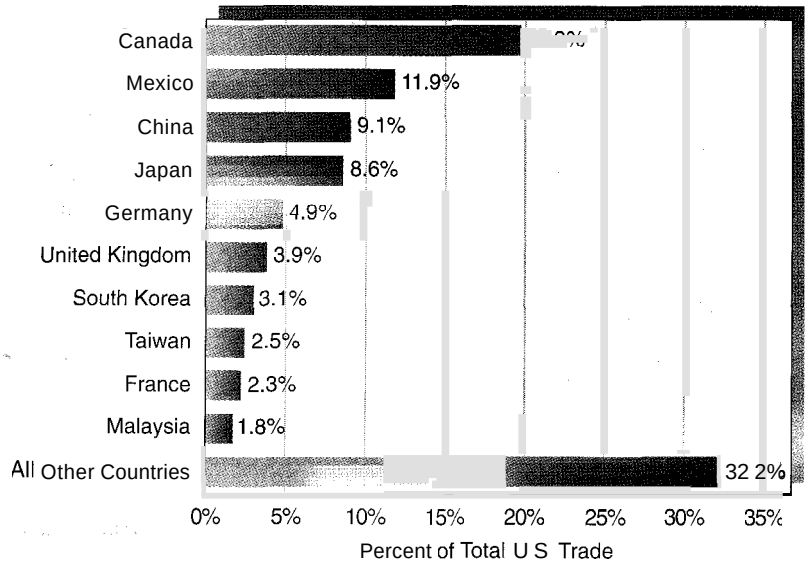


EXHIBIT 2
The 2002 Leading Trading Partners *of* the United States

Canada, Mexico, China, and Japan are the leading trading partners of the United States. Approximately one-half of all U.S. trade involves imports or exports to these four countries.

Source: http://www.census.gov.

account for approximately two-thirds of the total volume of U.S. trade. Canada, Mexico, China, and Japan are the four largest trading partners of Americans. Nearly half of all U.S. trade is with these four countries. The U.S. also conducts a substantial volume of trade with the nations of the European Union, particularly Germany, the United Kingdom, and France.

What are the leading imports and exports of the United States? Capital goods like automobiles, computers, semiconductors, telecommunications equipment, and industrial machines are bought and sold in worldwide markets. The U.S. both imports and exports substantial quantities of these goods. Civilian aircraft, electrical equipment, chemicals, and plastics are also among the leading products exported by the United States. Crude oil, textiles, toys, sporting goods, and pharmaceuticals are major products it imports.

Clearly, the impact of international trade differs across industries. In some industries, domestic producers find it very difficult to compete with their rivals abroad. For example, approximately 90 percent of the shoes purchased by Americans and nearly two-thirds of the radio and television sets, watches, and motorcycles are produced abroad. A high percentage of the clothing and textile products, paper, cut diamonds, and VCRs consumed in the United States are also imported. On the other hand, a large proportion of the aircraft, power-generating equipment, scientific instruments, construction equipment, and fertilizers produced in the United States are exported to purchasers abroad.

GAINS FROM SPECIALIZATION AND TRADE

As we discussed in Chapter 2, the law of **comparative advantage** explains why a group of individuals, regions, or nations can gain from specialization and exchange. ***International trade leads to mutual gains because it allows residents of different countries to: (1) specialize in the production of those things they do best, and (2) import goods foreign producers are willing to supply at a lower cost than domestic producers.*** Resources and labor-force skills differ substantially across countries, and these differences influence costs. A good that is quite costly to produce in one country might be cheaply produced in another. For example, the warm, moist climate of Brazil, Colombia, and Guatemala makes coffee production in these countries more economical than in other places. Countries such as Saudi Arabia and Venezuela with rich oil fields can produce petroleum cheaply. Countries with an abundance of fertile land, such as Canada and

Comparative advantage
The ability to produce a good at a lower opportunity cost than others can produce it. Relative costs determine comparative advantage.

ECONOMICS AT THE MOVIES

20TH CENTURY FOX/DREAMWORKS/ THE KOBAL COLLECTION

Cast Away *(2000)*

In Cast Away, Tom Hanks is stranded alone on an island *for four* years. Hanks must be self-sufficient and produce everything he consumes. Because he's unable to specialize and trade with others, his standard of living while on the island is clearly meager.

Look to here for assignment.

Australia, are able to produce products as wheat, feed grains, and beef at a low cost. In contrast, land is scarce in Japan, a nation with a highly skilled labor force. The Japanese, therefore, specialize in manufacturing, using their comparative advantage to produce cameras, automobiles, and electronic products for export. With international trade, the residents of different countries can gain by specializing in the production of goods they can produce economically. They can then sell those goods in the world market and use the proceeds to import other goods expensive to produce domestically.

The failure to comprehend the principle of mutual gains from trade is often a source of "fuzzy" economic thinking. Because of this, we will take the time to illustrate the principle in detail. To keep things simple, let's consider a case involving only 2 countries, the United States and Japan, and two products, food and clothing. Furthermore, let's assume that labor is the only resource used to produce these products. In addition, since we want to illustrate that gains from trade are nearly always possible, we are going to assume that Japan has an **absolute advantage**—that the Japanese workers are more efficient than the Americans—at producing both food and clothing. **Exhibit 3** illustrates this situation. Perhaps due to their prior experience or higher skill levels, Japanese workers can produce three units of food per day, compared to only two units per day for U.S. workers. Similarly, Japanese workers are able to produce nine units of clothing per day, compared to one unit of clothing per day for U.S. workers.

Absolute advantage
A situation in which a nation, as the result of its previous experience and/or natural endowments, can produce more of a good (with the same amount of resources) than another nation.

Can two countries gain from trade if one of them can produce both goods with fewer resources? The answer is "Yes." As long as the *relative* production costs of the two goods

EXHIBIT 3
Gains from Specialization and Trade

Columns 1 and 2 indicate the daily output of either food or clothing of each worker in the United States and Japan. If the United States moves three workers from the clothing industry to the food industry, it can produce six more units of food and three fewer units of clothing. Similarly, if Japan moves one worker from food to clothing, clothing output will increase by 9 units, while food output will decline by three units. With this reallocation of labor, the United States and Japan are able to increase their aggregate output of both food (three additional units) and clothing (six additional units).

	Output per Worker Day		Potential Change in Output[a]	
Country	Food (1)	Clothing (2)	Food (3)	Clothing (4)
United States	2	1	+6	−3
Japan	3	9	−3	+9
Change in Total			+3	+6

[a]Change in output if the United States shifts three workers from the clothing to the food industry and if Japan shifts one worker from the food to the clothing industry.

differ between Japan and the United States, gains from trade will be possible. Consider what would happen if the United States shifted three workers from the clothing industry to the food industry. This reallocation of labor would allow the United States to expand its food output by six units (two units per worker), while clothing output would decline by three units (one unit per worker). Suppose Japan reallocates labor in the opposite direction. When Japan moves one worker from the food industry to the clothing industry, Japanese clothing production expands by nine units, while food output declines by three units. The exhibit shows that this reallocation of labor *within* the two countries has increased their joint output by three units of food and six units of clothing.

The source of this increase in output is straightforward: aggregate output expands because the reallocation of labor permits each country to specialize more fully in the production of those goods that it can produce at a *relatively* low cost. Our old friend, the opportunity-cost concept, reveals the low-cost producer of each good. If Japanese workers produce one additional unit of food, they sacrifice the production of three units of clothing. Therefore, in Japan the opportunity cost of one unit of food is three units of clothing. On the other hand, one unit of food in the United States can be produced at an opportunity cost of only a half-unit of clothing. American workers are therefore the low-opportunity-cost producers of food, even though they cannot produce as much food per day as the Japanese workers. Simultaneously, Japan is the low-opportunity-cost producer of clothing. The opportunity cost of producing a unit of clothing in Japan is only a third of a unit of food, compared to two units of food in the United States. The reallocation of labor illustrated in Exhibit 3 expanded joint output because it moved resources in both countries toward areas where they had a comparative advantage.

To reiterate: as long as the relative costs of producing the two goods differ in the two countries, gains from specialization and trade will be possible. Both countries will find it cheaper to trade for goods they can produce only at a high opportunity cost. For example, both countries will gain if the United States trades food to Japan for clothing at a trading ratio greater than one unit of food to one half-unit of clothing (the U.S. opportunity cost of food) but less than one unit of food to three units of clothing (the Japanese opportunity cost of food). Any trading ratio between these two extremes will permit the United States to acquire clothing more cheaply than it could be produced within the country and simultaneously permit Japan to acquire food more cheaply than it could be produced domestically.

How Trade Expands Consumption Possibilities

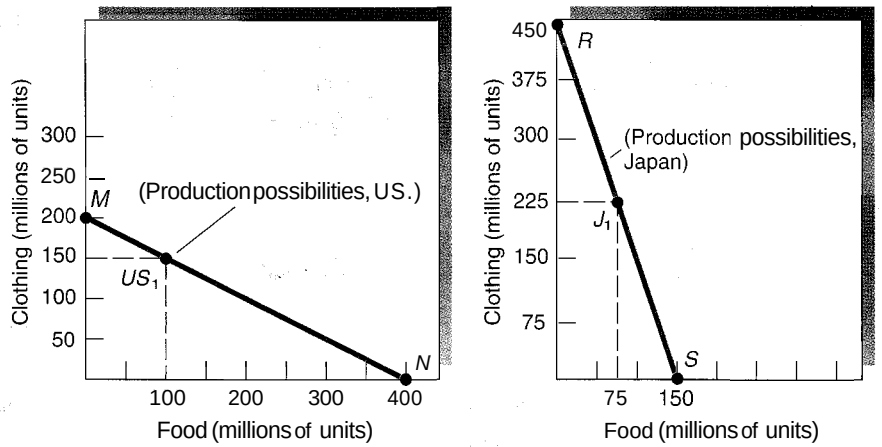
Because trade permits nations to expand their joint output, it also allows each nation to expand its consumption possibilities. The production possibilities concept can be used to illustrate this point. Suppose that there were 200 million workers in the United States and 50 million in Japan. Given these figures and the productivity of workers indicated in Exhibit 3, **Exhibit 4** presents the production possibilities curves for the two countries. If the United States used all of its 200 million workers in the food industry, it could produce 400 million units of food per day—two units per worker—and zero units of clothing (N). Alternatively, if the United States used all its workers to produce clothing, daily output would be 200 million units of clothing and no food (M). Intermediate output combinations along the production possibilities line (MN) between these two extreme points also could be achievable. For example, the United States could produce 150 million units of clothing and 100 million units of food (US_1).

Part (b) of Exhibit 4 illustrates the production possibilities of the 50 million Japanese workers. Japan could produce 450 million units of clothing and no food (R), 150 million units of food and no clothing (S), or various intermediate combinations, like 225 million units of clothing and 75 million units of food (J_1). The slope of the production possibilities constraint reflects the opportunity cost of food relative to clothing. Because Japan is the high-opportunity-cost producer of food, its production possibilities constraint is steeper than the constraint for the United States.

In the absence of trade, the consumption of each country is constrained by its production possibilities. Trade, however, expands the consumption possibilities of both. As we

EXHIBIT 4
The Production Possibilities of the United States and Japan Before Specialization and Trade

Here we illustrate the daily production possibilities of a U.S. labor force with 200 million workers and a Japanese labor force with 50 million workers, given the cost of producing food and clothing presented in Exhibit 3. In the absence of trade, consumption possibilities will be restricted to points such as ***US,*** in the United States and J_1 in Japan along the production possibilities curve of each country.



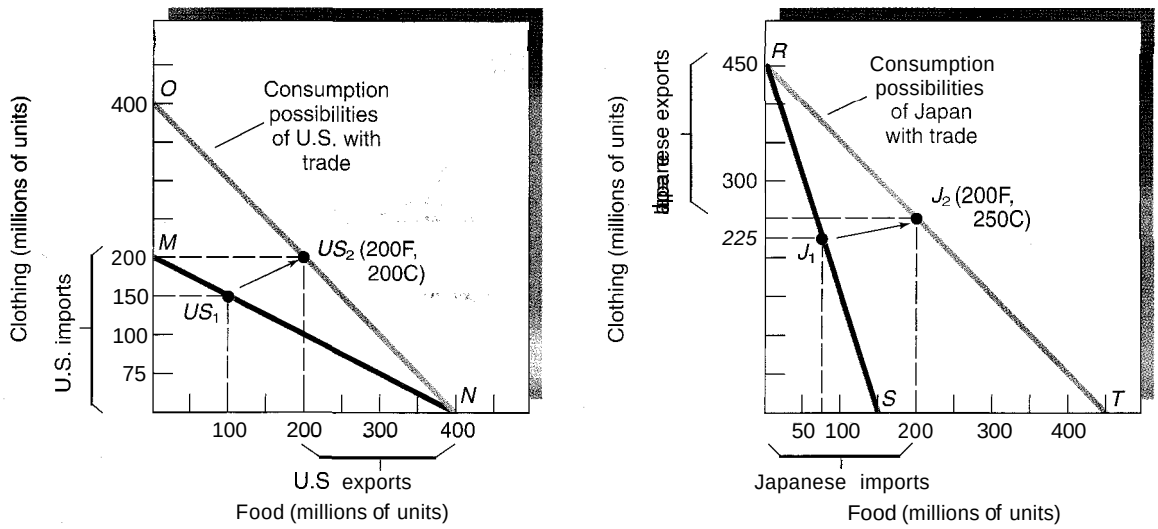
previously said, both countries can gain from specialization if the United States trades food to Japan at a price greater than one unit of food equals one half-unit of clothing but less than one unit of food equals three units of clothing. Suppose that they agree on an intermediate price of one unit of food equals one unit of clothing. As part (a) of **Exhibit 5** shows, when the United States specializes in the production of food (where it has a comparative advantage) and trades food for clothing (at the price ratio where one unit of food equals one unit of clothing), it can consume along the line *ON*. If the United States insisted on self-sufficiency, it would be restricted to consumption possibilities like *US*, (100 million units of food and 150 million units of clothing) along its production possibilities constraint of *MN*. With trade, however, the United States can achieve a combination like *US*, (200 million units of food and 200 million units of clothing) along the line *ON*. Trade permits the United States to expand its consumption of both goods.

Simultaneously, Japan is able to expand its consumption of both goods when it is able to trade clothing for food at the one-to-one price ratio. As part (b) of Exhibit 5 illustrates, Japan can specialize in the production of clothing and consume along the constraint *RT* when it can trade one unit of clothing for one unit of food. Without trade, consumption in Japan would be limited to points like J_1 (75 million units of food and 225 million units of clothing) along the line *RS*. With trade, however, it is able to consume combinations like J_2 (200 million units of food and 250 million units of clothing) along the constraint *RT*.

Look what happens when Japan specializes in clothing and the United States specializes in food. Japan can produce 450 million units of clothing, export 200 million to the United States (for 200 million units of food), and still have 250 million units of clothing remaining for domestic consumption. Simultaneously, the United States can produce 400 million units of food, export 200 million to Japan (for 200 million units of clothing), and still have 200 million units of food left for domestic consumption. Again, after specialization and trade, the United States is able to consume at the point of *US*, and Japan at point J_2, consumption levels that would be unattainable without trade. Specialization and exchange permit the two countries to expand their joint output, and, as a result, both countries can increase their consumption of both commodities.

EXHIBIT 5
Consumption Possibilities with Trade

The consumption possibilities of a country can be expanded with specialization and trade. If the United States can trade one unit of clothing for one unit of food, it can specialize in the production of food and consume along the *ON* line (rather than its original production possibilities constraint, *MN*). Similarly, when Japan is able to trade one unit of clothing for one unit of food, it can specialize in the production of clothing and consume any combination along the line *RT*. For example, with specialization and trade, the United States can increase its consumption from US_1 to US_2, gaining 50 million units of clothing and 100 million units of food. Simultaneously, Japan can increase consumption from J_1 to J_2, a gain of 125 million units of food and 25 million units of clothing.



(a) United States (b) Japan

The implications of the law of comparative advantage are clear: trade between nations will lead to an expansion in total output and mutual gain for each trading partner when each country specializes in the production of goods it can produce at a relatively low cost and uses the proceeds to buy goods that it could produce only at a high cost. It is comparative advantage that matters. As long as there is some variation in the relative opportunity cost of goods across countries, each country will always have a comparative advantage in the production of some goods.

Some Real-World Considerations

To keep things simple, we ignored the potential importance of transportation costs, which, of course, reduce the potential gains from trade. Sometimes transportation and other transaction costs, both real and artificially imposed, exceed the potential for mutual gain. In this case, exchange does not occur.

We also assumed that the cost of producing each good was constant in each country. This is seldom the case. Beyond some level of production, the opportunity cost of producing a good will often increase as a country produces more and more of it. Rising marginal costs as the output of a good expands will limit the degree to which a country will specialize in the production of a good. This situation would be depicted by a production possibilities curve that was convex, or bowed out from the origin. In a case like this, there will still be gains from trade, but generally such a situation won't lead to one country completely specializing in the production of the good.

International Trade

When people are permitted to engage freely in international trade, they are able to achieve higher income levels and living standards than would otherwise be possible.

Of course, domestic markets offer opportunities for specialization and trade also. This is particularly true for large countries like the United States. For example, most Americans specialize in relatively few productive activities and use the income from those activities to purchase most of the goods and services we consume. Many of these goods are domestically produced, but international trade makes it possible for people to buy and sell in an even larger market. Thus, it expands the opportunity for gains from trade. As we just explained, trading partners will be able to produce a larger joint output and consume a larger, more diverse bundle of goods when they each specialize in areas where they have a comparative advantage. Open markets also lead to gains from other sources. We will briefly discuss three of them.

1. More gains from large-scale production. International trade makes it possible for both domestic producers and consumers to derive larger gains from the lower per-unit costs that often accompany large-scale production, marketing, and distribution activities. When economies of scale are important in an industry, successful domestic firms will be able to produce larger outputs and achieve lower unit costs than they would if they were unable to sell their products internationally. This is particularly important for firms located in small countries. For example, textile manufacturers in Malaysia, Taiwan, and South Korea would face much higher per-unit costs if they could not sell abroad because the domestic markets of these countries are too small to support large-scale production. There simply aren't enough buyers. However, if the firms can access the world market, where there are many more buyers, they can operate on a large scale and compete quite effectively.

Domestic consumers also benefit because international trade often makes it possible for them to acquire goods at lower prices from large-scale producers in other countries. The aircraft industry vividly illustrates this point. Given the huge design and engineering costs it takes to produce a single jet, no firm would be able to produce them economically if it weren't able to sell them abroad. Because of international trade, however, consumers around the world are able to purchase planes economically from large-scale producers like Boeing, which is based in the United States.

2. Gains from more competitive markets. International trade promotes competition and encourages production efficiency and innovation. Competition from abroad keeps domestic producers on their toes and gives them a strong incentive to improve the quality of their products. The experience of the U.S. auto industry illustrates this point. Faced with stiff competition from Japanese firms during the 1980s, U.S. auto makers worked hard to improve the quality of their vehicles. As a result, the reliability of the automobiles and light trucks available to American consumers—including those produced by domestic manufacturers—is now much better than it used to be.

International trade also allows technologies and innovative ideas developed in one country to be disseminated to others. In many cases, local entrepreneurs will emulate production procedures and products that have been successful in other places and even further improve or adapt them for local markets. Dynamic competition of this type is an important source of growth and prosperity, particularly for less-developed countries.

3. More pressure to adopt sound institutions. Not only do firms in open economies face more intense competition, so, too, do their governments. The gains from trade and the prosperity that results from free trade motivate political officials to establish sound institutions and adopt constructive policies. If they do not, both labor and capital will move toward more favorable environments. For example, neither domestic nor foreign

investors will want to put their funds in countries characterized by hostile business conditions, monetary instability, legal uncertainty, high taxes, and inferior public services. When labor and capital are free to move elsewhere, implementing government policies that penalize success and undermine productive activities becomes more costly. This aspect of free trade is generally overlooked, but it may well be one of its most beneficial attributes.

SUPPLY, DEMAND, AND INTERNATIONAL TRADE

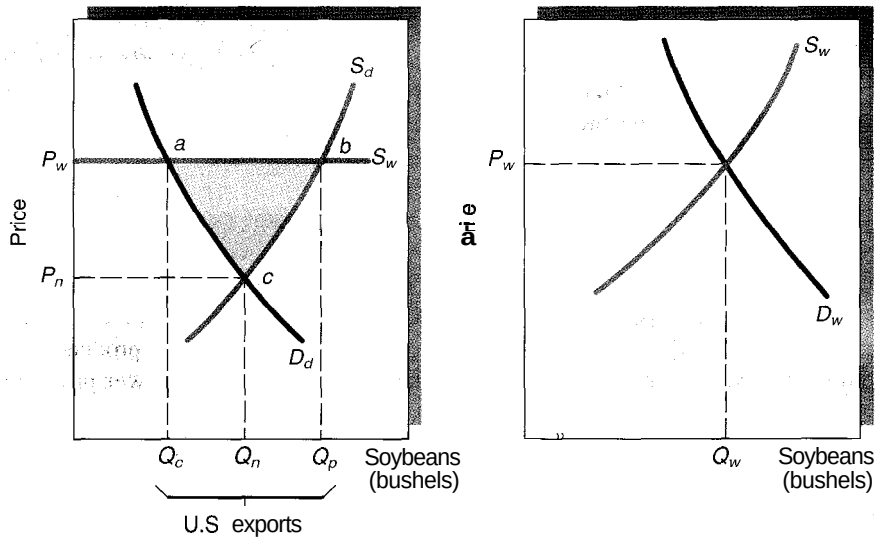
Like other things, international trade can be analyzed within the supply and demand framework. An analysis of supply and demand in international markets can show us how trade influences prices and output in domestic markets.

Consider the market for a good that U.S. producers are able to supply at a low cost. Using soybeans as an example, **Exhibit 6** illustrates the relationship between the domestic and world markets. The price of soybeans is determined by the forces of supply and demand in the world market. In an open economy, domestic producers are free to sell and domestic consumers are free to buy the product at the world market price (P_w). At this price, U.S. producers will supply Q, and U.S. consumers will purchase Q,. Reflecting their low cost (comparative advantage), U.S. soybean producers will export Q_p – Q, units at the world market price.

Let's compare this open-economy outcome with the outcome that would occur in the absence of trade. If U.S. producers were not allowed to export soybeans, the domestic price would be determined by the domestic supply (S,) and demand *(D,)* only. A lower "no-trade'' price *(P,)* would emerge.

EXHIBIT 6
Producer Benefits from Exports

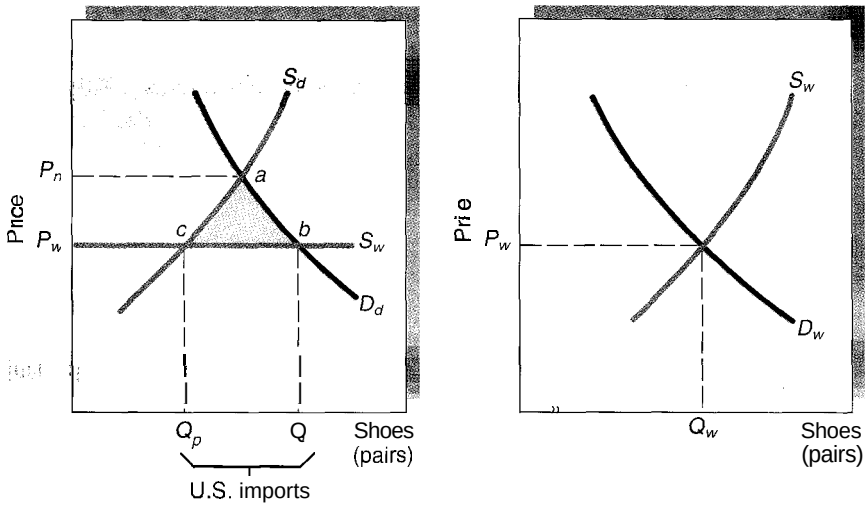
The price of soybeans and other internationally traded commodities is determined by the forces of supply and demand in the world market (b). If U.S. soybean producers are prohibited from selling to foreigners, the domestic price will be P_n (a). Free trade permits the U.S. soybean producers to sell Q_p units at the higher world price (P_w). The quantity $Q_p - Q_c$ is the amount U.S. producers export. Compared to the no-trade situation, the producers' gain from the higher price ($P_w bcP_n$) exceeds the cost imposed on domestic consumers ($P_w acP_n$) by the triangle *abc.*



(a) U.S. market for soybeans (b) World market for soybeans

EXHIBIT 7
Consumer Benefits from Imports

In the absence of trade, the domestic price of shoes would be P_n. Since many foreign producers have a comparative advantage in the production of shoes, international trade leads to lower prices. At the world price P_w, U.S. consumers will demand Q units, of which $Q_c - Q_p$ are imported. Compared to the no-trade situation, consumers gain $P_n abP_w$, while domestic producers lose $P_n acP_w$. A net gain of *abc* results.



Who are the winners and losers as the result of free trade in soybeans? Clearly, soybean producers gain. Free trade allows domestic producers to sell a larger quantity (Q_p rather than Q_n). As a result, the net revenues of soybean producers will rise by $P_w bcP_n$. On the other hand, domestic consumers of soybeans will have to pay a higher price under free trade. Soybean consumers will lose (1) because they have to pay P_w rather than P_n for the Q, units they purchase, and (2) because they lose the consumer surplus on the $Q_n - Q_c$ units now purchased at the higher price. Thus, free trade imposes a net cost of $P_w acP_n$ on consumers. As you can see in Exhibit 6, however, the gains of soybean producers outweigh the losses to the consumers by the triangle *abc*. In other words, free trade leads to a net welfare gain.

This exporting example makes it seem like free trade benefits producers relative to consumers, but this ignores the secondary effects: if foreigners do not sell goods to Americans, they will not have the purchasing power necessary to purchase goods from Americans. U.S. imports — the purchase of goods from low-cost foreign producers — provides foreigners with the dollar purchasing power necessary to buy U.S. exports. In turn, the lower prices in the import-competitive markets will benefit the U.S. consumers who appeared at first glance to be harmed by the higher prices (compared to the no-trade situation) in export markets.

Using shoes as an example, **Exhibit 7** illustrates the situation when the United States is a net importer. In the absence of trade, the price of shoes in the domestic market would be P_n, the intersection of the domestic supply and demand curves. However, the world price of shoes is P_w. In an open economy, many U.S. consumers would take advantage of the low shoe prices available from foreign producers. At the lower world price, U.S. consumers would purchase Q, units of shoes, importing Q, $- Q_p$ from foreign producers.

Compared to the no-trade situation, free trade in shoes results in lower prices and greater domestic consumption. The lower prices lead to a net consumer gain of $P_n abP_w$. Domestic producers lose $P_n acP_w$ in the form of lower sales prices and reductions in output. However, the net gain of the shoe consumers exceeds the net loss of producers by *abc*.

International competition will direct resources toward their area of comparative advantage. If domestic producers have a comparative advantage in the production of a good, they will be able to compete effectively in the world market and profit from the export of goods to foreigners. In turn, the exports will generate the purchasing power necessary to buy goods that foreigners can supply more economically.

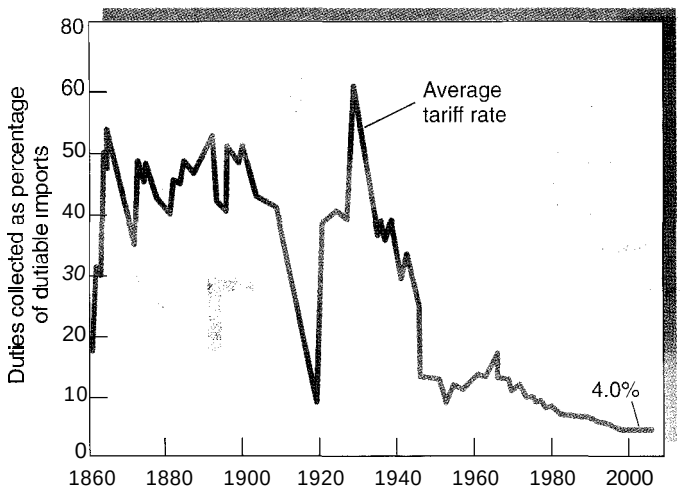


EXHIBIT 8
How High Are U.S. Tariffs?

Tariff rates in the United States fell sharply during the period from 1935 to 1950. After rising slightly during the 1950s, they have trended downward since 1960. In 2002, the average tariff rate on merchandise imports was 4.0 percent.

THE ECONOMICS OF TRADE RESTRICTIONS

Despite the potential benefits of free trade, almost all nations have erected trade barriers. Tariffs, quotas, and exchange rate controls are the most commonly used trade-restricting devices. Let's consider how various types of trade restrictions affect the economy.

Economics of Tariffs

Tariff
A tax levied on goods imported into a country.

A **tariff** is nothing more than a tax on imports from foreign countries. As **Exhibit 8** shows, average tariff rates of between 30 percent and 50 percent of product value were often levied on products imported to the United States prior to 1945. The notorious Smoot-Hawley Tariff Act of 1930 pushed the average tariff rate upward to 60 percent. Many economists believe that this legislation contributed significantly to the length and severity of the Great Depression. During the past sixty years, however, tariff rates in the United States have declined substantially. In 2002, the average tariff rate on imported goods was only 4 percent.

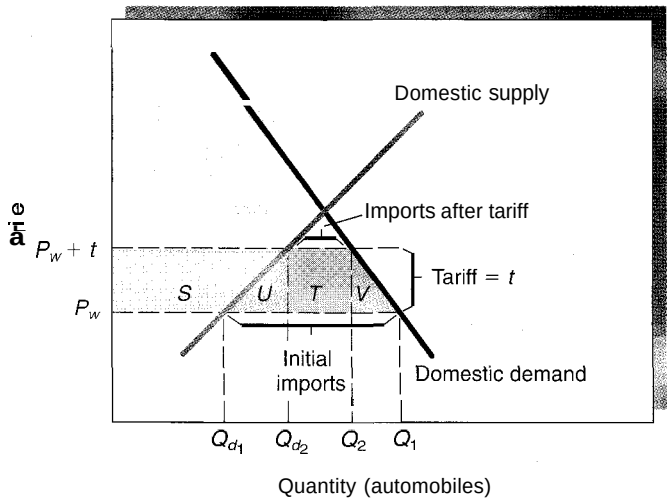
Exhibit 9 shows the impact of a tariff on automobiles. In the absence of a tariff, the world market price of P_w would prevail in the domestic market. At that price, U.S. consumers purchase Q, units. Domestic producers supply Q_{d_1}, while foreigners supply Q, $- Q_{d_1}$ units to the U.S. market. When the United States levies a tariff, *t,* on automobiles, Americans can no longer buy cars at the world price. U.S. consumers now have to pay $P_w + t$ to purchase an automobile from foreigners. At that price, domestic consumers demand Q_2 units (Q_{d_2} supplied by domestic producers and $Q_2 - Q_{d_2}$ supplied by foreigners). The tariff results in a higher domestic price and lower level of domestic consumption.

The tariff benefits domestic producers and the government at the expense of consumers. Because domestic producers don't have to pay the tariff, they will expand their output in response to the higher (protected) market price. In effect, the tariff acts as a subsidy to domestic producers. Domestic producers gain the area *S* (Exhibit 9) in the form of additional net revenues. The tariff raises revenues equal to the area *T* for the government. The areas *U* and V represent costs imposed on consumers and resource suppliers that do not benefit the government. Simply put, *U* and V represent *deadweight losses:* consumer and producer surpluses that could have been gained if the tariff hadn't been imposed.

As a result of the tariff, resources that could have been used to produce other U.S. goods more efficiently (compared to producing them abroad) are diverted to automobile production. Ultimately, we end up producing fewer products in areas where we have a comparative advantage and more products in areas where we are a high-cost producer. Because of this, some of the gains from specialization and trade go unrealized.

EXHIBIT 9
The Impact of a Tariff

Here we illustrate the impact of a tariff on automobiles. In the absence of the tariff, the world price of automobiles is P_w: U.S. consumers purchase Q_1 units (Q_{d_1} from domestic producers plus $Q_1 - Q$ from foreign producers). The tariff makes it more costly for Americans to purchase automobiles from foreigners. Imports decline and the domestic price increases. Higher prices reduce consumer surplus by the areas $S + U + T + V$. Producers gain the area S, and the tariff generates T tax revenues for the government. The areas U and Vare deadweight losses. Consumers lose the surplus associated with these two areas, but producers and the government don't gain it.



The Economics of Quotas

Import quota
A specific limit or maximum quantity (or value) of a good permitted to be imported into a country during a given period.

An **import quota,** like a tariff, is designed to restrict foreign goods and protect domestic industries. A quota places a ceiling on the amount of a product that can be imported during a given period (typically a year). The United States imposes quotas on several products, including brooms, shoes, sugar, dairy products, and peanuts. For example, since 1953, the United States has imposed an annual peanut quota of 1.7 million pounds. That's equivalent to just two imported peanuts per American. Like tariffs, the primary purpose of quotas is to protect domestic industries from foreign competition.

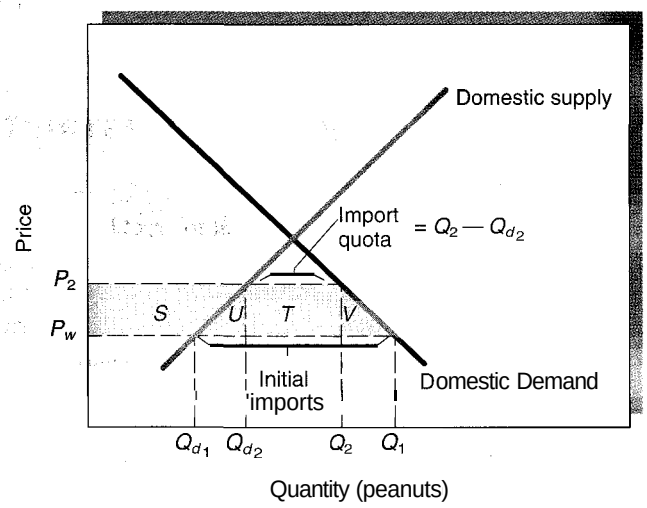
Using peanuts as an example, **Exhibit 10** illustrates the impact of a quota. If there were no trade restraints, the domestic price of peanuts would be equal to the world market price (P_w). Under those circumstances, Americans would purchase Q, units. At the price P_w, domestic producers would supply Q_{d_1}, and the amount Q, – Q, would be imported from foreign producers.

Now consider what happens when a quota limits imports to Q, – Q_{d_2}, a quantity well below the free-trade level of imports. Since the quota reduces the foreign supply of peanuts to the domestic market, the price of the quota-protected product increases (to P_2). At the higher price, U.S. consumers will reduce their purchases to Q_2, and domestic producers will happily expand their production to Q_{d_2}. With regard to the welfare of consumers, the impact of a quota is similar to that of a tariff. Consumers lose the area $S + U + T + V$ in the form of higher prices and the loss of consumer surplus. Similarly, domestic producers gain the area S, while the areas U and V represent deadweight losses in the form of reductions in consumer surplus, gains that buyers would have derived in the absence of the quota.

While the adverse impact of a quota on consumer welfare is similar to that of a tariff, there is a big difference with regard to the area T. Under a tariff, the U.S. government would collect revenues equal to T, representing the tariff rate multiplied by the number of units imported. With a quota, however, these revenues will go to foreign producers, who

EXHIBIT 10
The Impact of a Quota

Here we illustrate the impact of a quota, such as the one the United States imposes on peanuts. The world market price of peanuts is P_w. If there were no trade restraints, the domestic price would also be P_w, and the domestic consumption would be Q_1. Domestic producers would supply Q_{d_1} units, while $Q_1 - Q_{d_1}$ would be imported. A quota limiting imports to $Q_1 - Q_{d_2}$ would push up the domestic price to P_2. At the higher price, the amount supplied by domestic producers increases to Q_{d_2}. Consumers lose the sum of the area $S + U + T + V$, while domestic producers gain the area S. In contrast with tariffs, quotas generate no revenue for the government. The area T goes to foreign producers, who are granted permission to sell in the U.S. market.



are granted licenses (quotas) to sell various amounts in the U.S. market. Clearly, this right to sell at a premium price (since the domestic price exceeds the world market price) is extremely valuable. Thus, foreign producers will compete for the permits. They will hire lobbyists, make political contributions, and engage in other rent-seeking activities in an effort to secure the right to sell at a premium price in the U.S. market.

In many ways, quotas are more harmful than tariffs. With a quota, foreign producers are prohibited from selling additional units regardless of how much lower their costs are relative to those of domestic producers. In contrast to a tariff, a quota brings in no revenue for the government. While a tariff transfers revenue from U.S. consumers to the Treasury, quotas transfer these revenues to foreign producers. Rewarding domestic producers with higher prices and foreign producers with valuable import permits will create *two* strong interest groups. Both groups will lobby hard to keep the quota in place. As a result, lifting the quota will often be more difficult than lowering a tariff would be.

In addition to tariffs and quotas, governments sometimes use regulations and political pressure to restrain foreign competition. For example, the United States prohibits foreign airlines from competing in the domestic air travel market. Japanese regulations make it illegal for domestic automobile dealers to sell both foreign and domestically produced vehicles; this makes it more difficult for foreign manufacturers to establish the dealer networks they need to effectively penetrate the Japanese market. Like tariffs and quotas, regulatory barriers such as these reduce the supply to domestic markets and the gains from potential trades. Overall output is reduced, and domestic producers benefit at the expense of domestic consumers.

Exchange Rate Controls as a Trade Restriction

Some countries fix the exchange rate value of their currency above the market rate and impose restrictions on exchange rate transactions.[2] At the official (artificially high) exchange rate, the country's export goods will be extremely expensive to foreigners. As a result, foreigners will purchase goods elsewhere, and the country's exports will be small. In turn, the low level of exports will make it extremely difficult for domestic residents to obtain the foreign currency they need to purchase imports. Exchange rate controls both reduce the volume of trade and lead to black-market currency exchanges. Indeed, a large black-market premium indicates that the country's exchange rate policy is substantially limiting the ability of its citizens to trade with foreigners. While exchange rate controls have declined in popularity, they are still an important trade barrier in countries such as Myanmar and Zimbabwe.

WHY DO NATIONS ADOPT TRADE RESTRICTIONS?

As social philosopher Henry George noted over a century ago, trade restraints act like blockades. Why would political officials want to erect blockades against their own people? As we consider this question, we will take a look at three arguments often raised by the proponents of trade restrictions: the national-defense, infant-industry, and antidumping arguments. Finally, we will look at the politics of trade restrictions and analyze how the nature of the restraints influences their political popularity.

The National-Defense Argument

Protective tariffs are as much applications of force as are blockading squadrons, and their objective is the same—to prevent trade. The difference between the two is that blockading squadrons are a means whereby nations seek to prevent their enemies from trading; protective tariffs are a means whereby nations attempt to prevent their own people from trading.

—Henry George[3]

According to the national-defense argument, certain industries—aircraft, petroleum, and weapons, for example—are vital to a nation's defense. Therefore, these industries and their inputs should be protected from foreign competitors so that a domestic supply of necessary materials would be available in case of an international conflict. Would we want to be entirely dependent on Arabian or Russian petroleum? Would complete dependence on French aircraft be wise? Many Americans would answer "no," even if it meant imposing trade restrictions that would lead to higher prices on products they buy.

Although the national-defense argument has some validity, it is often abused. Relatively few industries are truly vital to our national defense. If a resource is important for national defense, often it would make more sense to stockpile the resource during peacetime rather than follow protectionist policies to preserve a domestic industry. Furthermore, fostering an economy robust enough to produce the mass quantity of goods necessary to sustain a war effort in the first place is, itself, part of a strong defense.

The Infant-Industry Argument

Infant-industry advocates believe that new domestic industries should be protected from foreign competition for a period of time so that they will have a chance to develop. As the new industry matures, it will be able to stand on its own feet and compete effectively with foreign producers, at which time the protection can be removed.

The infant-industry argument has a long and often notorious history. Alexander Hamilton used it to argue for the protection of early U.S. manufacturing. The major problem with the argument is that the protection, once granted, will be difficult to remove. For example, a century ago, this argument was used to gain tariff protection for the newly

[2]The most common exchange rate restriction is that individuals are required to obtain approval from the government before they engage in transactions involving foreign currency.

[3]Henry George, *Protection or Free Trade* (Washington, *D.C.*: U.S. Government Printing Office, 1886), 37.

APPLICATIONS IN ECONOMICS

Do More Open Economies Perform Better?

Economic theory indicates that more open economies will perform better than those with sizeable trade restrictions. Is this really true? In order to address this question, a measure of trade openness—the freedom of individuals to engage in voluntary exchange across national boundaries—is needed. Economist Charles Skipton, in a recent research project, developed a trade openness index (TOI) for eighty-one countries during the 1980–1999 period.[1] To achieve a high rating on the zero to ten TOI scale (with ten indicating more openness to free trade), a country had to maintain low tariff rates and a freely convertible currency (no exchange rate controls) as well as refrain from imposing quotas and other regulations reducing the size of its trade sector.

Exhibit 11 shows the ten countries with the highest and lowest trade openness ratings. The ratings reflect the *average* degree of openness for the entire 1980–1999 period. This is important because the gains from increased openness can only be realized over time. Expanding the openness of trade is a long-term growth strategy, not a short-term "quick fix," in other words. Hong Kong, Singapore, Belgium, Canada, and the Netherlands head the list of the most open of the eighty-one economies. By way of comparison, the United States ranked sixteenth. At the other end of the spectrum, the TOI indicates that Bangladesh, Burundi, Madagascar, Pakistan, and India were the least open economies during the period.

As Exhibit **11** shows, the 2002 average GDP per person of $26,518 of the ten most open economies was more than eight times the comparable figure for the ten least

(continued)

EXHIBIT 11
Trade Openness, Income, and Growth

	TOI	2002 GDP Per Capita (1995 Dollars)	Growth Rate, 1980–2002
10 Most Open Economies, 1980–1999			
Singapore	9.9	$21,296	4.2
Hong Kong	9.9	23,833	3.7
Belgium	9.2	24,418	1.7
Canada	8.9	26,114	1.6
Netherlands	8.9	25,778	1.8
Luxembourg	8.7	54,201	4.0
Malaysia	8.4	8,080	3.4
Norway	8.4	32,414	2.5
Australia	8.2	25,880	2.0
United Kingdom	8.1	23,166	2.2
Average	**8.9**	**$26,518**	**2.7**
Ten Least Open Economies, 1980–1999			
Algeria	3.4	$5,101	−0.0
Belize	3.4	5,386	2.1
Tanzania	3.3	513	1.0
Argentina	3.2	9,633	−0.6
Syria	3.1	3,205	0.7
India	2.8	2,365	3.6
Pakistan	2.8	1,719	2.3
Madagascar	2.8	659	−2.1
Burundi	1.4	561	−0.9
Bangladesh	0.9	1,501	2.1
Average	**2.7**	**$3,064**	**0.8**

Source: The trade openness index data are from Charles Skipton, *The Measurement of Trade Openness.* Doctoral Dissertation, Florida State University, 2003 (Table 4.6). The per capita GDP and growth Bank, World Development Indicators, CD-ROM, 2004. The per capita income figures were divided by the purchasing power parity method.

APPLICATIONS IN ECONOMICS

(continued)

open economies. Moreover, the more open economies also grew more rapidly. During 1980-2002, real GDP per person in the ten most open economies expanded at an annual rate of 2.7 percent, compared to 0.8 percent in the ten least open economies. The per capita GDP in every one of the most open economies grew at least 1.6 percent annually. Of course, the data of Exhibit 11 do not take into account other cross-country differences that theory indicates will influence growth. However, Skipton found that even after the differences in countries' inflation rates, legal structures, and similar factors were taken into account, trade openness continued to have a strong positive impact on both per capita GDP and growth rates.[2]

[1]Charles Skipton, "The Measurement of Trade Openness" (doctoral dissertation, Flonda State University 2001)

[2]For additional information on the relationship between international trdde dnd economic growth, see Jeffrey A Frankel and David Romer, 'Doe? Trade Cause Growth?" *American Economic Review* (June 1999) 379–399, dnd Jeffrey D Sachs dnd Andrew Warner, ' Economic Reform and the Process of Global Integration " *Brookings Papers on Economic Activity,* no. 1 (1995): 1–95.

emerging steel industry in the United States. Over time, the steel industry developed and became very powerful, both politically and economically. Despite its maturity, the tariffs remained. To this day, legislation continues to provide the steel industry with various protections that limit competition from abroad.

The Antidumping Argument

Dumping
Selling a good in a foreign country at a lower price than it's sold for in the domestic market.

Dumping involves the sale of goods by a foreign firm at a price below cost or below the price charged in the firm's home-base market. Dumping is illegal and if a domestic industry is harmed, current law provides relief in the form of antidumping duties (tariffs imposed against violators). In addition, under the recently enacted Byrd Amendment, the revenues collected from the antidumping duties are transferred to the firms and unions lodging antidumping complaints, further increasing their incentive to levy such charges.[4]

Proponents of the antidumping argument argue that foreign producers will temporarily cut prices, drive domestic firms out of the market, and then use their monopoly position to gouge consumers. However, there is reason to question the effectiveness of this strategy. After all, the high prices would soon attract competitors, including other foreign suppliers.

Antidumping cases nearly always involve considerable ambiguity. The prices charged in the home market generally vary, and the production costs of the firms charged with dumping are not directly observable. This makes it difficult to tell whether a dumping violation has really occurred. Furthermore, aggressive price competition is an integral part of the competitive process. When demand is weak and inventories are large, firms will lower the prices of their products below their average total cost of production. Domestic firms are permitted to engage in this practice, and consumers benefit from it. Why shouldn't foreign firms be allowed to do the same?

One thing is for sure: antidumping legislation gives politicians another way to channel highly visible benefits to powerful business and labor interests — another open invitation for rent seeking. Moreover, the dumping charges are adjudicated by International Trade Commission and U.S. Department of Commerce officials. Consequently, it's naive to believe that political considerations won't be an important element underlying the charges that are levied and how they are resolved. Unsurprisingly, the number of claimants bringing charges of dumping has increased substantially in recent years.

[4]In effect, the Byrd Amendment provides subsidies to firms and unions willing to lend their support to antidumping charges and places them at a competitive advantage relative to those who fail to support the petitions. Thus, it encourages the filing of antidumping charges. In August 2004, the World Trade Organization ruled that the Byrd Amendment was a violation of international trade rules and authorized Japan, Europe, Korea, and several other countries to levy specific levels of retaliation duties against U.S. products.

Special Interests and the Politics of Trade Restrictions

Regardless of the arguments made by the proponents of trade restrictions, in truth, the restrictions are primarily special-interest related. (See the quotation to the right from Professor Weidenbaum.) ***Trade restrictions typically provide highly visible, concentrated benefits for a small group of people, while imposing on the general citizenry costs that are widely dispersed and difficult to identify.*** **As** we discussed in Chapter 6, the political process handles such issues poorly. It often leads to their adoption, even when they lower income levels and living standards.

Protectionism is a politician's delight because it delivers visible benefits to the protected parties while imposing the costs as a hidden tax on the public.

—Murray L. Weidenbaum[5]

The politics of trade restrictions are straightforward and play out over and over again. Well-organized business and labor interests gain substantially from restrictions that limit competition from abroad. Because their personal gain is large, they will feel strongly about the issue and generally vote for or against candidates on the basis of their positions on trade restriction. Most important, the special-interest groups will be an attractive source of political contributions. When it comes to consumers, on the other hand, even if the total cost of the restrictions is quite large, it will be spread thinly among them; most consumers will be unaware that they are paying slightly higher prices for various goods because of the restrictions.

As you can see, courting special-interest groups helps politicians solicit campaign contributions and generate votes. On the other hand, little political gain can be derived from poorly organized and largely uninformed consumers. Given this incentive structure, the adoption of trade restrictions is not surprising.

The U.S. tariff code itself is a reflection of the politics of trade restrictions. It is both lengthy (the schedule fills 3,825 pages) and highly complex. This makes it difficult for even a well-educated citizen to figure out how it works. High tariffs are imposed on some products (for example, apparel, tobacco, and footwear), while low tariffs are imposed on others. Highly restrictive quotas limit the import of a few commodities, most notably agricultural products. Even though this complex system of targeted trade restrictions is costly to administer, it is no accident. It reflects the rent seeking of special-interest groups and the political side payments, particularly campaign contributions, made to politicians.

TRADE BARRIERS AND POPULAR TRADE FALLACIES

Fallacies abound in the area of international trade. Why? Failure to consider the secondary effects of international trade is part of the answer. Key elements of international trade are closely linked; you cannot change one element without changing the other. For example, you cannot reduce imports without simultaneously reducing the demand for exports. The political incentive structure is also a contributing factor. As business, labor, and political leaders seek to gain from trade restrictions, they will often use half-truths and wrong-headed ideas to achieve their political objectives. Two of the most popular trade fallacies involve the effects of imports on employment and the impact of trade with low-wage countries. Let's take a closer look at both.

Trade Fallacy 1: Trade restrictions that limit imports save jobs and expand employment. Like most fallacies, this one has just enough truth to give it some credibility. When tariffs, quotas, and other trade barriers limit imports, they are likely to foster employment in the industries shielded from competition. But this is only half of the story: simultaneously, jobs in other domestic sectors will be destroyed. Here's how: When trade barriers reduce the amount of goods Americans buy from foreigners, sales to foreigners will also fall. This is because our imports provide foreigners with the dollars they need to buy our exports. Because foreigners cut back on the items they would normally buy from us, other U.S. sectors will suffer job losses because they're selling less.

[5]Murray L Weidenbaum, personal correspondence with the authors Professor Weidenbaum is a former chairman of the President's Council of Economic Advisers and longtime director of the Center for the Study of American Business of Washington University

Furthermore, when trade restrictions are imposed on a resource domestic producers use as an input, they will have to pay a higher price for it than their foreign rivals. This will increase their costs and make it more difficult for them to compete internationally. As a result, they will have to lay off some of their employees. The import quotas imposed on steel during 2002–2004 vividly illustrate this point. The quotas helped the domestic steel industry, but they virtually wiped out the domestic industry producing steel barrels, a product the U.S. had exported prior to the quota being imposed. The quota also increased costs and reduced the competitiveness of industries that were major users of steel, like the automobile- and appliance-manufacturing industries. Employment in those industries fell as well. The same phenomenon occurred after the United States imposed sugar quotas. The import quotas pushed domestic sugar prices to two or three times the world price. As a result, several large candy makers relocated abroad so that they could buy sugar at the lower world price. Again, the jobs lost in U.S. industries using sugar were offset by any increase in employment by U.S. sugar producers.

On balance, there is no reason to expect that trade restrictions will either create or destroy jobs. Instead, they will reshuffle them. The restrictions artificially direct workers and other resources toward the production of things that we do poorly, as shown by our inability to compete effectively in the world market. Simultaneously, employment will decline in areas where American firms would be able to compete successfully in the world market if it were not for the side effects of the restrictions. In other words, more Americans will be employed producing things we do poorly and fewer will be employed producing things we do well. As a result, our overall income level will be lower than it would have been otherwise.

Unfortunately, the jobs "saved' by the import quotas are more visible than those destroyed in other sectors. This increases the political popularity of trade restraints and perpetuates the fallacy that the restraints increase employment. But it does not change the reality of the situation. As Exhibit 1 shows, imports increased from 6 percent of GDP in 1980 to 15 percent in 2003. If the growth of imports destroys jobs, as the proponents of trade restrictions argue, the rapid import growth should have adversely affected U.S. employment. But this was not the case. On the contrary, civilian employment in the United States rose from 99 million in 1980 to 119 million in 1990 and 137 million in 2003. Far from retarding employment, the unprecedented growth of imports during the last two decades was associated with unprecedented employment growth.

Trade Fallacy 2: Free trade with low-wage countries like Mexico and China will reduce the wages of Americans. Many Americans believe that, without trade restrictions, their wages will fall to the wage levels of workers in poor countries. How can Americans compete with workers in countries like Mexico and China who are willing to work for $1 or less per hour? This fallacy stems from a misunderstanding of both the source of high wages and the law of comparative advantage. Workers in the United States are well educated, possess high skill levels, and work with large amounts of capital equipment. These factors contribute to their high productivity, which is the source of their high wages. Similarly, in countries like Mexico and China, wages are low precisely because productivity is low. Workers are less skilled in these countries, and there is less capital equipment to make them more productive.

The key thing to remember, though, is that gains from trade emanate from comparative advantage, not absolute advantage (see Exhibits 3, 4, and 5). The United States cannot produce *everything* more cheaply than Mexico or China merely because U.S. employees are more productive and work with more capital. Neither can the Mexicans and Chinese produce *everything* more cheaply merely because their wage rates are low compared to those of U.S. workers.

As long as there are differences between countries when it comes to their comparative advantages, gains from trade will be possible, no matter what the wages of the employees in the two countries are. Trade reflects relative advantage, not wage levels. We can illustrate this point using trade between individuals. No one argues that trade between doctors and lawn service workers, for example, will cause the wages of doctors to fall. Because of their different skills and costs of providing alternative goods, both high-wage

doctors and low-wage lawn care workers can gain from trade. The same is also true for trade between rich and poor nations.

If foreigners (including low-wage foreigners) have a comparative advantage and can sell us a product for less than we ourselves can produce it, we can gain by buying it. This will give us more resources to invest in and produce other things. Perhaps an extreme example will illustrate this point. Suppose a foreign producer is willing to supply us automobiles free of charge (perhaps because its employees were willing to work for nothing). Would it make sense to impose tariffs or quotas to keep the automobiles from coming into the country? Of course not. Resources that were previously used to produce automobiles would then be freed up to produce other goods, and the real income and availability of goods would expand. It makes no more sense to erect trade barriers to keep out cheap foreign goods than it would to keep out the free autos.

THE CHANGING NATURE OF GLOBAL TRADE

Since World War II, there has been a gradual reduction in tariff rates and other trade barriers. Liberalized trade policies and lower transportation and communication costs have propelled the growth of international trade. The growth of trade—some might say the globalization of the economy—has also resulted in a changing institutional environment. This section will focus on the institutions of international trade and the prospects for future trade liberalization.

GATT and the WTO

Following World War II, the major industrial nations of the world established the **General Agreement on Tariffs and Trade (GATT).** For almost five decades, GATT played a central role in reducing tariffs and relaxing quotas. The average tariff rates of GATT members fell from approximately 40 percent in 1947 to less than 5 percent in 1998, for example.

General Agreement on Tariffs and Trade (GATT)
An organization formed after the Second World War to set the rules for the conduct of international trade and reduce trade barriers among nations.

Following 1993, GATT was given a new name: the **World Trade Organization (WTO).** This organization of almost 150 countries is now responsible for monitoring and enforcing the trade agreements developed through GATT. The WTO gives member nations a forum in which to discuss trade rules, and it settles trade disputes among them.

World Trade Organization (WTO)
The new name given to GATT in 1994; the WTO is currently responsible for monitoring and enforcing multilateral trade agreements among its 133 member countries.

NAFTA and Other Regional Trade Agreements

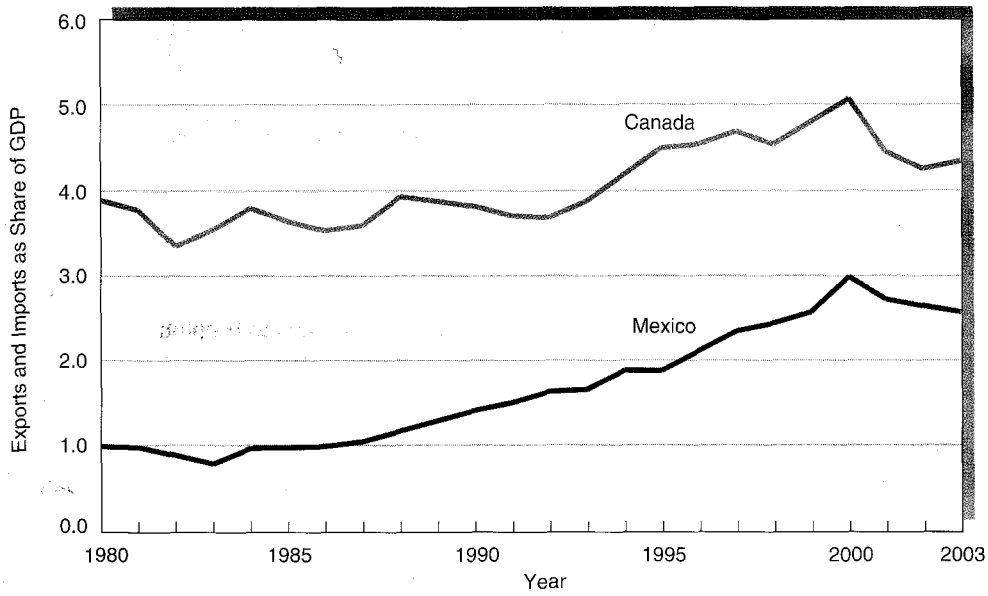
Canada has been a major trading partner of the United States for many decades. On the other hand, U.S. trade with Mexico was small prior to the 1990s. Historically, Mexico has been a relatively closed economy. This began to change in the mid-1980s, when Mexico began cutting its tariff rates and unilaterally removing other trade barriers. In 1988, the United States and Canada negotiated a trade agreement designed to reduce barriers limiting both trade and the flow of capital between the two countries. A few years later, the United States, Canada, and Mexico finalized the **North American Free Trade Agreement (NAFTA),** which took effect in 1994. As the result of NAFTA, the tariffs of most goods moving among the three countries have now been eliminated. The agreement will eventually remove restrictions on financial investments, liberalize trade in services like banking, and establish uniform legal requirements for the protection of intellectual property. In addition to its participation in NAFTA, Mexico has also adopted a free-trade agreement with the European Union. During the last fifteen years, Mexico has moved from one of the world's more protectionist countries to one of its more open economies.

North American Free Trade Agreement (NAFTA)
A comprehensive trade agreement between the United States, Mexico, and Canada that went into effect in 1994. Under the agreement, tariff barriers were to continue to be phased out until 2004.

As **Exhibit 12** shows, U.S. trade with both Mexico and Canada has grown rapidly in recent years. Measured as a share of GDP, trade with Mexico jumped from 1.4 percent in 1990 to 2.6 percent in 2003. During the same period, trade with Canada rose from 3.8 percent of GDP to 4.3 percent. This growth of trade, particularly with Mexico, has not been without controversy. Business and labor groups often blame employment contractions and plant closings on competition with Mexican firms. The news media generally give such stories ample exposure. However, there is no evidence that increased trade with Mexico has adversely affected the U.S. economy. During the last decade, the growth rate of the

EXHIBIT 12
U.S. Trade with Canada and Mexico, 1980–2003

Measured as a share of GDP, U.S. trade with both Canada and Mexico has increased sharply during the last fifteen years.



Source: Statistical Abstract of the United States (various years) and http://www.bea.gov.

U.S. has been strong and the unemployment rate relatively low. Clearly, the dire predictions about the "jobs going to Mexico" have not been realized.

The Future of Free Trade

For several decades following World War II, the United States and most other high-income countries were leaders among those pursuing and promoting more liberal trade policies. In contrast, India, China, and most of the less-developed economies of Africa and Latin America imposed sizeable trade restraints, and they were reluctant to relax them.

Since 1980, the situation has changed dramatically. Observing the success of open economies like Hong Kong and Singapore, many less-developed countries unilaterally reduced many of their trade restrictions during the last two decades. On average, the tariff rates of less-developed countries are now less than half their levels in the early 1980s. Exchange rate controls are becoming increasingly rare, and capital market controls are much less restrictive than they were a decade ago. Today, many leaders in less-developed countries recognize that free trade is the surest route to higher income levels and improved living standards. These countries are often the fiercest advocates of trade liberalization.

By contrast, the United States, Japan, and European Union nations have agricultural price support programs contrary to free trade. It will take considerable effort to reduce, let alone remove, the price supports and subsidies to agriculture interests. To date, these countries have been unwilling to do so, and their resistance has become a major stumbling block on the road to trade liberalization. Furthermore, protectionist proponents—particularly those in high-income countries like the Unites States—have successfully lobbied to impose labor and environmental regulations that block trade liberalization. Meanwhile, the Internet and other technological changes continue to reduce transport and communications costs, and thereby encourage the movement of goods, ideas, and people across national boundaries. All of this promises to enliven trade issues in the years ahead.

LOOKING AHEAD

There are many similarities between domestic trade and trade across national boundaries, but there is also a major difference: international trade generally involves exchanging foreign currencies. The next chapter deals with the foreign exchange market and other dimensions of international finance.

KEY POINTS

- The volume of international trade has grown rapidly in recent decades. In the United States, international trade (imports plus exports) summed to 25 percent of GDP in 2003, compared to 12 percent in 1980 and 6 percent in 1960.
- Comparative advantage rather than absolute advantage is the source of gains from trade. As long as relative production costs of goods differ, trading partners will be able to gain from trade. Specialization and trade make it possible for trading partners to produce a larger joint output and expand their consumption possibilities.
- Exports and imports are linked. The exports of a nation are the primary source of purchasing power used to purchase imported goods. When a nation restricts imports, it simultaneously limits the ability of foreigners to acquire the purchasing power necessary to buy its exports.
- International specialization and trade result in lower prices for imported products and higher domestic prices for exported products. However, the net effect is an expansion in the aggregate output and consumption possibilities of trading nations.
- Import restrictions, such as tariffs and quotas, reduce the supply of foreign goods to domestic markets. This causes domestic price to rise. Essentially, the restrictions are a subsidy to producers (and workers) in protected industries at the expense of (a) consumers and (b) producers (and workers) in export industries. Jobs protected by import restrictions are offset by jobs destroyed in export-related industries.
- Trade restrictions generally provide concentrated benefits to the producers in industries they're designed to protect. The costs are spread thinly among consumers in the form of higher prices. Even though the impact of trade restrictions on the economy as a whole is harmful, they help politicians reward special-interest groups for campaign contributions and support. This makes them attractive to politicians.
- Persistently open economies have grown more rapidly and achieved higher per capita income levels than economies more closed to international trade.

CRITICAL ANALYSIS QUESTIONS

1. Why do American households and businesses buy things from foreigners? What are the characteristics of the items we buy from foreigners? What are the characteristics of the things we sell to foreigners?

*2. "Trade restrictions limiting the sale of cheap foreign goods in the United States are necessary to protect the prosperity of Americans." Evaluate this statement made by an American politician.

3. Suppose as the result of the Civil War that the United States had been divided into two countries and that, through the years, high trade barriers had grown up between the two. How might the standard of living in the "divided" United States have been affected? Explain.

*4. Can both of the following statements be true? Why or why not?
   a. "Tariffs and import quotas promote economic inefficiency and reduce the real income of a nation. Economic analysis suggests that nations can gain by eliminating trade restrictions."
   b. "Economic analysis suggests that there is good reason to expect that trade restrictions will exist in the real world."

5. "The average American is hurt by imports and helped by exports." Do you agree or disagree with this statement?

*6. "An increased scarcity of a product benefits producers and harms consumers. In effect, tariffs and other trade restrictions increase the domestic scarcity of products by reducing the supply from abroad. Such policies benefit domestic producers of the restricted product at the expense of domestic consumers." Evaluate this statement.

7. Suppose that a very high tariff were placed on steel imported into the United States. How would that affect employment in the U.S. auto industry? *(Hint:* Think about how higher steel prices will impact the cost of producing automobiles.)

*8. "Getting more Americans to realize that it pays to make things in the United States is the heart of the competitiveness issue." (This is a quote from an American business magazine.)
   a. Would Americans be better off if more of them paid higher prices in order to "buy American" rather than purchase from foreigners? Would U.S. employment be higher? Explain.
   b. Would Californians be better off if they bought goods produced only in California? Would the employment in California be higher? Explain.

9. How do tariffs and quotas differ? Can you think of any reason why foreign producers might prefer a quota rather than a tariff? Explain your answer.

*10. It is often alleged that Japanese producers receive subsidies from their government permitting them to sell their products at a low price in the U.S. market. Do you think we should erect trade barriers to keep out cheap Japanese goods if the source of their low price is governmental subsidies? Why or why not?

11. In recent years, the European Union has reduced trade barriers among its members, and most EU members now use a common currency. What impact will these changes have on European economies?

*12. Does international trade cost American jobs? Does interstate trade cost your state jobs? What is the major effect of international and interstate trade?

13. "The U.S. is suffering from an excess of imports. Cheap foreign products are driving American firms out of business and leaving the U.S. economy in shambles." Evaluate this view.

*14. The United States uses an import quota to maintain the domestic price of sugar well above the world price. Analyze the impact of the quota. Use supply and demand analysis to illustrate your answer. To whom do the gains and losses of this policy accrue? How does the quota affect the efficiency of resource allocation in the United States? Why do you think Congress is supportive of this policy?

15. As U.S. trade with low-wage countries like Mexico increases, will wages in the United States be pushed down? Why or why not? Are low-wage workers in the United States hurt when there is more trade with Mexico? Discuss.

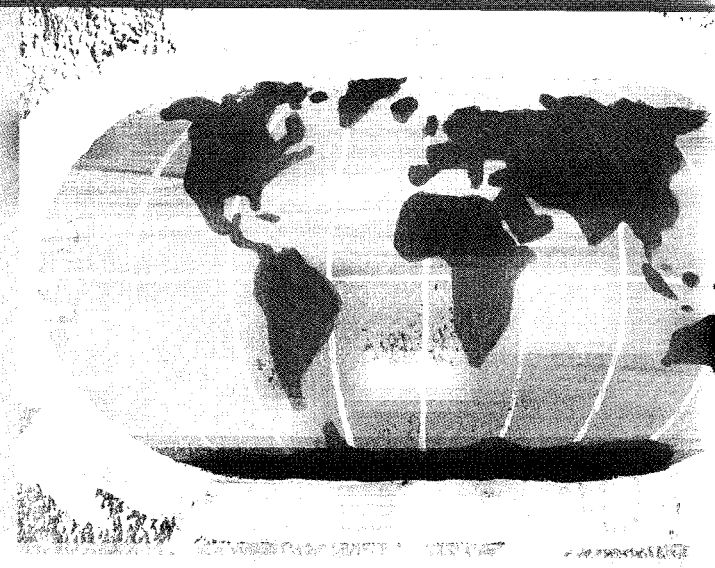
***16.** "Tariffs not only reduce the volume of imports, they also reduce the volume of exports." Is this statement true or false? Explain your answer.

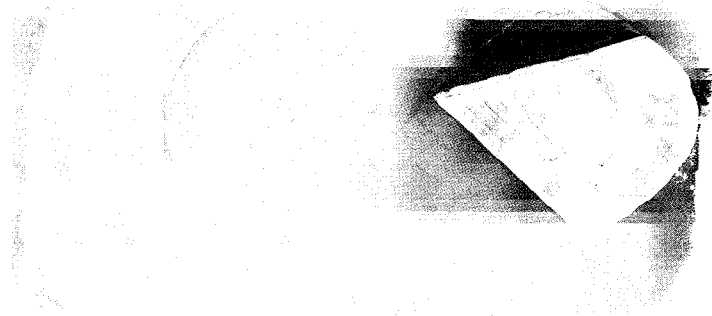
17. "Physical obstacles like bad roads and stormy weather increase transaction costs and thereby reduce the volume of trade. Tariffs, quotas, exchange rate controls, and other human-made trade restrictions have similar effects." Evaluate this statement. Is it true? Why or why not?

*Asterisk denotes questions for which answers are given in Appendix B

PART 4

Economics is about how the real world works





Applying the Basics: Special Topics in Economics

PART 4

Economics has a lot to say about current issues and real world events. What impact will the internet have on your life? Why does the current social security system face problems and what might be done to minimize them? Is ownership of stock risky? Why is the unemployment rate higher in Europe than the United States? What might be done to improve the quality of health care and education? How can we best protect the environment? This section will focus on these topics and several other current issues.

SPECIAL TOPIC 1

Government Spending and Taxation

[A] wise and frugal government, which shall restrain men from injuring one another, shall leave them otherwise free to regulate their own pursuits of industry and improvement, and shall not take from the mouth of labor the bread it has earned. This is the sum of good government. . . .

—*Thomas Jefferson*[1]

A taxpayer is someone who works for the federal government but doesn't have to take a civil service examination.

—*Ronald Reagan*

Focus

- **How has government spending per person changed historically in the United States?**
- **How has the composition of government spending changed in recent decades?**
- **Do taxes measure the cost of government?**
- **Do the rich pay their fair share *of* taxes? Do they pay a smaller share of taxes today than they did a couple of decades ago?**
- **How does the size of government in the United States compare with the size of governments of other countries?**

[1]First Inaugural Address, March 4, 1801.

In Chapters 5 and 6, we analyzed the economic role of government and the operation of the political process. We learned that while the political process and markets are alternative ways of organizing the economy, a sound legal system, secure property rights, and stable monetary regime are vitally important for the efficient operation of markets. We also noted that there may be advantages of using government to provide certain classes of goods that are difficult to supply efficiently through markets. However, as public-choice analysis indicates, the political process is not a corrective device. Even democratic representative government will often result in the misuse of resources—spending on programs that are counterproductive. This feature will take a closer look at government in the United States and will provide some additional details with regard to the characteristics of both its spending and taxation. ■

GOVERNMENT EXPENDITURES

As we noted in Chapter 6, total government spending (federal, state, and local) sums to approximately one-third of the U.S. economy. Government spending on the purchase of goods and services, including the payments made to employees, accounts for about 20 percent of the total economy, while spending on transfer payments sums to nearly 15 percent of total income. Moreover, government spending has risen rapidly during the last seventy-five years. Measured as a share of the economy, government spending rose from less than 10 percent in 1929 to nearly 35 percent in 2002. The bulk of this increase in spending has taken place at the federal level.

Approximately three-fifths of the spending by government now takes place at the federal level. Federal expenditures on just four things, (1) income transfers (including Social Security and other income security programs), (2) health care, (3) national defense, and (4) net interest on the national debt, accounted for 85 percent of federal spending in 2003. (See Chapter 6, Exhibit 2.) This means that expenditures on everything else—the federal courts, national parks, highways, education, job training, agriculture, energy, natural resources, federal law enforcement, and numerous other programs—were less than 15 percent of the federal budget. Major spending categories at the state and local level include education, public welfare and health, transportation and highways, utilities, and law enforcement.

Government Spending per Person, 1792–2003

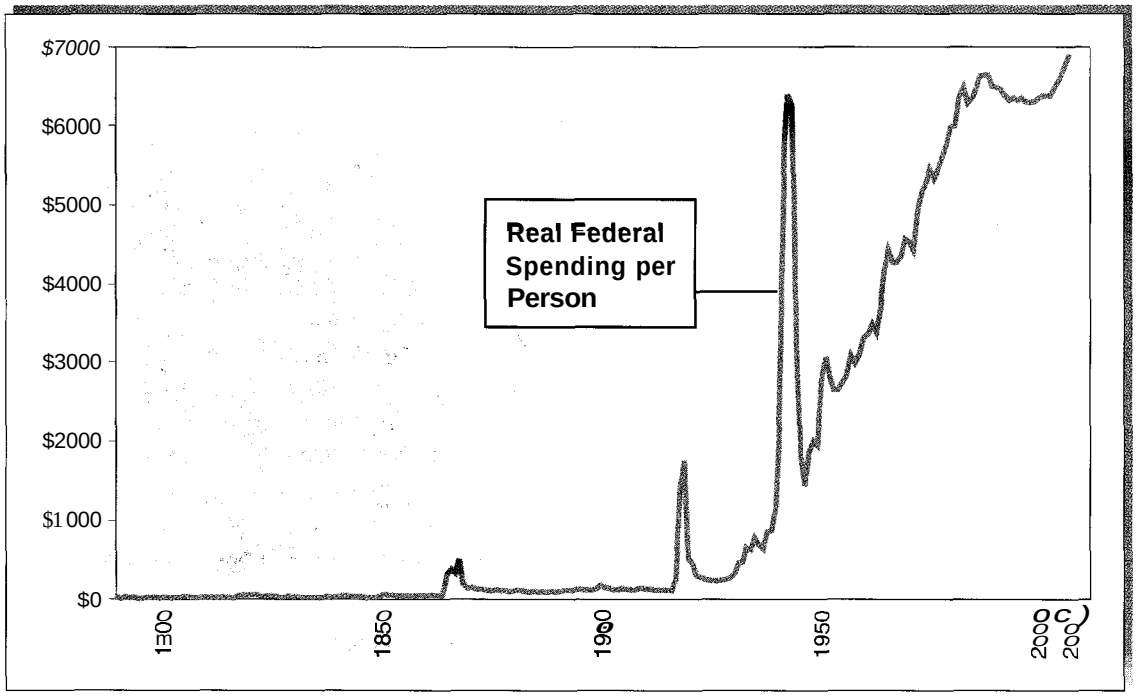
Article 1, section 8 of the U.S. Constitution outlined a limited set of functions that the federal government was authorized to perform. These included the authority to raise up an army and navy, establish a system of weights and measures, issue patents and copyrights, operate the Post Office, and regulate the value of money that it issued. Beyond this, the federal government was not authorized to do much else. The founders of the United States were skeptical of governmental powers, and they sought to limit those powers, particularly those at the federal level. (See the quotation by Thomas Jefferson at the beginning of this feature.)

During the United States' first 125 years, the constitutional limitations worked pretty much as planned; the economic role of the federal government was quite limited, and its expenditures were modest. In the nineteenth century, except during times of war, most government expenditures were undertaken at the state and local level. The federal government spent funds on national defense and transportation (roads and canals), but not much else.

Exhibit 1 presents data on real federal spending per person (measured in terms of the purchasing power of the dollar in 2000). Just prior to the Civil War, real federal expenditures were $50 per person, not much different than the $40 figure of 1800. Federal spending per person rose sharply during the Civil War, but it soon receded and remained in a range between $90 and $150 throughout the 1870–1916 period. Thus, prior to World War I, federal expenditures per person were low and the growth of government was modest.

EXHIBIT 1
Real Federal Expenditures per Capita: 1792–2003

Real federal spending per person (measured in 2000 dollars) was generally less than $50 prior to the Civil War, and it ranged from $90 to $150 throughout the 1870–1916 period. However, beginning with the spending buildup for World War I in 1917, real federal spending per person soared, reaching $6,938 in 2003—roughly sixty times the level of 1916.



Source: U.S. Census Bureau, *Historical Statistics of the United States* (Washington, D.C.: U.S. Dept. of Commerce, U.S. Bureau of the Census, 1975) and *Economic Report of the President* (Washington, D.C.: U.S. Government Printing Office, 2004).

Beginning with the World War I spending of 1917, however, the situation changed dramatically. Federal spending remained well above the prewar levels during the 1920s and rose rapidly during the 1930s. It soared during World War II, and after receding at the end of the war, federal spending continued to grow rapidly throughout the 1950–1990 period. After a brief reduction during the 1990s, per capita real federal spending is once again on the upswing. In 2003, it amounted to $6,940, roughly sixty times the $112 figure of 1916. The additional government expenditures came with a cost. On average, Americans paid more federal taxes in one week during 2003 than they paid during the entire year in 1916.

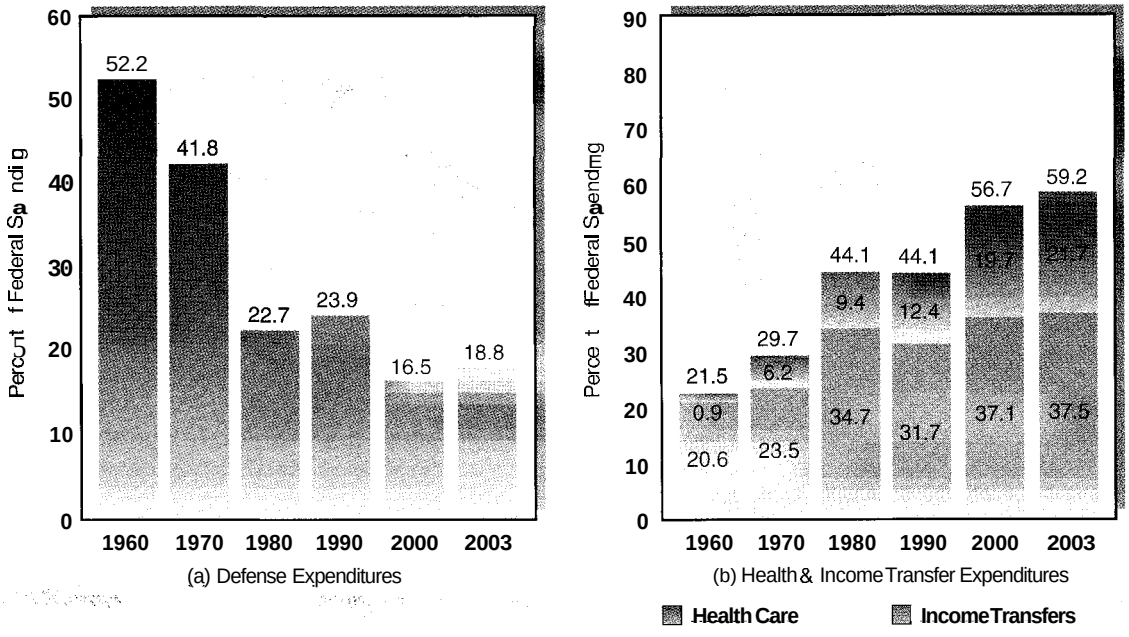
How Has the Composition of Federal Spending Changed?

Not only has federal spending grown rapidly, there has also been a dramatic shift in the composition of that spending. Since 1960, spending on defense has fallen as both a share of the budget and as a share of the economy, while expenditures on health care, transfer payments, and subsidies have soared.

As **Exhibit 2** illustrates, defense expenditures constituted more than half (52.2 percent) of federal spending in 1960. By 2000, defense spending was only 16.5 percent of the federal budget. Because of the war in Iraq, defense spending has risen slightly since 2000 to 18.8 percent of the federal budget. In contrast, government expenditures on income transfers (including Social Security, agriculture subsidies, and other income transfer programs) and health care (primarily Medicare and Medicaid) rose from 21.5 percent of the federal budget in 1960 to 59.2 percent in 2003.

EXHIBIT 2
The Changing Composition of Federal Spending

In 2003, national defense expenditures accounted for only 18.8 percent of the federal budget, down from 52.2 percent in 1960. In contrast, spending on income transfers and health care rose from 21.5 percent of the federal budget in 1960 to 59.2 percent in 2003.



Source: Economic Report of the President (Washington, D.C.: U.S. Government Printing Office, 2004).

Thus, there has been a dramatic change in the composition of federal spending during the last four decades. In contrast with earlier times, national defense is no longer the primary focus of the federal government. In essence, the federal government has become an entity that taxes working-age Americans in order to provide income transfers and health care benefits primarily for senior citizens. Furthermore, spending on the elderly is almost certain to increase once the baby boomers begin to retire, starting around 2010.

TAXES AND THE FINANCE OF GOVERNMENT

I'm proud to be paying taxes in the United States. The only thing is—I could be just as proud for half the money.

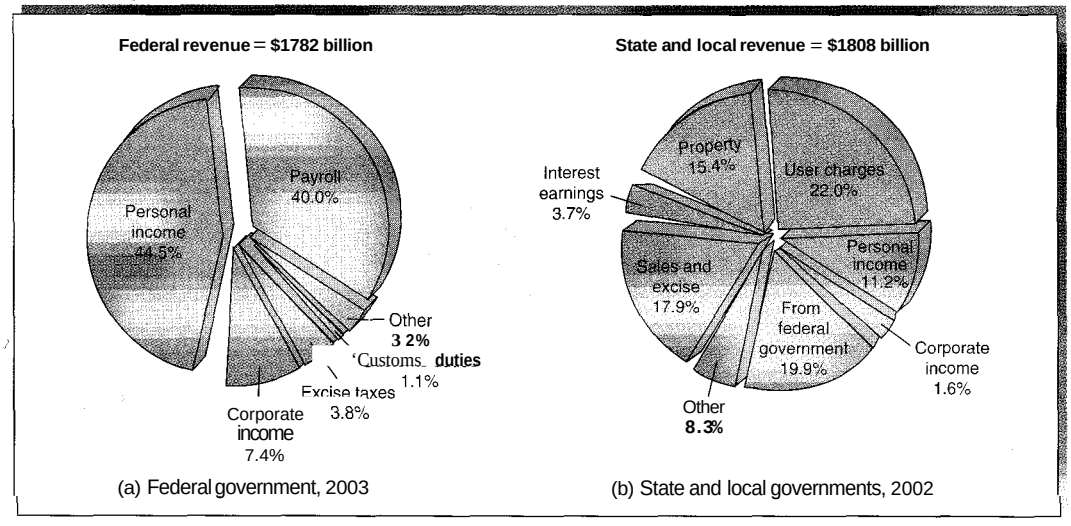
—Comedian Arthur Godfrey

Government expenditures must be financed through taxes, user charges, or borrowing.[2] Borrowing is simply another name for future taxes that will have to be levied to pay the interest on the borrowed funds. Thus, it affects the timing but not the level of taxes. In the United States, taxes are by far the largest source of government revenue. The power to tax sets governments apart from private businesses. Of course, a private business can put whatever price tag it wishes on its products; but no private business can force you to buy them. With its power to tax, a government can force citizens to pay, whether or not they receive something of value in return. As government expenditures have increased, so, too, have taxes. Taxes now take more than one-third of the income generated by Americans.

[2]In addition to user charges, taxes, and borrowing, the operations of government might be financed by printing money. But this is also a type of tax (it is sometimes called an "inflation tax") on those who hold money balances.

EXHIBIT 3
Sources of Government Revenue

Almost half of federal revenues are derived from the personal income tax. The payroll tax and corporate income tax are also major sources of federal revenue. The major revenue sources of state and local governments are sales and excise taxes, personal income taxes, user charges, grants from the federal government, property taxes.



Source. Economic Report of the President, 2004, and *Statistical Abstract of the United States, 2003.* http://www.census.gov.

Types of Taxes

Exhibit 3 indicates the major revenue sources for the federal and state and local levels of government. At the federal level, the personal income tax accounts for almost half of all revenue. Although income from all sources is covered by the income tax, only earnings derived from labor are subject to the payroll tax. Payroll taxes on the earnings of employees and self-employed workers finance Social Security, Medicare, and unemployment compensation benefits. The payroll tax accounts for 40 percent of federal revenue. The remaining sources of revenue, including the corporate income tax, excise taxes, and customs duties, account for less than 20 percent of federal revenue.

Both sales and income taxes are important sources of revenue for state governments. **A** sales tax is levied by forty-five of the fifty states (Alaska, Delaware, Montana, New Hampshire, and Oregon are the exceptions). State and local governments derive about 18 percent of their revenue from this source. Personal income taxes are imposed by forty-two states (Alaska, Florida, Nevada, New Hampshire, South Dakota, Texas, Washington, and Wyoming are the exceptions), and they provide approximately 11 percent of state and local government revenue.[3] Property taxes (levied mostly at the local level), grants from the federal government, and user charges (prices for services provided by the government) also provide substantial revenues for state and local governments.

Taxes and the Cost of Government

There are no free lunches. Regardless of how they are financed, activities undertaken by the government will incur costs. When governments purchase resources and other goods and services to provide rockets, education, highways, health care, and other goods, the resources used by the government will be unavailable to produce goods and services in the

[3]New Hampshire does levy a tax on income derived from dividends and interest.

private sector. As a result, private-sector output will be lower. This reduction in private-sector output is an opportunity cost of government. Furthermore, this cost will be present whether government activities are financed by taxes or borrowing.

Moreover, a tax dollar extracted from *an* individual or a business ends up costing the private economy much more than just one dollar. There are two main reasons why this is the case.

First, the collection of taxes is costly. The administration, enforcement, and compliance of tax legislation requires a sizable volume of resources, including the labor services of many highly skilled experts. The IRS itself employs 100,000 people. In addition, an army of bookkeepers, tax accountants, and lawyers is involved in the collection process. According to the Office of Management and Budget, each year individuals and businesses spend more than 6.7 billion hours (the equivalent of 3.3 million full-time year-round workers) keeping records, filling out forms, and learning the tax rules and other elements of the tax-compliance process.[4] More than half of U.S. families now retain tax-preparation firms like H&R Block and Jackson Hewitt to help them file the required forms and comply with the complex rules. Businesses spend roughly $5 billion each year in tax-consulting fees to the four largest accounting firms, to say nothing of the fees paid to other accounting, law, and consulting firms. In total, the resources involved amount to between 3 percent and 4 percent of national income (or 12 to 15 percent of the revenues collected). If these resources were not tied up with the tax-collection process, they could be employed producing goods and services for consumption.

Second, tuxes impose an additional burden on the economy because they eliminate some productive exchanges (and cause people to undertake some counterproductive activities). As we noted in Chapter 4, economists refer to this as an excess burden (or deadweight loss) because it imposes a burden over and above the tax revenue transferred to the government. It results because taxes distort incentives. When buyers pay more and sellers receive less due to the payment of a tax, trade and the production of output become less attractive and decline. Individuals will spend less time on productive (but taxed) market activities and more time on tax-avoidance and untaxed activities, such as leisure. Research indicates that these deadweight losses add between 9 percent and 16 percent to the cost of taxation.[5] This means that $1 in taxes paid to the government imposes a cost of somewhere between $1.20 and $1.30 on the economy. Thus, the cost of a $100 million government program financed with taxes is really somewhere between $120 million and $130 million. As a result, the government's supply of goods and services generally costs the economy a good bit more than either the size of the tax bill or the level of government spending implies.

When the burden of taxation is considered, it is also important to recognize that all taxes are paid by people. Politicians often speak of imposing taxes on "business" as if part of the tax burden could be transferred from individuals to a nonperson (business). This is not the case. Business taxes, like all other taxes, are paid by individuals. A corporation or business firm might write the check to the government, but it merely collects the money from someone else—from its customers in the form of higher prices, its employees in the form of lower wages, or its stockholders in the form of lower dividends—and transfers the money to the government.

How Has the Structure of the Personal Income Tax Changed?

The personal income tax is the largest single source of revenue for the federal government. The rate structure of the income tax is progressive; taxpayers with larger incomes face higher tax rates. However, the structure of the rates has changed substantially since 1960. In the early 1960s, there were twenty-four marginal tax brackets ranging from a low of 20 percent to a high of 91 percent. The Kennedy-Johnson tax cut reduced the lowest marginal

[4]Office of Management and Budget, *Information Collection Budget of the United States Government,* fiscal year 2004 Also see Tax Foundation *Special Brief* by Arthur Hall, March 1996

[5]The classic article on this topic is Edgar K. Browning, "The Marginal Cost of Public Funds," *Journal of Political Economy* 84, no 2 (April 1976) 283–98

rate to 14 percent and the top rate to 70 percent. The rate reductions during the Reagan years cut the top marginal rate initially to 50 percent in 1981 and later to approximately 30 percent during the period 1986–1988. During the 1990s, the top rate was increased to 39.6 percent, but the tax reductions during the administration of George W. Bush rolled back the top rate to 35 percent.

Thus, since the late 1980s, Americans with the highest incomes have paid sharply lower top marginal tax rates—rates in the 30 to 40 percent range, compared to top rates of 91 percent in the early 1960s and 70 percent prior to 1981. These reductions in the top rate make it tempting to jump to the conclusion that high-income Americans are now getting a free ride—that they now shoulder a smaller share of the personal income tax burden than in the past. But such a conclusion would be fallacious.

Exhibit 4 presents the Internal Revenue Service data on the share of the personal income tax paid by various classes of high-income taxpayers, as well as those in the bottom half of the income distribution, for the years 1963, 1980, 1990, and 2002. These data show that the share of the personal income tax paid by high-income Americans has increased substantially since 1963, and the increase has been particularly sharp since 1980. For example, the top 1 percent of earners paid 33.7 percent of the personal income tax in 2002, up from 19.1 percent in 1980 and 18.3 percent in 1963. The top 10 percent of income recipients paid 65.7 percent of the personal income tax in 2002, compared to 49.3 percent in 1980 and 47 percent in 1963. At the same time, the share of the personal income tax paid by the bottom half of the income recipients has steadily fallen from 10.4 percent of the total in 1963 to 7.1 percent in 1980 and 3.5 percent in 2002.

What is going on here? How can one explain the fact that high-income Americans are now paying more of the personal income tax even though their rates are now sharply lower than those in effect prior to 1981? Two major factors provide the answer. First, when marginal rates are cut by a similar percentage, the "incentive effects" are much greater in the top tax brackets. For example, when the top rate was cut from 91 percent to 70 percent during the Kennedy-Johnson years, high-income taxpayers in this bracket got to keep $30 out of every $100 of additional earnings after the tax cut, compared to only $9 before the rates were reduced. Thus, their incentive to earn additional income increased by a whopping 233 percent (30 minus 9 divided by 9)! On the other hand, the rate reduction in the lowest tax bracket from 20 percent to 14 percent meant that the low-

EXHIBIT 4
Share of Federal Income Taxes Paid by Various Groups, 1963–2002

Even though marginal tax rates have been reduced substantially during the last four decades, upper-income Americans pay a much larger share of the federal income tax today than was previously the case. In 2002, the richest 1 percent of Americans paid 33.7 percent of the federal income tax, up from 18.3 percent in 1963 and 19.1 percent in 1980. The richest 5 percent of Americans paid over half of the personal income tax, while the entire bottom half of the income distribution (the bottom 50 percent) paid only 3.5 percent of the total.

	SHARE OF TOTAL FEDERAL PERSONAL INCOME TAX PAID			
INCOME GROUP	1963	1980	1990	2002
Top 1%	18.3%	19.1%	25.1%	33.7%
Top 5%	35.6%	36.8%	43.6%	53.8%
Top 10%	47.0%	49.3%	55.4%	65.7%
Top 25%	68.8%	73.0%	77.0%	83.9%
Top 50%	89.6%	93.0%	94.2%	96.5%
Bottom 50%	10.4%	7.1%	5.8%	3.5%

Source: Internal Revenue Service (also available online at the Tax Foundation's Web site: http://www.taxfoundation.org/prtopincometable.htmI).

income taxpayers in this bracket now got to keep $86 of each additional hundred dollars that they earned compared to $80 prior to the tax cut. Their incentive to earn increased by a modest 7.5 percent (86 minus 80 divided by 80). Because the rate reductions increased the incentive to earn by much larger amounts in the top tax (and therefore highest income) brackets, the income base on which high-income Americans were taxed expanded substantially as their rates were reduced. As a result, the tax revenues collected from them declined only modestly. In the very highest brackets, the rate reductions actually increased the revenues collected from high-income Americans. (See Laffer Curve analysis of Chapter 4.) In contrast, the incentive effects were much weaker in the lower tax brackets and, as a result, rate reductions led to approximately proportional reductions in revenues collected from low- and middle-income taxpayers. This combination of incentive effects shifts the share of taxes paid toward those with higher incomes, the pattern observed in Exhibit 4.

Second, both the standard deduction and personal exemption have been increased substantially during the last couple of decades. This means that Americans are now able to earn more income before they face any tax liability. In 2002, for example, 30 percent (approximately 40 million returns) of those filing an income tax return either had zero tax liability or actually received funds from the IRS as the result of the **Earned Income Tax Credit.** This change in the structure of the personal income tax explains why people in the bottom half of income now pay such a small percentage of the personal income tax: 3.5 percent in 2002 compared to 10.4 percent in 1963.[6]

Earned Income Tax Credit
A provision of the tax code that provides a credit or rebate to persons with low earnings (income from work activities). The credit is eventually phased out if the recipient's earnings increase.

Income Levels and Overall Tax Payments

In addition to the personal income tax, the federal government also derives sizable revenues from payroll, corporate income, and excise taxes. How is the overall burden of federal taxes allocated among the various income groups? **Exhibit 5** presents Congressional Budget Office estimates for the average amount of federal taxes paid in 2001 according to income. On average, the top quintile (20 percent) of earners are estimated to pay 26.8 percent of their income in federal taxes. The average federal tax rate for the quintile with the next-highest level of income falls to 19.3 percent, and the average tax rate continues to fall as income declines. The average tax rate of the bottom quintile is 5.4 percent, about one-fifth of the average rate for the top quintile of earners. Clearly, the federal tax system is highly progressive, meaning that it takes a larger share of the income of those with higher incomes than from those with lower income levels.

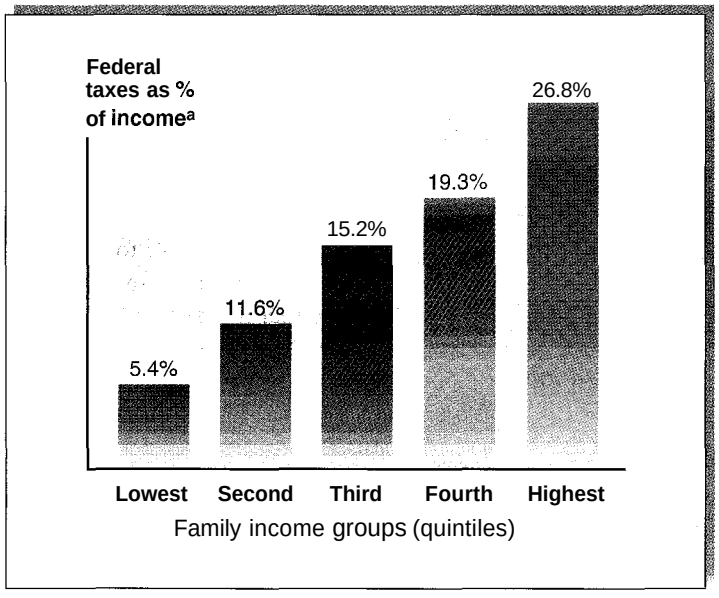
Does the Growth of Income Benefit the Federal Government?

The federal personal income tax brackets are indexed for inflation. Therefore, the tax brackets are widened as inflation increases the nominal incomes of individuals and families. However, no adjustment is made for increases in real incomes. Under a progressive tax system, a larger and larger share of income will be taxed at higher rates as real incomes rise. As a result, the growth of real income will automatically increase federal revenues more than proportionally. Thus, under the current progressive tax structure, the growth of real income will increase federal revenues as a share of total income if no offsetting action is taken. Some economists, particularly those with a public-choice perspective, argue that these automatic tax increases accompanying economic growth adversely affect the efficiency of political decision making. They believe that elected political officials would make better (more efficient) choices if they had to vote for higher taxes in order to adopt new spending programs and expand the relative size of government.

[6]The data of Exhibit 4 consider only the tax liability of taxpayers. They do not reflect the payments from the IRS to taxpayers as the result of the Earned Income Tax Credit, which was established in the mid-1980s. If these payments to taxpayers were taken into consideration, the net taxes paid by the bottom half of income recipients would have been less than 1 percent in 2002. Thus, the data of Exhibit 4 actually understate the reduction in the net share of taxes paid by the bottom half of income recipients during the last two decades.

EXHIBIT 5
Total Federal Taxes as a Share of Income, 2001

The federal income tax structure is highly progressive. Federal taxes take 26.8 percent of the income generated by the top quintile (20 percent) of earners, compared to 15.2 percent from the middle-income quintile and 5.4 percent from the lowest quintile of earners.



[a]Total federal taxes include income, payroll, and excise taxes.
Source: Congressional Budget Office, *Effective Federal Tax Rates: 1979–2001,* April 2004, http://www.cbo.gov/showdoc.cfm?index=5324&sequence=0.

SIZE OF GOVERNMENT: THE U.S. VERSUS OTHER COUNTRIES

There is substantial variation in the size of government across countries. As **Exhibit 6** illustrates, the relative size of government in most other high-income industrial countries is greater than that of the United States. In 2003, government spending summed to 58 percent of the economy in Sweden and 56 percent in Denmark. Government spending in France, Belgium, and Austria exceeded 50 percent of the economy. Compared to the United States, government was also quite large in Germany, Netherlands, and Italy. The size of government in Australia and Ireland was similar to that of the U.S., about 36 percent of the economy. Interestingly, the size of government was substantially smaller in South Korea, Singapore, Thailand, and Hong Kong—four Asian nations that have achieved rapid growth and substantial increases in living standards during the last four decades. As we proceed, we will return to this topic and investigate the effect of the size and functions of government on the growth and prosperity of nations.

The Future

A major share of government spending in the United States is now directed toward the elderly. The Social Security and Medicare programs constitute a huge share of federal spending. Once the baby boomers move into the retirement phase of life, beginning around 2010, federal expenditures on both of these programs are likely to balloon. This will make it very difficult to control the growth of government in the decades immediately ahead.

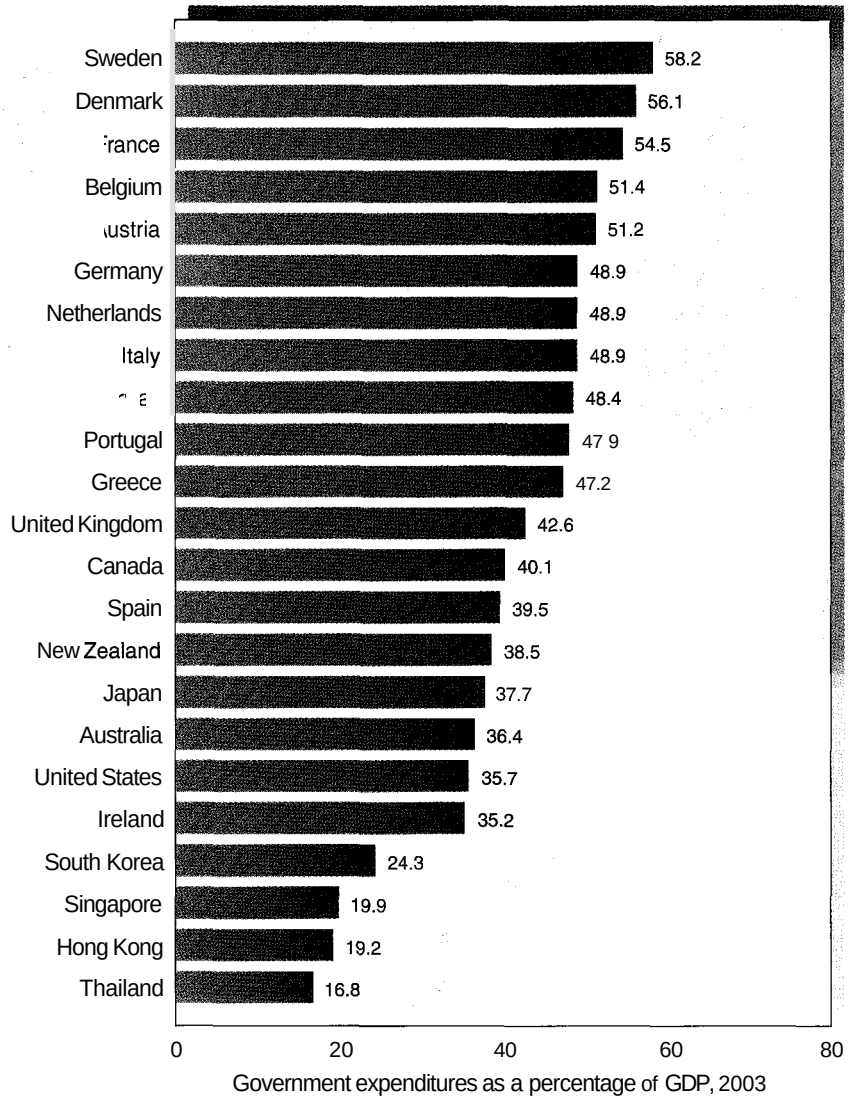


EXHIBIT 6
The Size of Governments—An International Comparison

The size of governments varies substantially across countries. In Sweden, government spending sums to almost three-fifths of the economy, compared to 35.7 percent in the United States and less than 20 percent in Singapore, Hong Kong, and Thailand.

Source: OECD, *OECD Economic Outlook*, No. 75, June 2004, Annex Table 25, and International Monetary Fund, *International Financial Statistics* and *Government finance Statistics Yearbook*. The data for Singapore and Hong Kong are for 2002 and 2001, respectively.

KEY POINTS

- During the first 125 years of U.S. history, federal expenditures per person were small and they grew at a relatively slow rate. But the size and nature of government has changed dramatically during the past 100 years. Today, the real (adjusted for inflation) spending per person of the federal government is roughly sixty times the level of 1916.
- During the last four decades, the composition of federal spending has shifted away from national defense and toward spending on income transfers and health care.
- As the size of government has grown, taxes have increased. Taxes impose a burden on the economy over and above the revenue transferred to the government because of (1) the administration and compliance costs and (2) the deadweight losses that accompany taxation.
- Overall, the federal tax system of the United States is highly progressive. Taxes as a percentage of income are approximately five times greater for the top quintile (20 percent) of families than for the bottom quintile.
- The size of government of the United States is smaller than that of the major Western European countries, but larger than for a number of high-growth Asian economies.

CRITICAL ANALYSIS QUESTIONS

*1. How do taxes influence the efficiency of resource use? How much does it cost for the government to raise an additional dollar (or $1 billion) of tax revenue?

2. During the last four decades, a smaller share of the federal budget has been allocated to national defense and a larger share to income transfers and health care. Does economics indicate that this change will help Americans achieve higher living standards?

3. Because the structure of the personal income tax is progressive, a larger share of income is taxed at higher rates as income grows. Therefore, economic growth automatically results in higher taxes unless offsetting legislative action is taken. Do you think this is an attractive feature of the current tax system? Why or why not?

4. Compared to the situation prior to 1981, the marginal tax rates imposed on individuals and families with high incomes are now lower. What was the top marginal personal income tax rate in 1980? What is the top rate now? Are you in favor of or opposed to the lower marginal rates? Why?

*5. As the result of changes during the last two decades, the bottom half of income recipients now pay little or no personal income tax. Rather than paying taxes, many of them now receive payments back from the IRS as the result of the Earned Income Tax Credit and Child Tax Credit programs. Do you think the increase in the number of people who pay no taxes will affect the efficiency of the political process? Why or why not?

*Asterisk denotes questions for which answers are given in Appendix B.

SPECIAL TOPIC 2

The Internet: How Is It Changing the Economy?

Focus

- Why is the development of the Internet an important economic phenomenon?
- How is the Internet changing product markets?
- What is the effect of the Internet on the labor market?

The Internet is kind of like a gold rush where there really is gold.

—*Bill Gates*[1]

[1]Bill Gates, *Microsoft Magazine*, January/February 1996.

The Internet is a gigantic library, super shopping mall, and extensive transportation system all wrapped into one. Far more documents can be obtained over the Internet than from even the largest brick-and-mortar library facility. Although a large mall can provide you with access to hundreds of shops in a given locality, the Internet provides access to millions of businesses located around the world. And for only a small fee, you can open your own shop in the world's super mall.

Music, movies, software, and financial services can be transported almost instantaneously over the Internet, something unheard of just a couple of decades ago. For example, online sales of goods and services such as airline tickets, computers, and books totaled $114 billion in 2003. This figure is expected to rise to $316 billion, or 12 percent of all projected retail sales, by 2010.[2] The Web is also influencing labor markets by changing how people find new jobs as well as how labor services are provided.

USE OF THE INTERNET

As **Exhibit 1** shows, the number of Web sites on the Internet has exploded in the past decade. In 1992, there were only 16,000 Web sites, but by 2004 there were more than 39 million sites. This tremendous growth in the number of Web sites has been matched by a dramatic rise in the number of Internet users. By October 2004, more than two-thirds of the population, or 200 million people, were using the Internet in the United States. The typical user employs the Internet intensively. The average user spends more than seven hours a week surfing the Web and visits twenty-two different Web sites per week.[3]

As **Exhibit 2** demonstrates, the use of the Internet has become a worldwide phenomenon. About one-third of the Internet users reside in Asia, compared to a little more than one-quarter in both North America and Europe. Residents of Latin America, the Middle East, and Africa account for a relatively small portion of the total users.

ECONOMIC GAINS FROM THE INTERNET

Why is the development of the Internet an important economic phenomenon? There are good reasons to believe that the Internet has improved productivity and efficiency — that it has helped us generate more value from available resources. There are three major sources of economic gains from the Internet.

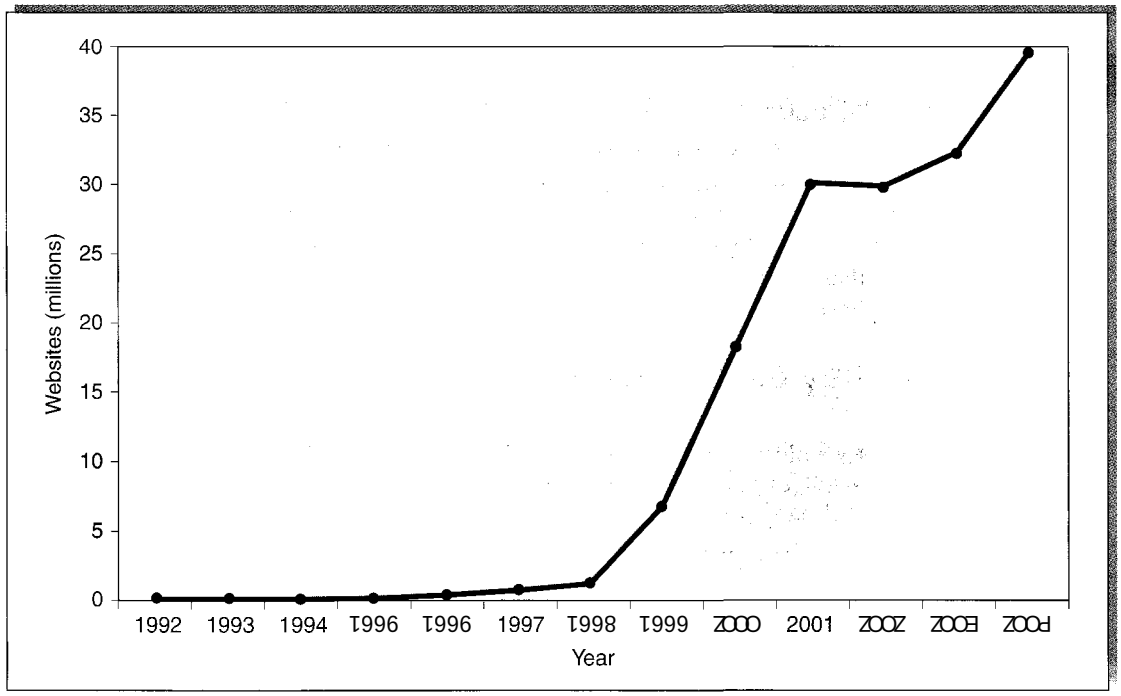
1. Gains from broader and more competitive markets. As we have previously discussed, gains from trade and competition are important sources of growth and prosperity. The Internet promotes the realization of gains from both. The cost of establishing an Internet firm is low, often only a few hundred dollars. The costs of identifying potential suppliers via the Internet are also low, and with the development of more efficient search devices, they are declining. Via the Internet, firms are able to compete over a much larger geographic area than they could a decade ago, and buyers are better able to purchase from sellers who are located far away. A case in point: a farm family in West Texas was able to start a successful business selling tumbleweeds to New York restaurants wanting southwestern décor. Markets are becoming more competitive and the location of both buyers and sellers less relevant. This is particularly true for goods that can be transported economically, either through the Internet or via other means of transportation.

[2] Carrie A. Johnson, "US eCommerce Overview: 2004 to 2010," August 2004. *Forrester Research Report,* Forrester Research Inc., Cambridge, Mass.

[3] The current number of Internet users as well as usage patterns can be obtained from http://www.nielsen-netratings.com/.

EXHIBIT 1
Number of Web Sites, 1992–2004

The number of Web sites rose dramatically between 1992 and 2004. By 2004, there were 39 million Web sites.



Source: http://www.zooknic.com/.

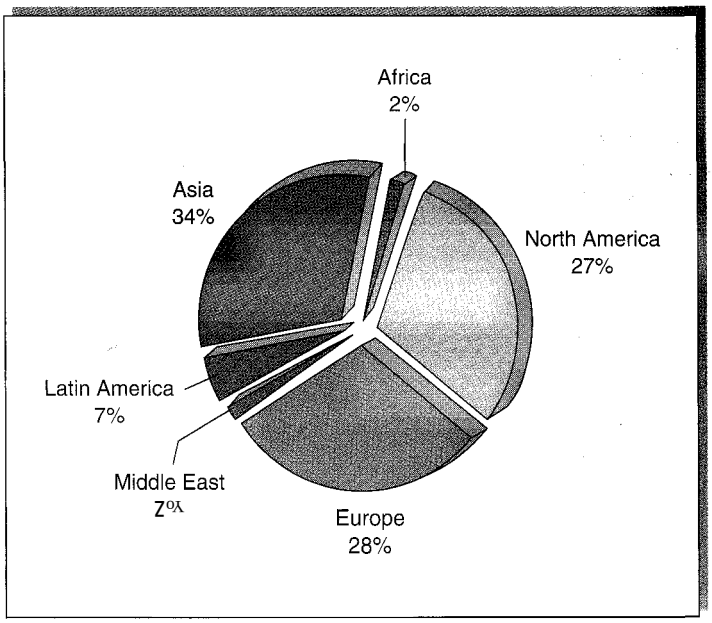


EXHIBIT 2
The Geographic Distribution of Internet Use, 2004

About one-third of Internet users live in Asia. About one-quarter reside in Europe, and another quarter are in North America.

Source: http://www.internetworldstats.com/.

2. Gains from lower transaction costs. Transaction costs are an obstacle to the realization of gains from trade. The Internet, however, often reduces the cost of transactions, including the cost of information. Think about, for example, the software you are able to download and the time and money it saves you from having to go to a store to buy it. The U.S. economy has already reaped substantial gains from these lower transcation costs. But as use of the Internet increases, there is reason to believe that the future gains will be even greater.

3. Gains from networking. Like the telephone of an earlier era, the Internet is a networking system. Telephones were not very valuable when only a few people had one, but their value increased dramatically as more and more people acquired them. The Internet has this same characteristic. The value of the system to current users increases as more and more people join the network. Growth of the network will make it more likely that Internet sellers will be offering products you want to buy and that potential Internet buyers will be searching for goods you are willing to sell.

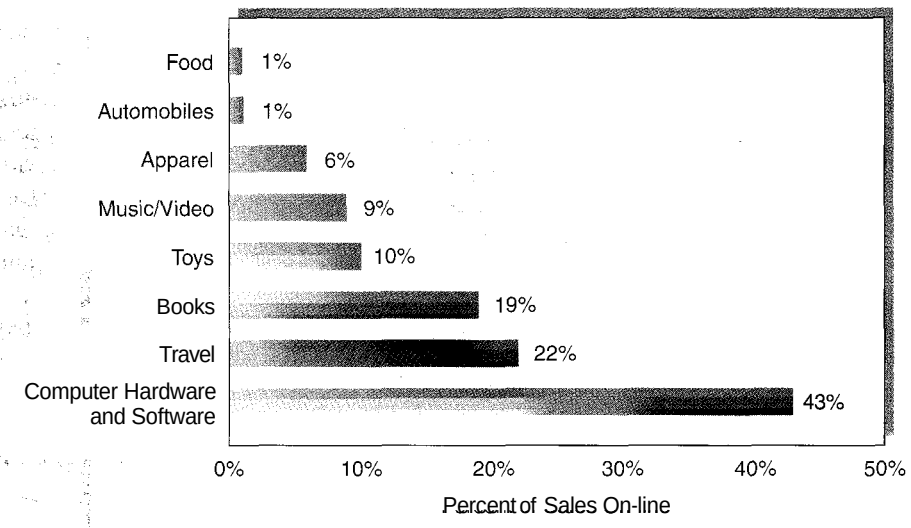
KEY SECTORS OF INTERNET GROWTH

Exhibit 3 shows that the Internet is now extensively used in several consumer product markets. The percentage of sales conducted online is about one-fifth or more in retail markets such as travel, books, and computer hardware and software. The rapid rise in the importance of the Internet in these markets is the result of the relatively low transportation costs for these items as well as the availability of information about these standardized products online. A much smaller percentage of the sales of automobiles, food, apparel, and toys and music items are conducted online. The low market penetration for online firms in the food market is partly because customers can't observe the condition of these items. Similarly, the sale of clothing over the Internet is hampered by the fact that one can't examine the fit of clothing online.

The Internet is also quickly transforming how banking is conducted. The number of households engaged in online banking has been rapidly increasing. In 2003, 19 million households paid their bills online, but this figure is projected to rise to 61 million by 2008.[4]

EXHIBIT 3
Online Market Penetration, 2004

The percentage of sales conducted online is about one-fifth or more in the travel, books, and computer hardware and software markets. The market share for online firms in the automobile, food, apparel, and toys and music/video markets is much smaller.



Source: Shop.org and Boston Consulting Group, State of Online Retailing 7.0, May 2004.

[4]http://www.clickz.com/.

The Web has revolutionized the market for used consumer durables. The Web site eBay has made it much easier for people to buy or sell used goods. Sales are conducted by an online auction process, and eBay earns money by charging a commission for each sale completed. Prior to purchasing an item on eBay, prospective buyers can examine earlier customers' ratings of a given seller. The ease with which trades can be conducted on the eBay Web site enabled the company to grow to 114.0 million users in 2004. Unlike most Internet firms, eBay has been profitable since its inception.[5]

MORE EFFICIENT CONSUMER MARKETS

The Internet helps create value in final product markets by either reducing sellers' costs or improving the matches between sellers and consumers.[6] First, several factors help online retail firms achieve lower costs. Their handling costs will often be lower because it will not be necessary for the firms to unpack products or put them on display. Losses due to shoplifting, which often are about 3 percent of sales, are eliminated. Web-based firms are generally able to use low-cost warehouses rather than expensive stores in urban or suburban areas. Online retailers often have lower sales comissions on their products than traditional brick-and-mortar firms do.

Second, the Internet can improve matches between sellers and consumers by making information about goods easier to get. The Internet is an excellent source of information for consumers about available goods and services because an individual can get very specific information about a product or service at a low cost any time of day or night. Web shoppers can obtain information from online versions of product catalogs, product reviews, and price comparisons. Sites like Amazon.com and eopinions.com also allow users to rate how well they like different products and provide product recommendations to customers based on their past purchases. Web shoppers can get product samples of books, music, and software, which can help them better decide whether to purchase products or not.

The Web can also make it easier for consumers to obtain access to hard-to-find goods, particularly specialty products and goods with unique characteristics. For example, it is very expensive for a chain such as The Gap to maintain a large inventory of its products at many different locations. As a consequence, each Gap store branch has only a limited inventory of different styles and sizes of jeans. This results in some consumers having to make compromises on the jeans they purchase. Online firms, however, can take advantage of economies of scale of centralized inventories and stock a much wider variety of products. This enables Web-based firms to provide a better match to customers' preferences.

Dell Computers provides the best example of the effect of product customization. Prior to Dell, the typical personal computer manufacturer forecasted demand for various computer models and then produced these computers in batches. The computers were then distributed to wholesalers and retail stores. Changes in inventories and prices were used to create a balance between supply and demand. Dell uses a different business model. Online, Dell customers can customize their computer purchases on several dimensions such as the processor, memory, hard disk size, monitor, and so on. After the order is placed, the computer is made and then shipped directly to the customer. As a result, Dell does not need to manufacture all possible configurations of its computers and therefore is able to achieve lower costs.

A major problem, of course, with purchasing goods online is that consumers can't touch, taste, smell, or try on the goods before purchasing them. A potential solution is hybrid stores that act as a showroom for a broad range of products. These stores have for display only a few of each product available for sale. Shoppers can examine the products and then place their order either in the store or later at home online.

[5]http://www.ebay.com/.

[6]For an overview of the effect of the Internet on product markets, see Severin Borenstein and Garth Saloner, "Economics and Electronic Commerce," *Journal of Economic Perspectives,* winter 2001: 3–12. This section is partly derived from this article.

Will Internet firms eventually dominate product markets? It depends on whether they are an efficient method of supplying goods and services. When Internet firms have lower costs or provide consumers with other benefits (for example, broader selection, faster delivery, or greater customization), they will be able to compete effectively. With time, they may even dominate some sectors of the economy. On the other hand, when traditional retailers have lower costs and provide consumers with other benefits (immediate access to goods, inspection of the items, and/or local service contacts, for example), they will survive and prosper. In a market economy, consumers are the ultimate judge. Their choices will determine which firms, be they traditional or online, will expand and prosper and which will be driven from the market.

MORE EFFICIENT INPUT MARKETS AND PRODUCTION PROCESSES

The ability of the Internet to improve matches between buyers and sellers also applies to business-to-business transactions. In fact, a higher volume of business-to-business transactions is being conducted on the Web today than business-to-consumer transactions. Some online firms serve as intermediaries between companies. For example, some Web sites auction off goods such as steel and advertising space. Other online companies have set up exchanges for a wide variety of goods. One such exchange is Covisint, which was started by DaimlerChrysler, Ford, and General Motors to handle their transactions with parts suppliers. Because these three companies purchase about $250 billion worth of parts each year, this Web site has the potential to become one of the largest firms on the Internet.[7]

These intermediaries can take the form of exchanges, online auctions, and brokers. The intermediaries can often reduce the buyer's search costs and facilitate one-stop shopping and thereby reduce the need for costly contacts with multiple suppliers.

Many companies are also increasingly conducting "reverse auctions" online. Vendors that supply inputs bid on the prices they will charge the companies conducting the auctions. In other words, instead of bidding to buy products online, vendors are bidding to sell their products online (which is why they are called "reverse" auctions). Sun Microsystems' Web site, for example, allows programmers and other vendors to bid on contracts to solve Sun's clients' software problems.

The Internet can also lower sales costs between companies. For example, switching to an electronic version of the purchasing process from a paper one can substantially reduce the cost of buying goods and services. The cost of completing a paper transaction has been estimated to be roughly $50 per transaction. In addition, improved information from a firm's suppliers about the availability of their products can enable the firm to lower its inventory of inputs and thus its costs. In addition, the Web can be used to take advantage of differences in time zones. For example, software projects can often be transferred over the Internet from programmers in the United States to their counterparts in India at the end of the workday. Via the Web, radiologists in India examine emergency medical scans overnight, so U.S. radiologists can get some sleep. Being able to outsource work anywhere around the world is also cost-effective for firms.

LABOR MARKETS AND THE INTERNET: FASTER AND BETTER EMPLOYEE-EMPLOYER MATCHES[8]

The explosive growth of the Internet has changed how people search for jobs and firms hire workers. There are now more than 3,000 job search sites. Monster.com, the leading

[7]David Lucking-Reiley and Daniel F Spulber, "Business to Business Electronic Commerce," *Journal of Economic Perspectives*, winter 2001 55–68

[8]This section draws on David H Autor, "Wiring the Labor Market," *Journal of Economic Perspectives*, winter 2001 25–40

job-posting site, indicated in October 2004 that it had more than 800,000 job openings posted on its site and the résumés of more than 25 million job seekers. Currently, the Internet is used in 15 percent of all job searches by unemployed workers. Half of job seekers with Internet access use it in their job search. Today, more people use the Internet to look for jobs than search methods such as contacting friends or relatives and using private employment agencies.

Job-posting Web sites have several advantages over traditional newspaper help-wanted ads. They contain more job openings and are easier to search. The job openings can be more current because employers can post ads immediately as well as edit them after their initial posting. Online **jobs** sites also permit individuals to advertise their skills to potential employers. Lastly, the cost to advertise a job opening is lower. The cost of a 30-day advertisement on Monster.com is less than 5 percent of the cost of a job advertisement in one issue of the Sunday *New York Times.*

Job-posting software on sites can also help match job seekers with employers. Software can compare the résumés of job seekers with descriptions of open positions. If an appropriate match occurs, then both the employer and the job seeker can be notified. Some Web programs advise applicants of new employment opportunities based on the job openings he or she has applied for in the past. Employers can also screen candidates by administering personality and skills tests over the Internet.

Because matches between employers and prospective employees can be made more quickly, it's possible that the Internet has the potential to reduce frictional unemployment in the economy. Moreover, online screening of candidates can lead to better matches, and better matches lead to higher productivity for firms.

The Internet's effect on employee turnover is less clear, however. On the one hand, better and faster matches can be made, as we have pointed out. On the other hand, because the Internet enables employed workers to easily search for a new position, turnover may increase. In fact, 7 percent of the employed indicate that they routinely use the Internet to search for potential new job opportunities.

The Delivery of Labor Services and the Training of Employees

The Internet has changed how workers provide labor services to employers. Remote access to documents and e-mail will permit some workers to provide part or all of their work at home or other locations. In 2001, about 15 percent of workers reported working at home at least once per week.[9] Less time is spent commuting, leaving more time for work activities that add economic value.

The Internet has also changed how workers obtain their skills. Students, of course, can now go to school and get their degrees online. Many employers provide formal and informal training to their workers online. This training plays an important role in workers' earnings. The online delivery of skills training has the potential to reduce the cost and increase the convenience of getting such training.

CONCLUDING THOUGHT

The Internet is an important technological change — perhaps as important as the development of electricity, the railroad, or the automobile. There are reasons to believe that it will improve economic efficiency and help us achieve higher living standards. It has increased the interaction of people around the world. It may also change lifestyles and alter cultural values around the globe. It will be exciting to follow these developments in the decades immediately ahead.

[9] http://www.bls.gov/.

KEY POINTS

- The use of the Internet has grown dramatically in the past decade and is now a worldwide phenomenon.
- The Internet tends to improve productivity and the efficiency of resource use because it (1) increases the breadth and competitiveness of markets, (2) lowers transaction costs, and (3) becomes more valuable as additional users and Web sites are added,
- The Internet can improve the operation of product markets by reducing costs and improving the matches between buyers and sellers. Costs are reduced due to lower distribution and production costs. Matches are improved through better information about available goods, greater access to goods, and increased customization.
- The volume of business-to-business sales being conducted on the Internet is even greater than the volume of retail-to-consumer sales. The Internet is making inputs cheaper and input markets more competitive, and is helping streamline production processes.
- The Internet has become an integral part of the job search process. Quicker and improved matches between employers and employees have the potential to reduce unemployment. The Internet also makes it possible for many employees to work at home and other locations and get training online.

CRITICAL ANALYSIS QUESTIONS

1. What effect does the Internet have on the efficiency of markets? Explain. How is the Internet likely to influence productivity and the growth of output in the years immediately ahead?

*2. The share of airline tickets bought over the Internet has grown rapidly, whereas the percentage of groceries purchased online remains minuscule. What factors likely explain this difference?

3. Indicate how the production, marketing, and distribution of each of the following are likely to be influenced by the development of the Internet: (a) popular music, (b) movies, (c) automobiles, (d) commercial employment agencies, (e) beautician services, and (f) health care. Briefly explain your response.

*Asterisk denotes questions for which answers are given in Appendix B

SPECIAL TOPIC 3

The Economics of Social Security

Focus

- Why is Social Security headed for problems?
- Will the Social Security Trust Fund lighten the tax burden of future generations?
- Does Social Security transfer income from the rich to the poor? How does it impact the economic status of blacks, Hispanics, and those with fewer years of life expectancy?
- Should the Social Security system be reformed?

The federal government's handling of [Social Security] pension monies is very different from that of private pension plans.

—Mark Weinberger[1]

[1]Mark Weinberger, *Social Security: Facing the Facts* (Washington, DC: Cato Institute, 1996), 2.

The ongoing debate about the future structure of Social Security is particularly important to younger people. It is their lives that will be most affected by how this issue is handled. In the United States, the program is officially known as Old Age and Survivors Insurance (OASI). It offers protection against the loss of income that usually accompanies old age or the death of a breadwinner. In spite of its official title, Social Security is not based on principles of insurance. Private insurance and pension programs invest the current payments of customers in buildings, farms, or other real assets. Alternatively, they buy stocks and bonds that finance the development of real assets. These real assets generate income that allows the pension fund (or insurance company) to fulfill its future obligations to its customers.

Social Security does not follow this saving-and-investment model. Instead, most of the funds flowing into the system are paid out to current retirees and survivors in the program. In essence, the Social Security system is an intergenerational income-transfer program. Most of the taxes collected from the present generation of workers are paid out to current beneficiaries. Thus, the system is based on "pay-as-you-go," rather than on the savings and investment principle.

The Social Security retirement program is financed by a flat-rate payroll tax of 10.6 percent applicable to employee earnings up to a cutoff level. In 2005, the earnings cutoff was $90,000. Thus, employees earning $90,000 or more paid $9,540 in Social Security taxes to finance the OASI retirement program.[2] The income cutoff is adjusted upward each year by the growth rate of nominal wages. Although the payroll tax is divided equally between employee and employer, it is clearly part of the employee's compensation package, and most economists believe that the burden of this tax falls primarily on the employee. The formula used to determine retirement benefits favors those with lower earnings during their working years. However, as we will discuss later, the redistributive effects toward those with lower incomes are more apparent than real.

When the program began in 1935, not many people lived past age 65, and the nation had lots of workers and few eligible retirees. As **Exhibit 1** illustrates, there were 16 workers for every Social Security beneficiary as recently as 1950. That ratio has declined sharply through the years. As a result, higher and higher taxes per worker have been required just to maintain a constant level of benefits. There are currently 3.3 workers per Social Security retiree. By 2030, however, that figure will decline to only 2.2.

Because there were many workers per beneficiary during the early years of Social Security, it was possible to provide retirees with generous benefits while maintaining a relatively low rate of taxation. Many of those who retired in the 1960s and 1970s received real benefits of three or four times the amount they paid into the system, far better than they could have done had they invested the funds privately. The era of high returns, however, is now over. The program has matured, and the number of workers per beneficiary has declined. Payroll taxes have risen greatly over the decades, and still higher taxes will be necessary merely to fund currently promised benefits.

Studies indicate that those now age 40 and younger can expect to earn a real rate of return of about 2 percent on their Social Security tax dollars, substantially less than what they could earn from personal investments. Thus, Social Security has been a good deal for current and past retirees. It is not, however, a very good deal for today's middle-aged and younger workers. ■

*Additional payroll taxes are levied for the finance of disability programs (1.8 percent) and Medicare (2.9 percent). Thus, although the total payroll tax sums to 15.3 percent, only revenues from the 10.6 percent rate are used to finance the benefits to retirees and surviving dependents. Note that the earnings cutoff does not apply to the Medicare portion of the payroll tax.

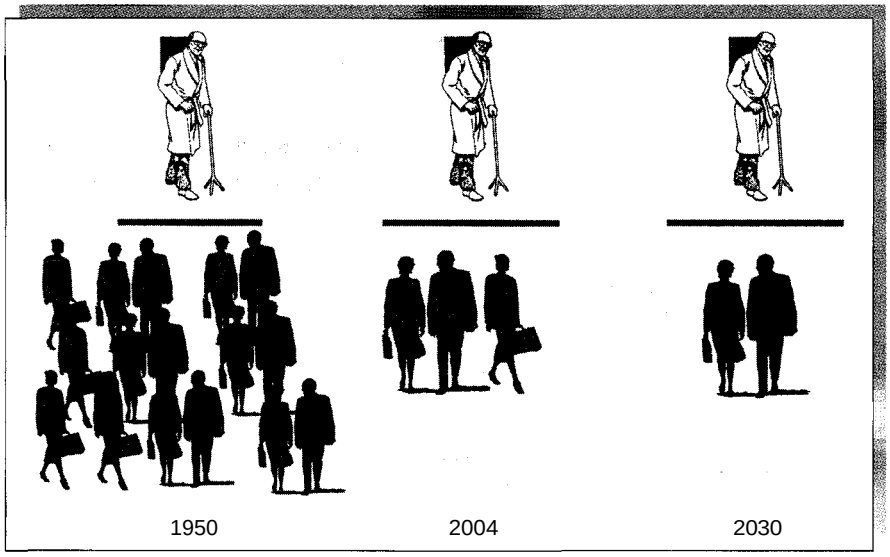


EXHIBIT 1
Workers per Social Security Beneficiary

In 1950, there were 16 workers per Social Security beneficiary. By 2004, the figure had fallen to only 3.3. By 2030, there will be only 2.2 workers per retiree. As the worker/beneficiary ratio falls under a pay-as-you-go system, either taxes must be increased or benefits reduced (or both).

Source: 2004 Annual Report of the Board of Trustees of the Federal Old Age and Survivors Insurance and Disability Insurance Trust Funds (Washington, DC: Government Printing Office, 2004), p. 47.

WHY IS SOCIAL SECURITY HEADED FOR PROBLEMS?

The flow of funds into and out of a pay-as-you-go retirement system is sensitive to demographic conditions. The Social Security system is currently enjoying a period of highly favorable demographics. The U.S. birthrate was low during the Great Depression and World War 11. The Great Depression and World War II group is now retiring, and because it is a relatively small generation, payments to it are also relatively small. The birthrate rose sharply during the two decades following World War 11. These baby boomers are now in their prime working years, and their large numbers are expanding the flow of revenues into the Social Security retirement system.

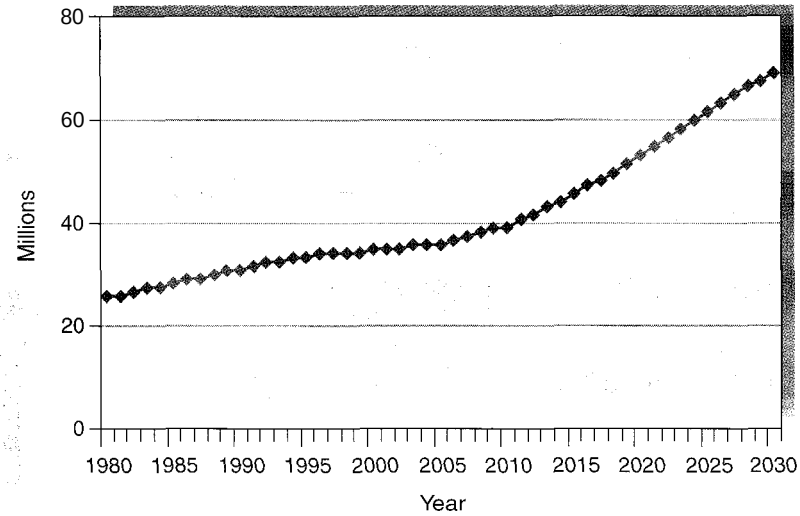
However, as **Exhibit 2** shows, the situation will change dramatically when the baby boomers start retiring around 201 1. Their retirement, combined with rising life expectancies, will substantially increase the number of retirees relative to the number of workers. As we previously noted, the number of workers per Social Security retiree will fall from the current 3.3 level to only 2.2 in 2030.

Exhibit 3 illustrates the effect of demographics on the pay-as-you-go Social Security system. Currently, the funds flowing into the system (pushed up by the large baby boom generation) exceed the expenditures on benefits for retirees (pulled down by the small Great Depression and World War II generation). But the retirement of the baby boomers around 201 1 will begin pushing the expenditures of the system upward at a rapid rate. The current surplus of revenues from the payroll tax relative to retirement benefits will dissipate around 2018. After 2018, the deficits will grow larger and larger as the number of beneficiaries relative to workers continues to grow in the decades ahead.

The revenues derived from the payroll tax have exceeded the benefits paid to current retirees since the mid-1980s (see Exhibit 3). Currently, only about 80 percent of the revenues are required for the payments to current beneficiaries. Thus, the system is currently running a surplus—about $138 billion per year. The surpluses are projected to continue for approximately another decade. If other elements of the federal budget were in balance, the Social Security surpluses could be used to pay off some of the federal government's outstanding debt. In turn, the debt reduction would reduce the government's future interest payments, which would make it easier to deal with rising Social Security expenditures when the baby boomers retire. However, this has not been the case. Most of the Social

EXHIBIT 2
U.S. Population Aged 65 and Over, 1980–2000 and Projections to 2030

As shown here, the growth rate of the elderly population will accelerate after 2010 as the baby boomers move into the retirement phase of life. This will place strong pressure on both the Social Security and Medicare programs.



Source: http://www.census.gov/.

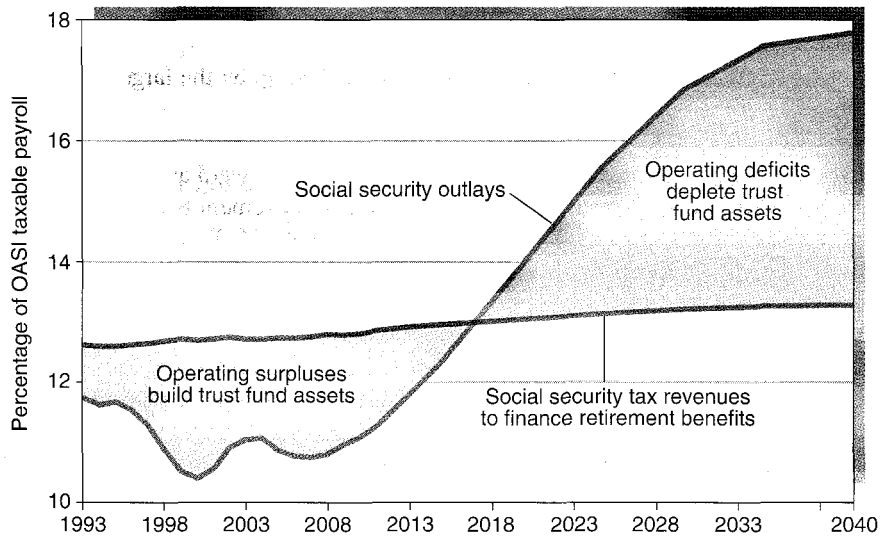
Security surpluses since 1980 have been used to finance current government operations rather than retiring the government debt.

WILL THE TRUST FUND LIGHTEN THE FUTURE TAX BURDEN?

Under current law, the surpluses are channeled into the Social Security Trust Fund (SSTF). The Trust Fund uses the revenue to buy special nonmarketable bonds from the U.S. Treasury. By 2018, the Social Security Trust Fund is expected to grow to more than $5 trillion. Social Security actuaries calculate that this will provide sufficient funds for payment of promised benefits until 2042.

EXHIBIT 3
The Forthcoming Deficit between Payroll Tax Revenues and Social Security Benefit Expenditures

Given current payroll taxes and retirement benefit levels, the system will run larger and larger deficits in the 2018–2030 period and beyond.



Source: Social Security Administration 2004 OASDI Annual Trustees Report, http://www.ssa.gov/.

Will a large trust fund make it easier to deal with the retirement of the baby boomers? Some people are surprised to learn that there is little reason to believe that it will. Unlike the bonds, stocks, and physical assets held by a private pension fund or insurance company, the SSTF bonds will not generate a stream of future income for the federal government. Neither are they a "pot of money" set aside for the payment of future benefits. Instead, the trust fund bonds are an IOU from one government agency—the Treasury—to another—the Social Security Administration. The federal government is both the payee and recipient of the interest and principal represented by the SSTF bonds. ***No matter how many bonds are in the trustfund, their net asset value to the federal government is zero!***

Thus, the number of IOUs in the trust fund is largely irrelevant.[3] The size of the trust fund could be doubled or tripled, but that would not give the government any additional funds for the payment of benefits. Correspondingly, the trust fund could be abolished, and the government would not be relieved of any of its existing obligations or commitments. To redeem the bonds and thereby provide the Social Security system with funds to cover future deficits, the federal government will have to raise taxes, cut other expenditures, or borrow from the public. These options will not change with the depletion of the trust fund.

THE REAL PROBLEM OF THE CURRENT SYSTEM

The real problem faced by the pay-as-you-go Social Security system will arise in about 2018 when the revenues from the payroll tax will begin to fall short of the benefits promised to retirees. The deficits of the system will become larger and larger throughout the 2020s and 2030s. Under current law, revenues will be sufficient to pay only about three-quarters of promised benefits by 2030, and less in later years.

There are only four ways to cover future shortfalls: (1) cut benefits, (2) increase taxes, (3) cut spending in other areas, or (4) borrow. None of these options are attractive and, regardless of how the gap is filled, a slowdown in the rate of economic growth is likely to occur. If benefits are reduced, current beneficiaries and persons near retirement will—quite correctly—feel that a commitment made to them has been broken. It will also be difficult to cover the shortfall with higher taxes. Once the baby boom generation retires, approximately a 50 percent increase in the payroll tax or a 30 percent increase in the personal income tax will be needed to cover Social Security deficits. Tax increases of this magnitude will have a negative effect on the economy. Neither will it be easy to cut expenditures in other areas of the federal budget. Defense spending was already cut substantially as a share of the economy during the 1990s, and it is likely to be pushed upward in the future by external threats, including those arising from terrorism. Furthermore, the growth of the elderly population is sure to put upward pressure on Medicare spending, another major federal program. Finally, borrowing to cover the shortfall will put upward pressure on interest rates, and taxes will have to be higher in the future merely to cover the interest obligations. Thus, it merely delays the problem.

Not even robust economic growth would eliminate the future shortfall. Retirement benefits are indexed to the average growth in nominal wages. If higher productivity enables *real* (inflation-adjusted) wages to rise quickly, then Social Security benefits will rise, too. For example, if inflation is zero, and real wages start growing at 2 percent a year instead of their previous level of 1 percent, then the formula used to calculate Social Security benefits will also begin to push up those benefits more rapidly. Higher economic growth may temporarily improve Social Security's finances, but under current law the improvement will not last.[4]

[3]Of course, the SSTF bonds represent funds borrowed by the Treasury from the Social Security system. This increases the legitimacy of claims on these funds by future Social Security recipients. It also indicates that the trust fund is similar to what is called budget authority, which provides the legal permission for the government to spend funds on an item.

[4]See Garth Davis, "Faster Economic Growth Will Not Solve the Social Security Crisis," Heritage Center for Data Analysis, Feb. 3.2000.

WHO IS HELPED AND WHO IS HURT BY SOCIAL SECURITY?

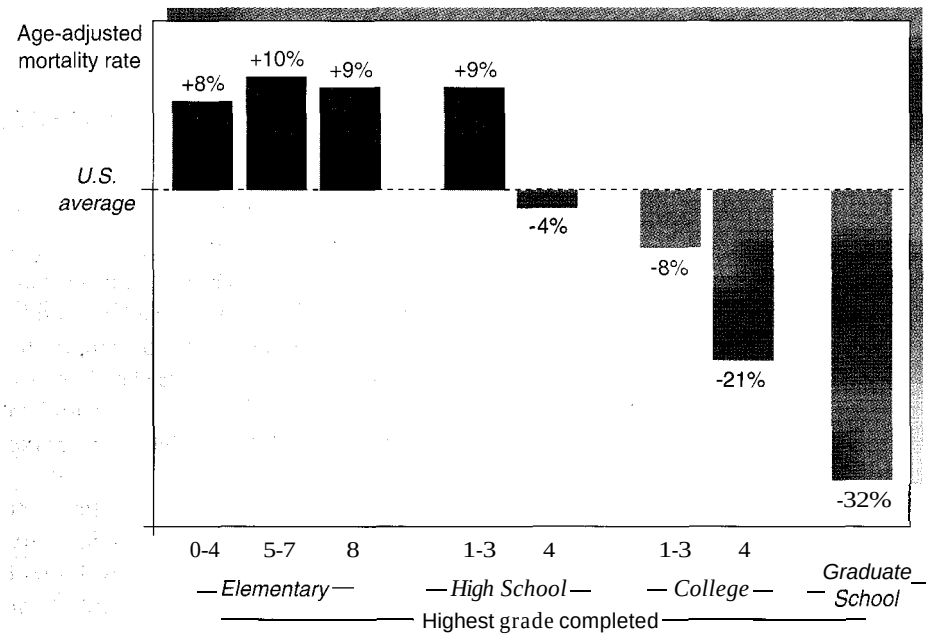
When Social Security was established in 1935, the population was growing rapidly, only a few Americans lived to age 65, and the labor force participation rate of women was low. Social Security was designed for that world. But today's world is dramatically different. Several aspects of the system now seem outdated, arbitrary, and, in some cases, unfair.

Does Social Security Help the Poor?

Social Security has gained many supporters because of the belief that it redistributes wealth from rich to poor. The system is financed with a flat tax rate up to the cutoff limit, but the formula used to calculate benefits disproportionately favors workers with low lifetime earning ~However, other aspects of the system tend to favor those with higher incomes. First, workers with more education and high earnings tend to live longer than those with less education and lower earnings. As **Exhibit 4** shows, the age-adjusted mortality rate of persons with less than a high school education is 8 to 10 percent higher than the average for all Americans. As years of schooling increase, mortality rates fall. The age-adjusted mortality rate of college graduates is 21 percent below the average for all Americans, whereas the rate for persons with advanced degrees is 32 percent below the average. Given the strong correlation between education and earnings, the age-adjusted mortality figures indicate that, on average, Americans with higher earnings live longer than their counterparts with less education and lower earnings. As a result, high-wage workers will, on average, draw Social Security benefits for a longer period of time than low-wage workers. Correspondingly, low-wage

EXHIBIT 4
Mortality Rates by Level of Education

As shown here, the age-adjusted mortality rates are lower for people with more education. Because of the close link between education and income, people with higher incomes tend to live longer and, therefore, draw Social Security benefits for a longer period of time than people with less education and income.



Source: Center for Data Analysis, Heritage Foundation.

[5]Retirement benefits are based on the best 35 years of earnings from a worker's career. Benefits are calculated by taking 90 percent of the first $7,524 a year of earnings, 32 percent of earnings between $7,524 and $45,348, and just 15 percent of earnings above $45,348 up to the earnings cutoff of $90,000. Therefore, as base earnings rise, benefits fall as a percentage of average earnings (and payroll taxes paid) during one's lifetime. For example, the retirement benefits of persons with base annual earnings of $10,000 sum to 76 percent of their average working year earnings. In contrast, the retirement benefits of those with base earnings of $60,000 are only 35 percent of their average preretirement earnings. These figures are based on the formula for 2005. The figures are adjusted each year for the growth of nominal wages.

workers are far more likely to pay thousands of dollars in Social Security taxes and then die before, or soon after, becoming eligible for retirement benefits.

Second, low-wage workers generally begin full-time work at a younger age. Many work full-time and pay Social Security taxes for years while future high-wage workers are still in college and graduate school. Low-wage workers generally pay more into the system earlier, and therefore forego more interest, than high-wage workers.

Third, labor participation tends to fall as spousal earnings increase. As a result, couples with a high-wage worker are more likely to gain from Social Security's spousal benefit provision, which provides the nonworking spouse with benefits equal to 50 percent of those the working spouse receives.

Two recent studies taking these and other related factors into consideration suggest that Social Security may actually transfer wealth from low-wage to high-wage workers. **A** study using data from the Social Security Administration and the Health and Retirement Study found that when Social Security benefits are assessed for family units, rather than for individuals, the progressiveness of the system disappears. Another study adjusted for differences in mortality rates, patterns of lifetime income, and other factors. It found that if a *2* percent real interest rate (discount rate) is used to evaluate the pattern of taxes paid and benefits received, the redistributive effects of Social Security are essentially neutral. However, at a more realistic 4 percent real interest rate, Social Security actually favors higher-income households.[6]

Social Security Adversely Affects Blacks and Other Groups with Below-average Life Expectancy

Currently, the average retiree reaching age 65 can expect to spend 18 years receiving Social Security benefits, after more than 40 years of paying into the system. But what about those people who do not make it into their 80s, or even to the normal retirement age of 65? Unlike private financial assets, Social Security benefits cannot be passed on to heirs. Thus, those who die before age 65 or soon thereafter, receive little or nothing from their payroll tax payments.

Social Security was not set up to transfer income from some ethnic groups to others, but under its current structure, it does so, nonetheless. Because of the shorter life expectancies of black Americans the Social Security system adversely affects their economic welfare. Compared to whites and Hispanics, blacks are far more likely to pay a lifetime of payroll taxes and then die without receiving much in the way of benefits. Thus, the system works to their disadvantage. On the other hand, Social Security is particularly favorable to Hispanics because of their above-average life expectancy and the progressive nature of the benefit formula. As a result, Hispanics derive a higher return than whites and substantially higher than blacks.[7]

Exhibit 5 presents the expected real returns for people born in 1975, according to gender, marital status, and ethnicity.[8] Single black men who were born in 1975 can expect to derive a real annual return of *negative* 1.3 percent on their Social Security tax payments, compared to returns of 0.2 percent for single white men and 1.6 percent for single

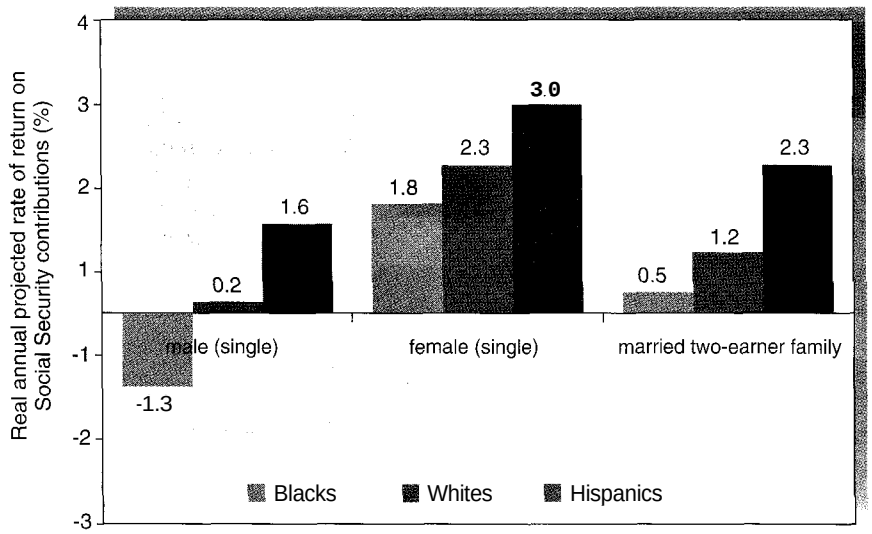
[6]See Alan Gustman and Thomas Steinmeier, "How Effective Is Redistribution Under The Social Security Benefit Formula?' *Journal of Public Economics* 82 (October 2001): 1–28; and Julia Lynn Coronado, Don Fullerton, and Thomas Glass, "Long Run Effects of Social Security Reform Proposals on Lifetime Progressivity," in *The Distributional Aspects of Social Security and Social Security Reform,* eds. Martin Feldstein and Jeffrey B. Liebman (Chicago: University of Chicago Press, 2002).

[7]For additional details on the redistributive effects of Social Security across ethnic groups, see William W. Beach and Gareth Davis, "More for Your Money: Improving Social Security's Rate of Return," in *Improving Retirement Security: A Handbook for Reformers,* ed. David C. John, (Washington DC: Heritage Foundation, 2000) 25–64 and Martin Feldstein and Jeffrey Liebman, "The Distributional Effects of an Investment-based Social Security System," in *The Distributional Aspects of Social Security and Social Security Reform,* eds. Martin Feldstein and Jeffrey B. Liebman (Chicago: University of Chicago Press, 2002).

[8] It is common to calculate **a** rate of return on financial investments by comparing initial investments with the stream of projected future income (or benefits). Social Security is not like a regular financial investment since there is no accumulation of assets and no legal right to benefits. Nonetheless, a rate of return can be calculated hy comparing the payroll taxes **a** worker pays with the future benefits he or she is promised. The rate-of-return figures of Exhibit 5 were derived in this manner. They assume that the current tax level and promised future benefits will he maintained. However, as we previously noted, projections indicate that current tax rates will cover only about three-fourths of promised benefits by 2030. Thus, higher taxes will he required to maintain the promised benefit levels. In turn, the higher taxes will lower rates of return. Therefore, the figures of Exhibit 5 probably overstate the rates of return for the various groups.

EXHIBIT 5
Rates of Return by Gender, Marital Status, and Ethnicity

The earnings of blacks are lower than whites, but their life expectancies are shorter. The latter effect dominates when it comes to Social Security payouts. As a result, blacks get a lower rate of return from Social Security than whites. On the other hand, the earnings of Hispanics are lower than whites', but their life expectancies are a little longer. Consequently, their returns from Social Security are higher than whites' and substantially higher than blacks'.



Source: Center for Data Analysis, Heritage Foundation.

Hispanic men. Similarly, a two-earner black couple born in 1975 can expect a real return of 0.5 percent, compared to returns of 1.2 percent and 2.3 percent for white and Hispanic couples born during the same year. A similar pattern exists when comparisons are made for people born in other years.

The Social Security retirement system also works to the disadvantage of those with life-shortening diseases. People with diabetes, heart disease, AIDS, and other diseases often spend decades paying 10.6 percent of their earnings into the system, only to die with loved ones unable to receive benefits from the Social Security taxes they have paid. (People with life-shortening diseases may receive disability insurance, but if they die before retirement, they collect nothing from their payments into the retirement system.)

Discrimination against Working Women

When Social Security was established, relatively few married women worked outside the home. Therefore, people were permitted to receive benefits based on either their own earnings or 50 percent of the benefits earned by their spouse, whichever was greater. This provision, which is still in place today, imposes a heavy penalty on women in the workforce. In the case of many working married women, the benefits based on the earnings of their spouses are approximately equal to, or in some cases greater than, benefits based on their own earnings. Thus, the payroll tax takes a big chunk of the earnings of many working women without providing them with any significant additional benefits.

PERSONAL RETIREMENT ACCOUNTS AND SOCIAL SECURITY REFORM

When the number of workers is growing rapidly and each successive generation is larger than the one that preceded it, a pay-as-you-go system can work well and yield a reasonable return. But we are now in an era when the number of retirees is growing more

© BY PERMISSION OF JOHN L. HART FLP, AND CREATORS SYNDICATE INC.

THE WIZARD OF ID

The projected real rate of return on Social Security taxes for persons born after 1950 is approximately 2 percent, far below the real rate of return on private-sector investments.

rapidly than the number of workers. In this environment, Social Security is not a good investment. As we previously mentioned, today's typical worker can expect a return of only 2.0 percent from the taxes paid into the system. Social Security uses a worker's highest 35 earning years to calculate benefits. Going back to 1873 (this is as far back as calculations can be made), the average real rate of return derived from a fixed annual payment over a 35-year period into the stock market has been 6.4 percent. The highest return over a 35-year period was 9.5, while the lowest was 2.7 percent.[7] Even this latter figure is higher than can be expected from Social Security. Social Security's low rate of return, along with its structure that sometimes seems outdated, has fostered an environment for change.

When the number of workers is small relative to the number of retirees, a system in which people finance their own retirement by saving and investing during their working years becomes more attractive. In varying degrees, several countries have already moved toward systems based on personal retirement accounts (PRAs). Beginning in the early 1980s, Chile shifted to a retirement system based on saving and investing through PRAs rather than pay-as-you-go. The Chilean plan was so successful that other Latin American countries, including Mexico, Bolivia, Colombia, and Peru, adopted similar plans in the 1990s. High-income countries have also moved in this direction. In 1986, the United Kingdom began allowing workers to channel 4.6 percentage points of their payroll tax into PRAs in exchange for accepting a lower level of benefits from the pay-as-you-go system. The PRA option is highly popular. Three-fourths of British workers now choose it. Other countries that now permit at least some substitution of PRAs for payroll taxes and pay-as-you-go benefits include the Netherlands, Australia, Sweden, and Germany.

A shift to a retirement security system based, at least partially, on personal saving accounts raises a number of issues. We will mention four of the most important.

1. Degree of Investor Choice Historical evidence indicates that, when held over a lengthy time period, a diverse holding of stocks is a low-risk investment that will generally outperform government bonds. (See the following Special Topic feature.) Furthermore, mutual funds now make it possible for even a small, novice investor to hold a diverse stock portfolio while still keeping administrative costs low. However, there is no assurance that investors, if left to their own discretion will choose to hold a diverse portfolio. Some proponents of PRAs say that providing individuals with a good deal of discretion concerning how their funds are invested is a good idea. Others say it is necessary to restrict people's choices in order to ensure that their retirement funds are not squandered on risky investments. The issue here is primarily about striking a balance between an attractive rate of return and minimal risk that the funds will be wiped out if invested unwisely or if the stock market were to crash.

[7] See Liqun Liu, Andrew J. Rettenmaier, and Zijun Wang, "Social Security and Stock Market Risk," NCPA Policy Report No. 244, National Center for Policy Analysis, July 23, 2001.

2. Share of Payroll Tax Allocated to PRAs and Current System How much of the payroll tax should workers be permitted to allocate into PRAs? Several plans would permit individuals to channel between 2 and 6 percentage points (out of the current 10.6 percent rate) into a PRA-based system in exchange for accepting lower benefits from the current system. Over time, other plans call for fully substituting a PRA-based system for the current pay-as-you-go system. Obviously, this issue is about the relative importance of the two systems in the future.

3. Protecting the Benefits of Current Retirees and People Near Retirement The benefits promised to current retirees and people near retirement age must be protected. Because most of the payroll tax is needed to finance current benefits, permitting workers to channel more than 1 or 2 percentage points into PRAs would create an immediate shortfall for the current system. Therefore, the transition to a system based more on personal investment may well require an increase in taxes and/or additional borrowing.

4. Property Rights and PRAs Most plans currently under consideration would give people the property rights to the funds in their accounts and allow them to be passed along to their heirs in case of death prior to their retirement. However, funds drawn from a PRA during one's retirement years would generally have to be converted to a lifetime annuity, an insurance instrument paying the retiree a regular income for the remainder of his or her life.

Policy makers are now searching for ways to provide income security for future retirees without having to increase payroll taxes to levels that will retard work incentives and endanger future economic growth. Personal retirement accounts may help achieve this objective. The disincentive effects of having to make payments into a PRA—which can enhance the retiree's income and be passed along to his or her heirs—are much less severe than those associated with higher taxes. Furthermore, PRAs will tend to encourage saving and investment, which will help grow the economy.

KEY POINTS

- Social Security does not follow the saving-and-investment model. Rather, most of the funds flowing into the system are paid out to current retirees and survivors.
- Although the current tax revenues exceed the payments to retirees, this will change dramatically as the baby boomers begin to move into the retirement phases of their lives. Beginning in about 2018, the system's current surplus will become a deficit, which will persist for several decades.
- The current surplus of the Social Security system is used to purchase U.S. Treasury bonds. However, because the federal government is both the payee and recipient of these bonds, their net asset value to the federal government is zero. The bonds will not reduce the level of future taxes needed to cover the Social Security deficit when the baby boomers begin to retire.
- The big problem with the current pay-as-you-go system is that large tax increases, spending cuts, and/or additional borrowing will be needed to cover the Social Security deficits following the retirement of the baby boom generation. Dealing with these deficits is likely to adversely affect the economy.
- Although the Social Security benefit formula favors those with lower lifetime earnings, low-wage workers have lower life expectancies, begin work at a younger age, and gain less from the spousal benefit provisions of the current system. These latter factors largely, if not entirely, offset the egalitarian effects of the benefit formula.
- Because of their shorter life expectancies, blacks get a lower rate of return from Social Security than whites and a substantially lower return than Hispanics.

CRITICAL ANALYSIS QUESTIONS

1. Is the Social Security system based on the same principles as private insurance? Why or why not?

*2. Why does the Social Security system face a crisis? **Are** there real assets in the Social Security Trust Fund that can be used to pay future benefits? Will the trust fund help avert higher future taxes and/or benefit reductions when the baby boomers retire? Why or why not?

3. Do you think workers should be permitted to invest all or part of their Social Security contribution in private investment funds? What are the advantages and disadvantages of a private option system? If given the opportunity, would you choose the private option or stay with the current system? Why?

4. How does Social Security affect the economic well-being of blacks relative to whites and Hispanics? Explain.

5. Does the current Social Security system promote income equality? Why or why not?

6. The Social Security payroll tax is split equally between the employee and the employer. Would it make any difference if the entire tax was imposed on employees? Would employees be helped if all the tax was imposed on employers? *(Hint:* You may want to consult the section on tax incidence in Chapter 4.)

Asterisk denotes questions for which answers are given in Appendix B