GLOBAL EDITION

Microeconomics

EIGHTH EDITION

Jeffrey M. Perloff

Pearson

Microeconomics

8th edition

Global edition

THE PEARSON SERIES IN ECONOMICS

Abel/Bernanke/Croushore
*Macroeconomics**†

Acemoglu/Laibson/List
*Economics**†

Bade/Parkin
*Foundations of Economics**†

Berck/Helfand
The Economics of the Environment

Bierman/Fernandez
Game Theory with Economic Applications

Blanchard
*Macroeconomics**†

Boyer
Principles of Transportation Economics

Brander/Perloff
*Managerial Economics and Strategy**†

Branson
Macroeconomic Theory and Policy

Bruce
Public Finance and the American Economy

Carlton/Perloff
Modern Industrial Organization†

Case/Fair/Oster
*Principles of Economics**†

Chapman
Environmental Economics: Theory, Application, and Policy

Daniels/VanHoose
International Monetary & Financial Economics

Downs
An Economic Theory of Democracy

Farnham
Economics for Managers†

Froyen
Macroeconomics: Theories and Policies

Fusfeld
The Age of the Economist

Gerber
*International Economics**†

Gordon
*Macroeconomics**†

Greene
Econometric Analysis†

Gregory/Stuart
Russian and Soviet Economic Performance and Structure

Hartwick/Olewiler
The Economics of Natural Resource Use

Heilbroner/Milberg
The Making of the Economic Society

Heyne/Boettke/Prychitko
The Economic Way of Thinking

Hubbard/O'Brien
*Economics**†

InEcon

*Money, Banking, and the Financial System**

Hubbard/O'Brien/Rafferty
*Macroeconomics**

Hughes/Cain
American Economic History

Husted/Melvin
International Economics

Jehle/Reny
Advanced Microeconomic Theory

Keat/Young/Erfle
Managerial Economics

Klein
Mathematical Methods for Economics

Krugman/Obstfeld/Melitz
*International Economics: Theory & Policy**†

Laidler
The Demand for Money

Lynn
Economic Development: Theory and Practice for a Divided World

Miller
*Economics Today**

Miller/Benjamin
The Economics of Macro Issues

Miller/Benjamin/North
The Economics of Public Issues

Mishkin
*The Economics of Money, Banking, and Financial Markets**†

*The Economics of Money, Banking, and Financial Markets, Business School Edition**†

*Macroeconomics: Policy and Practice**

Murray
Econometrics: A Modern Introduction

O'Sullivan/Sheffrin/Perez
*Economics: Principles, Applications and Tools**†

Parkin
*Economics**†

Perloff
*Microeconomics**†

*Microeconomics: Theory and Applications with Calculus**†

Pindyck/Rubinfeld
*Microeconomics**†

Riddell/Shackelford/Stamos/Schneider
Economics: A Tool for Critically Understanding Society

Roberts
The Choice: A Fable of Free Trade and Protection

Scherer
Industry Structure, Strategy, and Public Policy

Schiller
The Economics of Poverty and Discrimination

Sherman
Market Regulation

Stock/Watson
Introduction to Econometrics†

Studenmund
Using Econometrics: A Practical Guide†

Todaro/Smith
Economic Development

Walters/Walters/Appel/Callahan/Centanni/Maex/O'Neill
Econversations: Today's Students Discuss Today's Issues

Williamson
Macroeconomics†

*denotes MyLab Economics titles. Visit www.myeconlab.com to learn more.

†denotes the availability of a Global Edition.

Microeconomics

8th edition
Global edition

Jeffrey M. Perloff

University of California, Berkeley

Pearson

Harlow, England • London • New York • Boston • San Francisco • Toronto • Sydney • Dubai • Singapore • Hong Kong
Tokyo • Seoul • Taipei • New Delhi • Cape Town • Sao Paulo • Mexico City • Madrid • Amsterdam • Munich • Paris • Milan

For Alexx, Gregg, Laura, and Spenser

Vice President, Business Publishing: Donna Battista
Director of Portfolio Management: Adrienne D'Ambrosio
Senior Portfolio Manager: Christina Masturzo
Editorial Assistant: Courtney Paganelli
Senior Editor, Global Edition: Daniel Luiz
Editor, Global Edition: Punita Kaur Mann
Vice President, Product Marketing: Roxanne McCarley
Director of Strategic Marketing: Brad Parkins
Strategic Marketing Manager: Deborah Strickland
Product Marketer: Tricia Murphy
Field Marketing Manager: Ramona Elmer
Field Marketing Assistant: Kristen Compton
Product Marketing Assistant: Jessica Quazza
Vice President, Production and Digital Studio, Arts and Business: Etain O'Dea
Director of Production, Business: Jeff Holcomb
Managing Producer, Business: Alison Kalil
Content Producer: Kathryn Brightney
Content Producer, Global Edition: Nitin Shankar
Senior Manufacturing Controller, Global Edition: Angela Hawksbee
Operations Specialist: Carol Melville
Creative Director: Blair Brown
Manager, Learning Tools: Brian Surette
Managing Producer, Digital Studio, Arts and Business: Diane Lombardo
Digital Studio Producer: Melissa Honig
Digital Studio Producer: Alana Coles
Digital Content Team Lead: Noel Lotz
Digital Content Project Lead: Courtney Kamauf
Media Editor, Global Edition: Naina Singh
Full-Service Project Management and Composition: SPi Global
Project Management and Text Design: Gillian Hall, The Aardvark Group
Design Manager: Laurie Entringer
Cover Design: Lumina Datamatics Ltd.
Cover Art: FOTOGRIN/Shutterstock

Pearson Education Limited
KAO Two
KAO Park
Harlow
CM17 9NA
United Kingdom

and Associated Companies throughout the world

Visit us on the World Wide Web at: www.pearsonglobaleditions.com

ISBN 10: 1-292-21562-3
ISBN 13: 978-1-292-21562-4

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Print edition typeset in 10/11.5pt Sabon LT Pro by SPi Global
Printed and bound by Vivar in Malaysia

Brief Contents

Chapter 1	Introduction	25
Chapter 2	Supply and Demand	33
Chapter 3	Applying the Supply-and-Demand Model	67
Chapter 4	Consumer Choice	97
Chapter 5	Applying Consumer Theory	134
Chapter 6	Firms and Production	172
Chapter 7	Costs	203
Chapter 8	Competitive Firms and Markets	245
Chapter 9	Applying the Competitive Model	287
Chapter 10	General Equilibrium and Economic Welfare	331
Chapter 11	Monopoly	367
Chapter 12	Pricing and Advertising	406
Chapter 13	Oligopoly and Monopolistic Competition	448
Chapter 14	Game Theory	490
Chapter 15	Factor Markets	528
Chapter 16	Interest Rates, Investments, and Capital Markets	553
Chapter 17	Uncertainty	585
Chapter 18	Externalities, Open-Access, and Public Goods	619
Chapter 19	Asymmetric Information	649
Chapter 20	Contracts and Moral Hazards	677
Chapter Appendixes		707
Answers to Selected Questions and Problems		735
Sources for Challenges and Applications		753
References		767
Definitions		777
Index		782
Credits		812

Contents

Chapter 1 Introduction 25

1.1 **Microeconomics: The Allocation of Scarce Resources** 25
- Trade-Offs 26
- Who Makes the Decisions 26
- Prices Determine Allocations 26
- APPLICATION *Twinkie Tax* 26

1.2 **Models** 27
- APPLICATION *Income Threshold Model and China* 27
- Simplifications by Assumption 28
- Testing Theories 28
- Maximizing Subject to Constraints 29
- Positive Versus Normative 29

1.3 **Uses of Microeconomic Models** 31
- *Summary* 31

Chapter 2 Supply and Demand 33

- CHALLENGE *Quantities and Prices of Genetically Modified Foods* 33

2.1 **Demand** 34
- The Demand Curve 35
- APPLICATION *Calorie Counting* 38
- The Demand Function 38
- Solved Problem 2.1 40
- Summing Demand Curves 41
- APPLICATION *Aggregating Corn Demand Curves* 41

2.2 **Supply** 42
- The Supply Curve 42
- The Supply Function 44
- Summing Supply Curves 45
- How Government Import Policies Affect Supply Curves 45
- Solved Problem 2.2 46

2.3 **Market Equilibrium** 47
- Using a Graph to Determine the Equilibrium 47
- Using Math to Determine the Equilibrium 47
- Forces That Drive the Market to Equilibrium 49

2.4 **Shocking the Equilibrium** 50
- Effects of a Shock to the Supply Curve 50
- Solved Problem 2.3 51
- Effects of a Shock to the Demand Curve 52

2.5 **Effects of Government Interventions** 52
- Policies That Shift Supply Curves 52
- APPLICATION *Occupational Licensing* 52
- Solved Problem 2.4 54
- Policies That Cause the Quantity Demanded to Differ from the Quantity Supplied 55
- APPLICATION Venezuelan Price Ceilings and Shortages 57
- Solved Problem 2.5 58
- Why the Quantity Supplied Need Not Equal the Quantity Demanded 59

2.6 **When to Use the Supply-and-Demand Model** 60
- CHALLENGE SOLUTION *Quantities and Prices of Genetically Modified Foods* 61
- *Summary* 62 ■ *Questions* 63

Chapter 3 Applying the Supply-and-Demand Model 67

- CHALLENGE *Who Pays the Gasoline Tax?* 67

3.1 **How Shapes of Supply and Demand Curves Matter** 68

3.2 **Sensitivity of the Quantity Demanded to Price** 69
- Price Elasticity of Demand 69
- Solved Problem 3.1 70
- APPLICATION *The Demand Elasticities for Google Play and Apple Apps* 71
- Elasticity Along the Demand Curve 71
- Demand Elasticity and Revenue 74
- Solved Problem 3.2 74
- APPLICATION *Amazon Prime* 75
- Demand Elasticities over Time 75
- Other Demand Elasticities 76
- APPLICATION *Anti-Smoking Policies May Reduce Drunk Driving* 77

3.3 **Sensitivity of the Quantity Supplied to Price** 78
- Elasticity of Supply 78
- Elasticity Along the Supply Curve 79
- Supply Elasticities over Time 80
- APPLICATION *Oil Drilling in the Arctic National Wildlife Refuge* 80
- Solved Problem 3.3 81

3.4 **Effects of a Sales Tax** 82
- Effects of a Specific Tax on the Equilibrium 83
- The Equilibrium Is the Same No Matter Whom the Government Taxes 84
- Solved Problem 3.4 85
- Firms and Customers Share the Burden of the Tax 85
- APPLICATION *Taxes to Prevent Obesity* 87

Solved Problem 3.5 87
Ad Valorem and Specific Taxes Have Similar Effects 88
Solved Problem 3.6 89
Subsidies 90
APPLICATION *Subsidizing Ethanol* 90
CHALLENGE SOLUTION *Who Pays the Gasoline Tax?* 91
Summary 92 ■ *Questions* 93

Chapter 4 Consumer Choice 97

CHALLENGE *Why Americans Buy More Ebooks Than Do Germans* 97
4.1 Preferences 99
Properties of Consumer Preferences 99
APPLICATION *You Can't Have Too Much Money* 100
Preference Maps 101
Solved Problem 4.1 104
APPLICATION *Indifference Curves Between Food and Clothing* 107
4.2 Utility 108
Utility Function 108
Ordinal Preferences 108
Utility and Indifference Curves 109
Marginal Utility 110
Utility and Marginal Rates of Substitution 111
4.3 Budget Constraint 112
Slope of the Budget Constraint 114
Solved Problem 4.2 114
Effect of a Change in Price on the Opportunity Set 115
Effect of a Change in Income on the Opportunity Set 116
Solved Problem 4.3 116
4.4 Constrained Consumer Choice 116
The Consumer's Optimal Bundle 117
APPLICATION *Substituting Alcohol for Marijuana* 119
Solved Problem 4.4 119
Solved Problem 4.5 120
★ Optimal Bundles on Convex Sections of Indifference Curves 121
Buying Where More Is Better 122
Food Stamps 122
APPLICATION *Benefiting from Food Stamps* 124
4.5 Behavioral Economics 125
Tests of Transitivity 125
Endowment Effect 125
APPLICATION *Opt In Versus Opt Out* 126
Salience and Bounded Rationality 127
APPLICATION *Unaware of Taxes* 128
CHALLENGE SOLUTION *Why Americans Buy More Ebooks Than Do Germans* 128
Summary 129 ■ *Questions* 130

Chapter 5 Applying Consumer Theory 134

CHALLENGE *Per-Hour Versus Lump-Sum Childcare Subsidies* 134
5.1 Deriving Demand Curves 135
Indifference Curves and a Rotating Budget Line 136
Price-Consumption Curve 137
APPLICATION *Smoking Versus Eating and Phoning* 138
The Demand Curve Corresponds to the Price-Consumption Curve 138
Solved Problem 5.1 139
5.2 How Changes in Income Shift Demand Curves 139
Effects of a Rise in Income 140
Solved Problem 5.2 141
Consumer Theory and Income Elasticities 142
APPLICATION *Fast-Food Engel Curve* 145
5.3 Effects of a Price Change 146
Income and Substitution Effects with a Normal Good 146
Solved Problem 5.3 148
Solved Problem 5.4 149
Income and Substitution Effects with an Inferior Good 150
Solved Problem 5.5 150
★ Compensating Variation and Equivalent Variation 151
APPLICATION *What's Your Smart Phone Worth to You?* 152
5.4 Cost-of-Living Adjustments 152
Inflation Indexes 152
Effects of Inflation Adjustments 154
APPLICATION *Paying Employees to Relocate* 156
5.5 Deriving Labor Supply Curves 158
Labor-Leisure Choice 158
Income and Substitution Effects 160
Solved Problem 5.6 161
Shape of the Labor Supply Curve 162
APPLICATION *Working After Winning the Lottery* 163
Income Tax Rates and Labor Supply 164
CHALLENGE SOLUTION *Per-Hour Versus Lump-Sum Childcare Subsidies* 166
Summary 167 ■ *Questions* 168

Chapter 6 Firms and Production 172

CHALLENGE *More Productive Workers During Downturns* 172
6.1 The Ownership and Management of Firms 173
Private, Public, and Nonprofit Firms 173
APPLICATION *Chinese State-Owned Enterprises* 174
The Ownership of For-Profit Firms 174
The Management of Firms 175
What Owners Want 175

6.2 **Production** **176**
Production Functions 176
Varying Inputs over Time 176
6.3 **Short-Run Production** **177**
Total Product 177
Marginal Product of Labor 178
Solved Problem 6.1 179
Average Product of Labor 179
Graphing the Product Curves 179
Law of Diminishing Marginal Returns 181
APPLICATION *Malthus and the Green Revolution* 182
6.4 **Long-Run Production** **183**
Isoquants 184
APPLICATION *A Semiconductor Integrated Circuit Isoquant* 187
Substituting Inputs 188
Solved Problem 6.2 189
6.5 **Returns to Scale** **191**
Constant, Increasing, and Decreasing Returns to Scale 191
Solved Problem 6.3 192
APPLICATION *Returns to Scale in Various Industries* 192
Varying Returns to Scale 194
6.6 **Productivity and Technical Change** **195**
Relative Productivity 195
Innovations 196
APPLICATION *Robots and the Food You Eat* 197
APPLICATION *A Good Boss Raises Productivity* 198
CHALLENGE SOLUTION *Labor Productivity During Downturns* 198
Summary 199 ■ *Questions* 200

Chapter 7 Costs 203

CHALLENGE *Technology Choice at Home Versus Abroad* 203
7.1 **The Nature of Costs** **204**
Opportunity Costs 204
APPLICATION *The Opportunity Cost of an MBA* 205
Solved Problem 7.1 206
Opportunity Cost of Capital 206
Sunk Costs 207
7.2 **Short-Run Costs** **208**
Short-Run Cost Measures 208
APPLICATION *The Sharing Economy and the Short Run* 209
Short-Run Cost Curves 210
Production Functions and the Shape of Cost Curves 212
APPLICATION *A Beer Manufacturer's Short-Run Cost Curves* 215
Effects of Taxes on Costs 215
Solved Problem 7.2 216
Short-Run Cost Summary 217
7.3 **Long-Run Costs** **218**
All Costs Are Avoidable in the Long Run 218
Minimizing Cost 218
Isocost Line 219
Combining Cost and Production Information 221
Solved Problem 7.3 222
Factor Price Changes 224
Solved Problem 7.4 225
The Long-Run Expansion Path and the Long-Run Cost Function 225
Solved Problem 7.5 227
The Shape of Long-Run Cost Curves 227
APPLICATION *3D Printing* 230
Estimating Cost Curves Versus Introspection 230
7.4 **Lower Costs in the Long Run** **231**
Long-Run Average Cost as the Envelope of Short-Run Average Cost Curves 231
APPLICATION *A Beer Manufacturer's Long-Run Cost Curves* 232
APPLICATION *Should You Buy an Inkjet or a Laser Printer?* 233
Short-Run and Long-Run Expansion Paths 234
The Learning Curve 235
7.5 **Cost of Producing Multiple Goods** **236**
APPLICATION *Medical Economies of Scope* 238
CHALLENGE SOLUTION *Technology Choice at Home Versus Abroad* 238
Summary 239 ■ *Questions* 240

Chapter 8 Competitive Firms and Markets 245

CHALLENGE *The Rising Cost of Keeping On Truckin'* 245
8.1 **Perfect Competition** **246**
Price Taking 246
Why the Firm's Demand Curve Is Horizontal 247
Deviations from Perfect Competition 248
Derivation of a Competitive Firm's Demand Curve 249
Solved Problem 8.1 250
Why We Study Perfect Competition 251
8.2 **Profit Maximization** **251**
Profit 251
Two Decisions for Maximizing Profit 253
8.3 **Competition in the Short Run** **255**
Short-Run Output Decision 256
Solved Problem 8.2 258
Short-Run Shutdown Decision 259
APPLICATION *Fracking and Shutdowns* 261
Solved Problem 8.3 262
Short-Run Firm Supply Curve 262
Short-Run Market Supply Curve 262
Short-Run Competitive Equilibrium 265
Solved Problem 8.4 266
8.4 **Competition in the Long Run** **267**
Long-Run Competitive Profit Maximization 267
Long-Run Firm Supply Curve 268

APPLICATION *The Size of Ethanol Processing Plants* 269
Long-Run Market Supply Curve 269
APPLICATION *Entry and Exit of Solar Power Firms* 270
APPLICATION *Upward-Sloping Long-Run Supply Curve for Cotton* 273
APPLICATION *Reformulated Gasoline Supply Curves* 277
Solved Problem 8.5 278
Long-Run Competitive Equilibrium 279
CHALLENGE SOLUTION *The Rising Cost of Keeping On Truckin'* 280
Summary 281 ■ *Questions* 282

Chapter 9 Applying the Competitive Model 287

CHALLENGE *Liquor Licenses* 287
9.1 **Zero Profit for Competitive Firms in the Long Run** 288
Zero Long-Run Profit with Free Entry 288
Zero Long-Run Profit When Entry Is Limited 289
APPLICATION *What's a Name Worth?* 291
The Need to Maximize Profit 291
9.2 **Consumer Welfare** 291
Measuring Consumer Welfare Using a Demand Curve 292
APPLICATION *Willingness to Pay on eBay* 294
Effect of a Price Change on Consumer Surplus 295
APPLICATION *Goods with a Large Consumer Surplus Loss from Price Increases* 296
Solved Problem 9.1 297
9.3 **Producer Welfare** 297
Measuring Producer Surplus Using a Supply Curve 298
Using Producer Surplus 299
Solved Problem 9.2 300
9.4 **Competition Maximizes Welfare** 300
Solved Problem 9.3 302
APPLICATION *Deadweight Loss of Christmas Presents* 304
9.5 **Policies That Shift Supply and Demand Curves** 305
Entry Barrier 306
APPLICATION *Welfare Effects of Allowing Fracking* 306
9.6 **Policies That Create a Wedge Between Supply and Demand** 307
Welfare Effects of a Sales Tax 307
APPLICATION *The Deadweight Loss from Gas Taxes* 309
Solved Problem 9.4 309
Welfare Effects of a Subsidy 310
Solved Problem 9.5 311
Welfare Effects of a Price Floor 312
Solved Problem 9.6 314
APPLICATION *How Big Are Farm Subsidies and Who Gets Them?* 315
Welfare Effects of a Price Ceiling 316
Solved Problem 9.7 316
APPLICATION *The Social Cost of a Natural Gas Price Ceiling* 317
9.7 **Comparing Both Types of Policies: Imports** 318
Free Trade Versus a Ban on Imports 318
APPLICATION *Russian Food Ban* 320
Free Trade Versus a Tariff 320
Free Trade Versus a Quota 323
Rent Seeking 323
CHALLENGE SOLUTION *Liquor Licenses* 324
Summary 326 ■ *Questions* 327

Chapter 10 General Equilibrium and Economic Welfare 331

CHALLENGE *Anti-Price Gouging Laws* 331
10.1 **General Equilibrium** 333
Feedback Between Competitive Markets 334
Solved Problem 10.1 336
Minimum Wages with Incomplete Coverage 337
Solved Problem 10.2 339
APPLICATION *Urban Flight* 340
10.2 **Trading Between Two People** 340
Endowments 340
Mutually Beneficial Trades 342
Solved Problem 10.3 344
Bargaining Ability 344
10.3 **Competitive Exchange** 344
Competitive Equilibrium 345
The Efficiency of Competition 346
Obtaining Any Efficient Allocation Using Competition 347
10.4 **Production and Trading** 347
Comparative Advantage 347
Solved Problem 10.4 349
Efficient Product Mix 351
Competition 351
10.5 **Efficiency and Equity** 353
Role of the Government 353
APPLICATION *The 1% Grow Wealthier* 354
Efficiency 356
Equity 356
APPLICATION *How You Vote Matters* 359
Efficiency Versus Equity 361
CHALLENGE SOLUTION *Anti-Price Gouging Laws* 362
Summary 363 ■ *Questions* 363

Chapter 11 Monopoly 367

CHALLENGE *Brand-Name and Generic Drugs* 367
11.1 **Monopoly Profit Maximization** 368
Marginal Revenue 369
Solved Problem 11.1 371
Choosing Price or Quantity 373
Graphical Approach 373
Mathematical Approach 375
APPLICATION *Apple's iPad* 376

Solved Problem 11.2 376
Effects of a Shift of the Demand Curve 377
11.2 Market Power 378
Market Power and the Shape of the Demand Curve 379
APPLICATION *Cable Cars and Profit Maximization* 380
Lerner Index 381
Solved Problem 11.3 381
Sources of Market Power 381
11.3 Market Failure Due to Monopoly Pricing 382
Solved Problem 11.4 384
11.4 Causes of Monopoly 385
Cost-Based Monopoly 385
Solved Problem 11.5 387
Government Creation of a Monopoly 387
APPLICATION *Botox Patent Monopoly* 388
11.5 Government Actions That Reduce Market Power 390
Regulating Monopolies 390
Solved Problem 11.6 393
APPLICATION *Natural Gas Regulation* 394
Increasing Competition 395
APPLICATION *Movie Studios Attacked by 3D Printers!* 395
Solved Problem 11.7 396
11.6 Networks, Dynamics, and Behavioral Economics 397
Network Externalities 397
Network Externalities and Behavioral Economics 398
Network Externalities as an Explanation for Monopolies 398
APPLICATION *eBay's Critical Mass* 399
A Two-Period Monopoly Model 399
CHALLENGE SOLUTION *Brand-Name and Generic Drugs* 400
Summary 401 ■ *Questions* 402

Chapter 12 Pricing and Advertising 406

CHALLENGE *Sale Prices* 406
12.1 Conditions for Price Discrimination 408
Why Price Discrimination Pays 408
APPLICATION *Disneyland Pricing* 410
Which Firms Can Price Discriminate 410
Preventing Resale 411
APPLICATION *Preventing Resale of Designer Bags* 411
Not All Price Differences Are Price Discrimination 412
Types of Price Discrimination 412
12.2 Perfect Price Discrimination 413
How a Firm Perfectly Price Discriminates 413
Perfect Price Discrimination Is Efficient but Harms Some Consumers 414
APPLICATION *Botox and Price Discrimination* 416
Solved Problem 12.1 417
APPLICATION *Google Uses Bidding for Ads to Price Discriminate* 417
Transaction Costs and Perfect Price Discrimination 418
12.3 Group Price Discrimination 418
APPLICATION *Warner Brothers Sets Prices for a* Harry Potter *DVD* 419
Solved Problem 12.2 420
Prices and Elasticities 421
Solved Problem 12.3 421
Preventing Resale 422
APPLICATION *Reselling Textbooks* 422
Solved Problem 12.4 422
Identifying Groups 424
APPLICATION *Buying Discounts* 425
Welfare Effects of Group Price Discrimination 426
12.4 Nonlinear Price Discrimination 427
12.5 Two-Part Pricing 429
Two-Part Pricing with Identical Customers 429
Two-Part Pricing with Nonidentical Consumers 430
APPLICATION *iTunes for a Song* 432
12.6 Tie-In Sales 433
Requirement Tie-In Sale 433
APPLICATION *Ties That Bind* 433
Bundling 434
Solved Problem 12.5 436
12.7 Advertising 437
The Decision Whether to Advertise 437
How Much to Advertise 438
APPLICATION *Super Bowl Commercials* 439
CHALLENGE SOLUTION *Sale Prices* 440
Summary 441 ■ *Questions* 442

Chapter 13 Oligopoly and Monopolistic Competition 448

CHALLENGE *Government Aircraft Subsidies* 448
13.1 Market Structures 450
13.2 Cartels 451
Why Cartels Form 451
Why Cartels Fail 453
Laws Against Cartels 453
APPLICATION *The Apple-Google-Intel-Adobe-Intuit-Lucasfilms-Pixar Wage Cartel* 455
Maintaining Cartels 455
APPLICATION *Cheating on the Maple Syrup Cartel* 456
Mergers 457
APPLICATION *Mergers to Monopolize* 458
13.3 Cournot Oligopoly 458
The Duopoly Nash-Cournot Equilibrium 459
Equilibrium, Elasticity, and the Number of Firms 463
APPLICATION *Mobile Number Portability* 465
Nonidentical Firms 465
Solved Problem 13.1 467
Solved Problem 13.2 468

13.4	**Stackelberg Oligopoly**	**469**
	Graphical Model	470
	Solved Problem 13.3	471
	Why Moving Sequentially Is Essential	472
	Comparison of Competitive, Stackelberg, Cournot, and Collusive Equilibria	472
13.5	**Bertrand Oligopoly**	**474**
	Identical Products	474
	Differentiated Products	475
	APPLICATION *Bottled Water*	477
13.6	**Monopolistic Competition**	**478**
	APPLICATION *Monopolistically Competitive Food Truck Market*	478
	Equilibrium	479
	Solved Problem 13.4	480
	Fixed Costs and the Number of Firms	481
	Solved Problem 13.5	482
	APPLICATION *Zoning Laws as a Barrier to Entry by Hotel Chains*	482
	CHALLENGE SOLUTION *Government Aircraft Subsidies*	483
	Summary 485 ■ *Questions* 485	

Chapter 14 Game Theory 490

	CHALLENGE *Intel and AMD's Advertising Strategies*	490
14.1	**Static Games**	**492**
	Normal-Form Games	493
	Failure to Maximize Joint Profits	497
	APPLICATION *Strategic Advertising*	499
	Multiple Equilibria	500
	Solved Problem 14.1	500
	Mixed Strategies	501
	APPLICATION *Tough Love*	503
	Solved Problem 14.2	503
14.2	**Repeated Dynamic Games**	**504**
	Strategies and Actions in Dynamic Games	504
	Cooperation in a Repeated Prisoners' Dilemma Game	505
	Solved Problem 14.3	506
14.3	**Sequential Dynamic Games**	**506**
	Game Tree	506
	Subgame Perfect Nash Equilibrium	508
	Credibility	509
	Dynamic Entry Game	510
	Solved Problem 14.4	512
	APPLICATION *Keeping Out Casinos*	513
	Solved Problem 14.5	513
14.4	**Auctions**	**514**
	Elements of Auctions	515
	Bidding Strategies in Private-Value Auctions	516
	Winner's Curse	517
	APPLICATION *Bidder's Curse*	518
14.5	**Behavioral Game Theory**	**519**
	APPLICATION *GM's Ultimatum*	519
	An Experiment	519
	Reciprocity	520
	CHALLENGE SOLUTION *Intel and AMD's Advertising Strategies*	520
	Summary 521 ■ *Questions* 522	

Chapter 15 Factor Markets 528

	CHALLENGE *Athletes' Salaries and Ticket Prices*	528
15.1	**Competitive Factor Market**	**529**
	Short-Run Factor Demand of a Firm	529
	Solved Problem 15.1	532
	Solved Problem 15.2	534
	Long-Run Factor Demand	534
	Factor Market Demand	535
	Competitive Factor Market Equilibrium	537
15.2	**Effects of Monopolies on Factor Markets**	**538**
	Market Structure and Factor Demands	538
	A Model of Market Power in Input and Output Markets	539
	APPLICATION *Unions and Profits*	542
	Solved Problem 15.3	543
15.3	**Monopsony**	**544**
	Monopsony Profit Maximization	544
	APPLICATION *Monopsony and the Gender Wage Gap*	546
	Welfare Effects of Monopsony	546
	Solved Problem 15.4	547
	CHALLENGE SOLUTION *Athletes' Salaries and Ticket Prices*	548
	Summary 549 ■ *Questions* 550	

Chapter 16 Interest Rates, Investments, and Capital Markets 553

	CHALLENGE *Does Going to College Pay?*	553
16.1	**Comparing Money Today to Money in the Future**	**554**
	Interest Rates	555
	Using Interest Rates to Connect the Present and Future	557
	APPLICATION *Power of Compounding*	557
	Stream of Payments	558
	Solved Problem 16.1	560
	APPLICATION *Saving for Retirement*	561
	Inflation and Discounting	562
	APPLICATION *Winning the Lottery*	564
16.2	**Choices over Time**	**564**
	Investing	565
	Solved Problem 16.2	566
	Solved Problem 16.3	567
	Rate of Return on Bonds	567
	★ Behavioral Economics: Time-Varying Discounting	568
	APPLICATION *Falling Discount Rates and Self-Control*	569
16.3	**Exhaustible Resources**	**570**
	When to Sell an Exhaustible Resource	570

Price of a Scarce Exhaustible Resource 571
APPLICATION *Redwood Trees* 574
Why Price May Be Constant or Fall 574
16.4 Capital Markets, Interest Rates, and Investments 576
Solved Problem 16.4 577
CHALLENGE SOLUTION *Does Going to College Pay?* 578
Summary 580 ■ *Questions* 580

Chapter 17 Uncertainty 585

CHALLENGE *BP and Limited Liability* 585
17.1 Assessing Risk 587
Probability 587
Expected Value 588
Solved Problem 17.1 589
Variance and Standard Deviation 590
17.2 Attitudes Toward Risk 591
Expected Utility 591
Risk Aversion 592
Solved Problem 17.2 594
APPLICATION *Stocks' Risk Premium* 595
Risk Neutrality 595
Risk Preference 596
APPLICATION *Gambling* 596
17.3 Reducing Risk 597
Just Say No 598
Obtain and Use Information 598
Diversify 598
APPLICATION *Failure to Diversify* 600
Buy Insurance 601
Solved Problem 17.3 602
APPLICATION *Flight Insurance* 603
APPLICATION *Limited Insurance for Natural Disasters* 604
17.4 Investing Under Uncertainty 605
Risk-Neutral Investing 606
Risk-Averse Investing 607
Solved Problem 17.4 607
17.5 Behavioral Economics of Uncertainty 608
Biased Assessment of Probabilities 608
APPLICATION *Biased Estimates* 609
Violations of Expected Utility Theory 609
Prospect Theory 611
APPLICATION *Loss Aversion Contracts* 613
CHALLENGE SOLUTION *BP and Limited Liability* 613
Summary 614 ■ *Questions* 615

Chapter 18 Externalities, Open-Access, and Public Goods 619

CHALLENGE *Trade and Pollution* 619
18.1 Externalities 620
APPLICATION *Negative Externalities from Spam* 621
18.2 The Inefficiency of Competition with Externalities 621
APPLICATION *Global Warming* 624
18.3 Regulating Externalities 624
Emissions Standard 626
Emissions Fee 626
Solved Problem 18.1 627
APPLICATION *Why Tax Drivers* 628
Benefits Versus Costs from Controlling Pollution 628
APPLICATION *Protecting Babies* 628
18.4 Market Structure and Externalities 629
Monopoly and Externalities 629
Monopoly Versus Competitive Welfare with Externalities 630
Solved Problem 18.2 631
Taxing Externalities in Noncompetitive Markets 631
18.5 Allocating Property Rights to Reduce Externalities 631
Coase Theorem 632
APPLICATION *Buying a Town* 633
Markets for Pollution 634
APPLICATION *Acid Rain Cap-and-Trade Program* 634
Markets for Positive Externalities 635
18.6 Rivalry and Exclusion 635
Open-Access Common Property 635
APPLICATION *Road Congestion* 636
Club Goods 637
APPLICATION *Software Piracy* 637
Public Goods 638
Solved Problem 18.3 639
APPLICATION *Free Riding on Measles Vaccinations* 640
APPLICATION *What's Their Beef?* 641
CHALLENGE SOLUTION *Trade and Pollution* 643
Summary 644 ■ *Questions* 645

Chapter 19 Asymmetric Information 649

CHALLENGE *Dying to Work* 649
19.1 Adverse Selection 651
Insurance Markets 651
Products of Unknown Quality 652
Solved Problem 19.1 655
Solved Problem 19.2 655
19.2 Reducing Adverse Selection 656
Equalizing Information 656
APPLICATION *Discounts for Data* 657
APPLICATION *Adverse Selection and Remanufactured Goods* 659
Restricting Opportunistic Behavior 660
19.3 Price Discrimination Due to False Beliefs About Quality 660

APPLICATION *Reducing Consumers' Information* 661
19.4 **Market Power from Price Ignorance** **661**
Tourist-Trap Model 662
Solved Problem 19.3 663
Advertising and Prices 664
19.5 **Problems Arising from Ignorance When Hiring** **664**
Cheap Talk 664
APPLICATION *Cheap Talk in eBay's Best Offer Market* 666
Education as a Signal 666
Solved Problem 19.4 668
Screening in Hiring 670
CHALLENGE SOLUTION *Dying to Work* 672
Summary 673 ■ *Questions* 674

Chapter 20 Contracts and Moral Hazards 677

CHALLENGE *Clawing Back Bonuses* 677
20.1 **The Principal-Agent Problem** **679**
Efficiency 680
Symmetric Information 680
Asymmetric Information 681
Solved Problem 20.1 682
APPLICATION *Honest Cabbie?* 683
20.2 **Using Contracts to Reduce Moral Hazard** **683**
Fixed-Fee Contracts 684
Contingent Contracts 685
APPLICATION *Health Insurance and Moral Hazard* 685
Solved Problem 20.2 686
Solved Problem 20.3 687
APPLICATION *Sing for Your Supper* 688
Solved Problem 20.4 690
Choosing the Best Contract 691
20.3 **Monitoring to Reduce Moral Hazard** **691**
Bonding 692
Solved Problem 20.5 693
APPLICATION *Capping Oil and Gas Bankruptcies* 694
Deferred Payments 695
Efficiency Wages 696
Monitoring Outcomes 697
20.4 **Checks on Principals** **697**
APPLICATION *Layoffs Versus Pay Cuts* 698
20.5 **Contract Choice** **700**
CHALLENGE SOLUTION *Clawing Back Bonuses* 701
Summary 702 ■ *Questions* 703

Chapter Appendixes 707

Appendix 2A: Regressions 707
Appendix 3A: Effects of a Specific Tax on Equilibrium 709
Appendix 4A: Utility and Indifference Curves 710
Appendix 4B: Maximizing Utility 712
Appendix 5A: The Slutsky Equation 714
Appendix 5B: Labor-Leisure Model 715
Appendix 6A: Properties of Marginal and Average Product Curves 716
Appendix 6B: The Slope of an Isoquant 716
Appendix 6C: Cobb-Douglas Production Function 716
Appendix 7A: Minimum of the Average Cost Curve 717
Appendix 7B: Japanese Beer Manufacturer's Short-Run Cost Curves 717
Appendix 7C: Minimizing Cost 718
Appendix 8A: The Elasticity of the Residual Demand Curve 720
Appendix 8B: Profit Maximization 721
Appendix 9A: Demand Elasticities and Surplus 721
Appendix 11A: Relationship Between a Linear Demand Curve and Its Marginal Revenue Curve 722
Appendix 11B: Incidence of a Specific Tax on a Monopoly 722
Appendix 12A: Perfect Price Discrimination 723
Appendix 12B: Group Price Discrimination 724
Appendix 12C: Block Pricing 724
Appendix 12D: Two-Part Pricing 725
Appendix 12E: Profit-Maximizing Advertising and Production 725
Appendix 13A: Nash-Cournot Equilibrium 726
Appendix 13B: Nash-Stackelberg Equilibrium 728
Appendix 13C: Nash-Bertrand Equilibrium 729
Appendix 15A: Factor Demands 730
Appendix 15B: Monopsony 731
Appendix 16A: The Present Value of Payments over Time 732
Appendix 18A: Welfare Effects of Pollution in a Competitive Market 732
Appendix 20A: Nonshirking Condition 734

Answers to Selected Questions and Problems 735
Sources for Challenges and Applications 753
References 767
Definitions 777
Index 782
Credits 812

Preface

When I was a student, I fell in love with microeconomics because it cleared up many mysteries about the world and provided the means to answer new questions. I wrote this book to illustrate that economic theory has practical, problem-solving uses and is not an empty academic exercise.

This book shows how individuals, policy makers, lawyers and judges, and firms can use microeconomic tools to analyze and resolve problems. For example, students learn that

- individuals can draw on microeconomic theories when deciding about issues such as whether to invest and whether to sign a contract that pegs prices to the government's measure of inflation;
- policy makers (and voters) can employ microeconomics to predict the impact of taxes, regulations, and other measures before they are enacted;
- lawyers and judges can use microeconomics in antitrust, discrimination, and contract cases; and
- firms can apply microeconomic principles to produce at minimum cost and maximize profit, select strategies, decide whether to buy from a market or to produce internally, and write contracts to provide optimal incentives for employees.

My experience in teaching microeconomics for the departments of economics at MIT; the University of Pennsylvania; and the University of California, Berkeley; the Department of Agricultural and Resource Economics at Berkeley; and the Wharton Business School has convinced me that students prefer this emphasis on real-world issues.

How This Book Differs from Others

This book differs from other microeconomics texts in three main ways:

- It places greater emphasis than other texts on *modern theories*—such as industrial organization theories, game theory, transaction cost theory, information theory, contract theory, and behavioral economics—that are useful in analyzing actual markets.
- It uses *real-world economic examples* to present the basic theory and offers many more Applications to a variety of real-world situations.
- It employs step-by-step *problem-based learning* to demonstrate how to use microeconomic theory to solve business problems and analyze policy issues.

Modern Theories

This book has all of the standard economic theory, of course. However, what sets it apart is its emphasis on modern theories that are particularly useful for understanding how firms behave and the effects of public policy.

Industrial Organization How do firms differentiate their products to increase their profits? When does market outcome depend on whether firms set prices or quantities? What effects do government price regulations have on firms' behavior? Industrial organization theories address these and many other questions.

Game Theory What's the optimal way to bid in an auction? How do firms set prices to prevent entry of rival firms? What strategy should parents use when their college-graduate child moves back in with them? Game theory provides a way of thinking about strategies and it provides methods to choose optimal strategies.

Contract Theory What kind of a contract should a firm offer a worker to induce the employee to work hard? How do people avoid being exploited by other people who have superior information? Modern contract theory shows how to write contracts to avoid or minimize such problems.

Behavioral Economics Should a firm allow workers to opt in or opt out of a retirement system? How should people respond to ultimatums? We address questions such as these using behavioral economics—one of the hottest new areas of economic theory—which uses psychological research and theory to explain why people deviate from rational behavior.

Real-World Economics

This book demonstrates that economics is practical and provides a useful way to understand actual markets and firms' and consumers' decisions in two ways. First, it presents the basic theory using models estimated with real-world data. Second, it uses the theory to analyze hundreds of real-world applications.

Using Estimated Models to Illustrate Theory The text presents the basic theory using estimated demand curves, supply curves, production functions, and cost functions in most chapters. For example, students learn how imported oil limits the price that U.S. oil producers can charge based on estimated supply and demand curves, derive a Japanese beer manufacturer's cost curve using an estimated production function, examine the regulation of natural gas monopolies employing estimated demand and cost curves, and analyze oligopoly firms' strategies using estimated demand curves and cost and profit data from the real-world rivalries between United Airlines and American Airlines and between Coke and Pepsi.

Applications Applications use economic theory to predict the price effect of allowing drilling in the Arctic National Wildlife Refuge based on estimated demand and supply curves, demonstrate how iTunes price increases affect music downloads using survey data, explain why some top-end designers limit the number of designer bags customers can buy, analyze why Amazon raised the price for its Prime service, and measure the value of using the Internet.

Problem-Based Learning

People, firms, and policy makers have to solve economic problems daily. This book uses a problem-solving approach to demonstrate how economic theory can help them make good decisions.

Solved Problems After the introductory chapter, each chapter provides an average of over five Solved Problems. Each Solved Problem poses a qualitative or quantitative question and then uses a step-by-step approach to model good problem-solving techniques. These issues include whether Peter Guber and Joe Lacob should have bought the Golden State Warriors, how to determine Intel's and AMD's profit-maximizing quantities and prices using their estimated demand curves and marginal costs, and how regulating a monopoly's price affects consumers and firms.

Challenges Starting with Chapter 2, each chapter begins with a Challenge that presents information about an important, current real-world issue and concludes with a series of questions about that material. At the end of the chapter, a Challenge Solution answers these questions using methods presented in that chapter. That is, a Challenge combines an Application and a Solved Problem to motivate the material in the chapter. The issues covered include the price and quantity effects from introducing genetically modified foods, why Americans buy more ebooks than do Germans, whether higher salaries for star athletes raise ticket prices, whether it pays to go to college, and how Heinz can use sales to increase its profit on ketchup.

End-of-Chapter Questions Starting with Chapter 2, each chapter ends with an extensive set of Questions, many of which draw on topical, real-world issues. Each Solved Problem and Challenge has at least one associated end-of-chapter question that references them and asks students to extend or reapply their analyses. Many of the Questions relate to the Applications. Answers to selected end-of-chapter Questions appear at the end of the book, and select end-of-chapter Questions are available in MyLab Economics for self-assessment, homework, or testing.

What's New in the Eighth Edition

The Eighth Edition of *Microeconomics* is substantially updated and modified based on the extremely helpful suggestions of faculty and students who used the first seven editions. The major changes in this edition are:

- All the Challenges and almost all the examples and Applications throughout the book are updated or new.
- The book has several new and many revised Solved Problems.
- Many of the end-of-chapter Questions are new, updated, or revised.
- All chapters are revised.
- Each substantive chapter has a new feature, in which we analyze a *Common Confusion*.

(And, possibly most importantly, the book has two new cartoons.)

Challenges, Solved Problems, and Questions

All of the Challenges are new or updated. Because users requested more Solved Problems, I increased the number of Challenges and Solved Problems in this edition to 111, many of which are new or substantially revised. Every Solved Problem has at least one associated Question at the end of the chapter.

About 40% of these Solved Problems refer to real-world events. Many of these are associated with an adjacent Application or example in the text. In addition to the Challenges, examples of a paired Application and Solved Problem include the effect of oil drilling in the Arctic National Wildlife Refuge on prices, the opportunity cost of getting an MBA, the social cost of a natural gas price ceiling, Apple's iPad pricing, and the price effects of reselling textbooks bought abroad in the United States.

Starting with Chapter 2, the end of each chapter has an average of over 42 verbal, graphical, and mathematical Questions. This edition has 810 Questions, 47 more than in the previous edition. Over 12% of the Questions are new or updated. Many of these Questions refer to recent real-life events and issues drawn from newspapers, journal articles, and other sources.

Applications

The Eighth Edition has 133 Challenges and Applications, 4 more than in the previous edition. Of these, 35% are new and 53% are updated, so that 87% are new or updated. The vast majority of the Applications cover events in 2015 and 2016, a few deal with historical events, and the remaining ones examine timeless material.

To make room for the new Applications, older Applications from the Seventh Edition were moved to MyLab Economics. Also, several new ones have been added to the hundreds of Applications and other materials in MyLab Economics.

New and Revised Material in Chapters

I have revised every chapter—including most sections. Every chapter has new and updated Applications and Challenges. Virtually every chapter has updated examples and statistics. Some of the larger changes include:

- Chapters 2 and 3 are substantially rewritten. They illustrate the basic supply-and-demand theory using empirical estimates from the avocados, coffee, corn, ethanol, and oil markets. The major coffee example is new to this edition. Three Solved Problems are significantly revised.
- Chapters 4 and 5 are reorganized and significantly revised, particularly the section *Cost-of-Living-Adjustments* in Chapter 5.
- Chapter 6 has a substantially modified section, *Production*, and light revisions elsewhere.
- Chapter 7 is moderately revised, particularly the material associated with Figure 7.2.
- Chapter 8 is substantially revised, particularly the beginning of the *Two Steps to Maximizing Profit* section, the discussion of the shutdown decision, the *Competition in the Short Run* section, the section on *Entry and Exit*, and the section *Long-Run Market Supply When Input Prices Vary with Output*, as well as several figures.
- Chapter 9 is substantially rewritten, particularly the introduction, the section on *Policies That Shift Supply and Demand Curves*, the discussion of trading oil, which uses a new estimated model. In this chapter and following chapters, deadweight loss is expressed as a negative number consistently.
- Chapter 10 a Solved Problem, the comparison of Pareto superiority, and the material of *Efficiency and Equity* are rewritten.
- Chapter 11 has two new Solved Problems and all the material on Apple is revised.
- Chapter 12 is moderately revised throughout, the group discrimination material is reorganized and significantly revised, and the two-part pricing material is lightly revised. It contains a new Solved Problem.
- Chapter 13 is significantly revised. The section *Cartels* is reorganized and rewritten. The sections *Cournot Oligopoly* and *Comparison of Competitive, Stackelberg, Cournot, and Collusive Equilibria* are rewritten.
- Chapter 14 has two new Solved Problems. The static game section is completely reorganized and rewritten. The dynamic game section is significantly revised. The chapter has a new discussion of double auctions.

- Chapter 15 has a modified Table 15.1, Figures 15.9 and 15.10, and a discussion of monopsony.
- Chapter 16 has two rewritten Solved Problems and many facts are updated.
- Chapter 17 has a rewritten discussion of framing, a significantly modified Solved Problem, a new introduction to the section on *Reducing Risk*, and a new subsection *Just Say No*.
- Chapter 18 has substantially revised sections on the Coase Theorem, club goods, and public goods.
- Chapter 19 is moderately rewritten and has a substantially revised Solved Problem.
- Chapter 20 has a new Challenge Solution.

A New Feature: Common Confusions

A new feature in this edition are discussions of *Common Confusions*—widely held but false beliefs. After Chapter 1, every chapter has at least one Common Confusion. We use economic theory to explain why these beliefs are incorrect.

Alternative Organizations

Because instructors cover material in different orders, I designed this textbook for maximum flexibility. The most common approach to teaching microeconomics is to follow the sequence of the chapters in the first half of this book: supply and demand (Chapters 2 and 3), consumer theory (Chapters 4 and 5), the theory of the firm (Chapters 6 and 7), and the competitive model (Chapters 8 and 9). Many instructors then cover monopoly (Chapter 11), price discrimination (Chapter 12), oligopoly (Chapters 13 and 14), input markets (Chapter 15), uncertainty (Chapter 17), and externalities (Chapter 18).

A common variant is to present uncertainty (Sections 17.1 through 17.3) immediately after consumer theory. Many instructors like to take up welfare issues between discussions of the competitive model and noncompetitive models, as Chapter 10, on general equilibrium and economic welfare, does. Alternatively, some instructors cover this chapter at the end of the course. Faculty can assign material on factor markets earlier: Section 15.1 could follow the chapters on competition, and the remaining sections could follow Chapter 11. The material in Chapters 14–20 can be presented in a variety of orders, though Chapter 20 should follow Chapter 19 if both are covered, and Section 17.4 should follow Chapter 16.

Many business school courses skip consumer theory (and possibly some aspects of supply and demand, such as Chapter 3) to allow more time for consideration of the topics covered in the second half of this book. Business school faculty may want to place particular emphasis on game and theory strategies (Chapter 14), capital markets (Chapter 16), and modern contract theory (Chapters 19 and 20).

Optional, technically demanding sections are marked with a star (★). Subsequent sections and chapters can be understood even if these sections are skipped.

MyLab Economics

MyLab Economics's powerful assessment and tutorial system works hand-in-hand with the Eighth Edition of *Microeconomics*. It includes comprehensive homework, quiz, test, and tutorial options, allowing students to test their knowledge and instructors to manage all assessment needs in one program. Students and instructors can register, create, and access all of their MyLab courses, regardless of discipline, from one convenient online location: **http://www.pearsonmylab.com**.

Key features in the MyLab Economics course for *Microeconomics* Eighth Edition include the following resources for students and instructors:

- **Pearson eText.** The eText helps students to learn on their own by helping them better understand course material. The worked examples, animations, and interactive tutorials bring learning to life. Students can apply these concepts using self-assessment, algorithmic exercises in MyLab, providing students with a complete digital learning experience.
- **MyLab Economics Videos.** Key figures and concepts from the textbook are presented in step-by-step animations with audio explanations of the action.
- **MyLab Economics Solved Problems.** Many students have difficulty applying economics concepts to solving problems. The goal of this digital resource is to help students overcome this hurdle by learning how to solve an economic problem by breaking it down into steps. Each Solved Problem in MyLab Economics also includes at least one additional graded practice exercise for students. These interactive tutorials help students apply basic problem-solving skills to homework, quizzes, and exams. The goal is for students to build skills they can use to analyze real-world economic issues they hear and read about in the news.
- **Additional Readings (Applications, Supplemental Material, and Solved Problems).** Additional Applications, Supplemental Material, and Solved Problems are available in MyLab Economics.
- **NEW: Math Review Exercises in MyLab Economics.** MyLab Economics now offers a rich array of assignable and auto-graded exercises covering fundamental math concepts geared specifically to principles and intermediate economics students.
- **Practice.** Algorithmically generated homework and study plan exercises with instant feedback ensure varied and productive practice that help students improve their understanding and prepare for quizzes and tests. Many exercises require the student to draw figures or solve mathematic problems.
- **Learning Resources.** Personalized learning aids such as Help Me Solve This problem walkthroughs, Teach Me explanations of the underlying concept, and figure Animations provide on-demand help when students need it most.
- **Personalized Study Plan.** The Study Plan acts as a tutor, providing personalized recommendations for each student based on his or her demonstrated ability to master the learning objectives in your course. Consequently, students focus on precisely the areas they need to review using customized practice and learning aids—such as videos and eText, tutorials. You can use the report available in the Gradebook to lecture on the material for which students need the most support.
- **Learning Catalytics.** Learning Catalytics™ helps you generate class discussion, customize your lecture, and promote peer-to-peer learning with real-time analytics. Learning Catalytics allows students to use their smart phones, tablets, or laptops to respond to your questions or to engage in interactive tasks during class. You can use this information to adjust your teaching strategy in real time.
- **Current News Exercises.** Every week, Pearson scours the news, finds a current article appropriate for a microeconomics course and adds a gradable exercise based on it to MyLab Economics.
- **Reporting Dashboard.** Faculty can view, analyze, and report learning outcomes clearly and easily using the Reporting Dashboard. It is available via the Gradebook and fully mobile-ready. The Reporting Dashboard presents student performance data at the class, section, and program levels in an accessible, visual manner.

- **LMS Integration.** You can now link to MyLab Economics from Blackboard Learn, Brightspace by D2L, Canvas, or Moodle. Thus, you can directly access MyLab Economics assignments, rosters, and resources. You can also synchronize grades with your LMS gradebook. Students need sign-on only once to obtain access to all their personalized learning resources.
- **Mobile Ready.** Students and instructors can access multimedia resources and complete assessments from any mobile device.
- **Experiments in MyLab Economics.** Flexible, easy to assign, auto-graded, and available in Single and Multiplayer versions, the Experiments in MyLab Economics make learning fun and engaging.

For more information, visit **http://www.myeconlab.com**.

Supplements

A full range of supplementary materials to support teaching and learning accompanies this book.

- The *Online Instructor's Manual* revised by Leonie Stone has many useful and creative teaching ideas. It also offers a chapter outline, additional discussion questions, additional questions and problems, and solutions for all additional questions and problems.
- The *Online Solutions Manual* provides solutions for all the end-of-chapter questions in the text.
- The *Online Test Bank* by Lourenço Paz of Syracuse University features problems of varying levels of complexity, suitable for homework assignments and exams. Many of these multiple-choice questions draw on current events.
- The *Computerized Test Bank* reproduces the Test Bank material in the TestGen software, which is available for Windows and Macintosh. With TestGen, instructors can easily edit existing questions, add questions, generate tests, and print the tests in a variety of formats.
- The *Online PowerPoint Presentation* by Ting Levy of Florida Atlantic University contains text figures and tables, as well as lecture notes. These slides allow instructors to walk through examples from the text during in-class presentations.

These teaching resources are available online for download at the Instructor Resource Center, **www.pearsonglobaleditions.com/Perloff**, and on the catalog page for *Microeconomics*. **This title is available as an eBook and can be purchased at most eBook retailers.**

Acknowledgments

My greatest debt is to my students. My students at MIT; the University of Pennsylvania; and the University of California, Berkeley, patiently dealt with my various approaches to teaching them microeconomics and made useful (and generally polite) suggestions.

Over the years, many excellent research assistants—Hayley Chouinard, R. Scott Hacker, Guojun He, Nancy McCarthy, Enrico Moretti, Yann Panassie, Lisa Perloff, Asa Sajise, Hugo Salgado, Gautam Sethi, Edward Shen, Klaas van 't Veld, and Ximing Wu—worked hard to collect facts, develop examples and figures, and check material.

Many people were very generous in providing me with data, models, and examples, including among others: Thomas Bauer (University of Bochum), Peter Berck (University of California, Berkeley), James Brander (University of British Columbia), Leemore Dafny (Northwestern University), Lucas Davis (University of California, Berkeley), James Dearden (Lehigh University), Farid Gasmi (Université des Sciences Sociales), Avi Goldfarb (University of Toronto), Claudia Goldin (Harvard University), Rachel Goodhue (University of California, Davis), William Greene (New York University), Nile Hatch (University of Illinois), Larry Karp (University of California, Berkeley), Ryan Kellogg (University of Michigan), Arthur Kennickell (Federal Reserve, Washington), Fahad Khalil (University of Washington), Lutz Kilian (University of Michigan), Christopher Knittel (University of California, Davis), Jean-Jacques Laffont (deceased), Ulrike Malmendier (University of California, Berkeley), Karl D. Meilke (University of Guelph), Eric Muehlegger (Harvard University), Giancarlo Moschini (Iowa State University), Michael Roberts (North Carolina State University), Wolfram Schlenker (Columbia University), Junichi Suzuki (University of Toronto), Catherine Tucker (MIT), Harald Uhlig (University of Chicago), Quang Vuong (Université des Sciences Sociales, Toulouse, and University of Southern California), and Joel Waldfogel (University of Minnesota).

Writing a textbook is hard work for everyone involved. I am grateful to the many teachers of microeconomics who spent untold hours reading and commenting on proposals and chapters. Many of the best ideas in this book are due to them.

I am particularly grateful to Jim Brander of the University of British Columbia who provided material for Chapters 13 and 14, has given me many deep and insightful comments on many editions of this book, and with whom I wrote another, related book. Much of the new material in this edition was jointly written with him. My other biggest debt is to James Dearden, Lehigh University, who made extremely insightful comments on all previous editions and wrote some of the end-of-chapter questions.

In earlier editions, Peter Berck made major contributions to Chapter 16. Charles F. Mason made particularly helpful comments on many chapters. Larry Karp helped me to develop two of the sections and carefully reviewed the content of several others. Robert Whaples, Wake Forest University, read many chapters in earlier editions and offered particularly useful comments. He also wrote the first draft of one of my favorite Applications.

For this edition, my biggest debts are to Tony Lima and Gordon Lenjosek. Tony prepared the many, many MyLab Economics Videos and made substantive comments. In addition to carefully checking for typographic and substantive errors in the previous edition and suggesting better ways to present many topics, Gordon also contributed to creating the Global Edition for this book. I am grateful to the following people who reviewed the book or sent me valuable suggestions at various stages:

M. Shahid Alam, Northeastern University
Anne Alexander, University of Wyoming
Samson Alva, Boston College
Richard K. Anderson, Texas A&M University
Niels Anthonisen, University of Western Ontario
Wilma Anton, University of Central Florida
Emrah Arbak, State University of New York at Albany
Nestor M. Arguea, University of West Florida
Scott E. Atkinson, University of Georgia
Talia Bar, Cornell University
Raymond G. Batina, Washington State University
Anthony Becker, St. Olaf College
Robert A. Berman, American University
Gary Biglaiser, University of North Carolina, Chapel Hill
S. Brock Blomberg, Wellesley College
Hein Bogaard, George Washington University
Vic Brajer, California State University, Fullerton
Jurgen Brauer, Augusta State University
Bruce Brown, Cal Polytech Pomona and UCLA
Byoung Heon Jun, Korea University
Cory S. Capps, University of Illinois, Urbana-Champaign
John Cawley, Cornell University

Cedric Ceulmans, SUNY Purchase
Indranil Chakraborty, University of Oklahoma
Leo Chan, University of Kansas
Joni S. Charles, Southwest Texas State University
Kwang Soo Cheong, University of Hawaii at Manoa
Darlene Chisholm, Suffolk University
Joy L. Clark, Auburn University, Montgomery
Dean Croushore, Federal Reserve Bank of Philadelphia
Douglas Dalenberg, University of Montana
Shooshan Danagoulian, Wayne State University
Andrew Daughety, Vanderbilt University
Carl Davidson, Michigan State University
Ronald Deiter, Iowa State University
Manfred Dix, Tulane University
John Edgren, Eastern Michigan University
Patrick Emerson, University of Colorado, Denver
Xin Fang, Hawai'i Pacific University
Bernard Fortin, Université Laval
Tom Friedland, Rutgers University
Roy Gardner, Indiana University
Rod Garratt, University of California, Santa Barbara
Wei Ge, Bucknell University
Lisa Giddings, University of Wisconsin, La Crosse
J. Fred Giertz, University of Illinois, Urbana-Champaign
Haynes Goddard, University of Cincinnati
Steven Goldman, University of California, Berkeley
Julie Gonzalez, University of California, Santa Cruz
Rachel Goodhue, University of California, Davis
Srihari Govindan, University of Western Ontario
Gareth Green, Seattle University
Thomas A. Gresik, Pennsylvania State University
Jonathan Gruber, MIT
Steffan Habermalz, University of Nebraska, Kearney
Claire Hammond, Wake Forest University
John A. Hansen, State University of New York, Fredonia
Bradley Hardy, American University
Philip S. Heap, James Madison University
L. Dean Hiebert, Illinois State University
Kathryn Ierulli, University of Illinois, Chicago
Mike Ingham, University of Salford, U.K.
Samia Islam, Boise State University
D. Gale Johnson, University of Chicago
Erik Jonasson, Lund University, Sweden
Charles Kahn, University of Illinois, Urbana-Champaign
Courtney Kamauf, Pearson
Vibha Kapuria-Foreman, Colorado College
Paula M. Kazi, Bucknell University
Carrie Kerekes, Florida Gulf Coast University
Alan Kessler, Providence College
Kristin Kiesel, University of California, Davis
Kate Krause, University of New Mexico
Robert Lemke, Lake Forest College
Jing Li, University of Pennsylvania
Qihong Liu, University of Oklahoma
Zhou Lu, City College of New York
Fred Luk, University of California, Los Angeles
Robert Main, Butler University
David Malueg, Tulane University
Steve Margolis, North Carolina State University
Kate Matraves, Michigan State University
James Meehan, Colby College
Claudio Mezzetti, University of North Carolina, Chapel Hill
Chun-Hui Miao, University of South Carolina
Janet Mitchell, Cornell University
Felix Munoz-Garcia, Washington State University
Babu Nahata, University of Louisville
Kathryn Nantz, Fairfield University
Zafar Nazarov, Indiana University – Purdue Fort Wayne
Jawwad Noor, Boston University
Yuka Ohno, Rice University
Patrick B. O'Neil, University of North Dakota
John Palmer, University of Western Ontario
Christos Papahristodoulou, Uppsala University
Silve Parviainen, University of Illinois, Urbana-Champaign
Lourenço Paz, Syracuse University
Sharon Pearson, University of Alberta
Ingrid Peters-Fransen, Wilfrid Laurier University
Jaishankar Raman, Valparaiso University
Sunder Ramaswamy, Middlebury College
Lee Redding, University of Michigan, Dearborn
David Reitman, Department of Justice
Luca Rigotti, Tillburg University
S. Abu Turab Rizvi, University of Vermont
Bee Yan Aw Roberts, Pennsylvania State University
Richard Rogers, Ashland University
Nancy Rose, Sloan School of Business, MIT
Joshua Rosenbloom, University of Kansas
Roy Ruffin, University of Houston
Matthew Rutledge, Boston College
Alfonso Sanchez-Penalver, University of Massachusetts, Boston
George Santopietro, Radford College
David Sappington, University of Florida
Rich Sexton, University of California, Davis
Quazi Shahriar, San Diego State University
Jacques Siegers, Utrecht University, The Netherlands
Alasdair Smith, University of Sussex
William Doyle Smith, University of Texas at El Paso
Philip Sorenson, Florida State University
Peter Soule, Park College
Robert Stearns, University of Maryland
Jennifer Lynn Steele, Washington State University
Shankar Subramanian, Cornell University
Albert J. Sumell, Youngstown State University
Beck A. Taylor, Baylor University
Scott Templeton, Clemson University
Mark L. Tendall, Stanford University
Justin Tevie, University of New Mexico
Wade Thomas, State University of New York, Oneonta

Judith Thornton, University of Washington
Vitor Trindade, Syracuse University
Russell Triplett, University of North Florida
Nora Underwood, University of California, Davis
Burcin Unel, University of Florida
Kay Unger, University of Montana
Alan van der Hilst, University of Washington
Bas van der Klaauw, Free University Amsterdam and Tinbergen Institute
Andrew Vassallo, Rutgers University
Jacob L. Vigdor, Duke University
Peter von Allmen, Moravian College
Eleanor T. von Ende, Texas Tech University
Curt Wells, Lund University
Lawrence J. White, New York University
John Whitehead, East Carolina University
Ryan Williams, Texas Tech University
Colin Wright, Claremont McKenna College
Bruce Wydick, University of San Francisco
Peter Zaleski, Villanova University
Artie Zillante, Florida State University
Mark Zupan, University of Arizona

In addition, I thank Bob Solow, the world's finest economics teacher, who showed me how to simplify models without losing their essence. I've also learned a great deal over the years about economics and writing from my coauthors on other projects, especially Dennis Carlton (my coauthor on *Modern Industrial Organization*), Jackie Persons, Steve Salop, Michael Wachter, Larry Karp, Peter Berck, Amos Golan, and Dan Rubinfeld (whom I thank for still talking to me despite my decision to write this book).

The first edition of this book benefited from the work of the two best development editors in the business, Jane Tufts and Sylvia Mallory. Jane Tufts reviewed drafts of the first edition of this book for content and presentation. By showing me how to present the material as clearly, orderly, and thoroughly as possible, she greatly strengthened this text. Sylvia Mallory worked valiantly to improve my writing style and helped to shape and improve every aspect of the book's contents and appearance in each of the first four editions.

It was a pleasure to work with the good people at Pearson, who were incredibly helpful in producing this book. Adrienne d'Ambrosio, Director of Quantitative Business and Economics, has worked on planning and revising many editions. I am extremely grateful to Christina Masturzo, Senior Portfolio Manager; Kathryn Brightney, Content Producer; Meredith Gertz; and Gillian Hall, who helped me plan this revision and made very valuable suggestions at each stage of the process. Christina managed all aspects of planning, writing, and producing this textbook. Kathryn assisted with all aspects of revision planning and project oversight. Meredith did her usual outstanding job of supervising the production process, assembling the extended publishing team, and managing the design of the handsome interior. She and Gillian, and the rest of the team at The Aardvark Group Publishing Services have my sincere gratitude for designing the book and keeping the project on track and on schedule. As always, I'm particularly thankful to work with Gillian, who is wonderfully flexible and committed to producing the best book possible.

I thank Laurie Entringer for managing the cover design. I also want to acknowledge, with appreciation, the efforts of Melissa Honig, Noel Lotz, and Courtney Kamauf in developing MyLab Economics, the online assessment and tutorial system for the book.

Finally, and most importantly, I thank my wife, Jackie Persons; and daughter, Lisa Perloff, for their great patience and support during the nearly endless writing process. And I apologize for misusing their names—and those of my other relatives and friends—in the book!

J. M. P.

This page intentionally left blank

1 Introduction

An Economist's Theory of Reincarnation: If you're good, you come back on a higher level. Cats come back as dogs, dogs come back as horses, and people—if they've been really good like George Washington—come back as money.

microeconomics
the study of how individuals and firms make themselves as well off as possible in a world of scarcity and the effects of their actions on markets and the entire economy

If each of us could get all the food, clothing, and toys we wanted without working, no one would study economics. Unfortunately, most of the good things in life are scarce—we can't all have as much as we want. Thus, scarcity is the mother of economics. **Microeconomics** is the study of how individuals and firms make themselves as well off as possible in a world of scarcity and the effects of their actions on markets and the entire economy. Microeconomics is often called *price theory* to emphasize the important role that prices play. Microeconomics explains how the actions of all buyers and sellers determine prices and how prices influence the decisions and actions of individual buyers and sellers.

In this chapter, we examine three main topics

1. **Microeconomics: The Allocation of Scarce Resources.** Microeconomics is the study of the allocation of scarce resources.
2. **Models.** Economists use models to make testable predictions.
3. **Uses of Microeconomic Models.** Individuals, governments, and firms use microeconomic models and predictions in decision making.

1.1 Microeconomics: The Allocation of Scarce Resources

Individuals and firms allocate their limited resources to make themselves as well off as possible. Consumers pick the mix of goods and services that makes them as happy as possible given their limited wealth. Firms decide which goods and services to produce, where to produce them, how much to produce to maximize their profits, and how to produce those a given level of output at the lowest cost by using more or less of various inputs such as labor, capital, materials, and energy. The owners of a depletable natural resource such as oil decide when to use it. Government decision makers—to benefit consumers, firms, or government bureaucrats—decide which goods and services the government produces and whether to subsidize, tax, or regulate industries and consumers.

Trade-Offs

People make trade-offs because they can't have everything. A society faces three key trade-offs:

- **Which goods and services to produce:** If a society produces more cars, it must produce fewer of other goods and services, because society has a limited amount of *resources*—workers, raw materials, capital, and energy—available to produce goods.
- **How to produce:** To produce a given level of output, a firm must use more of one input if it uses less of another input. For example, cracker and cookie manufacturers switch between palm oil and coconut oil, depending on which is less expensive.
- **Who gets the goods and services:** The more of society's goods and services you get, the less someone else gets.

Who Makes the Decisions

The government may explicitly make these three allocation decisions for society. Alternatively, the interaction of independent choices by many individual consumers and firms may determine society's allocation decisions.

In the former Soviet Union, the government told manufacturers how many cars of each type to make and which inputs to use to make them. The government also decided which consumers would get a car.

In most other countries, how many cars of each type are produced and who gets them are determined by how much it costs to make cars of a particular quality in the least expensive way and how much consumers are willing to pay for them. More consumers would own a handmade Rolls-Royce Phantom and fewer would buy a mass-produced Toyota Camry if a Rolls-Royce was not 18 times more expensive than a Camry.

Prices Determine Allocations

Prices link the decisions about *which goods and services to produce*, *how to produce them*, and *who gets them*. Prices influence the decisions of individual consumers and firms, and the interactions of those decisions by consumers, firms, and the government determine price.

market
an exchange mechanism that allows buyers to trade with sellers

Interactions between consumers and firms take place in a **market**, which is an exchange mechanism that allows buyers to trade with sellers. A market may be a town square where people go to trade food and clothing, or it may be an international telecommunications network over which people buy and sell financial securities. Typically, when we talk about a single market, we refer to trade in a single good or group of goods that consumers view as very similar, such as Fuji apples, soft drinks, movies, novels, or automobiles.

Most of this book concerns how prices are determined within a market. We show that the number of buyers and sellers in a market and the amount of information they have help determine whether the price equals the cost of production. We also show that if no market is available—and hence no market price—serious problems, such as high levels of pollution, result.

Application

Twinkie Tax

Many American, Australian, British, Canadian, New Zealand, and Taiwanese jurisdictions are proposing a *Twinkie tax* on unhealthful fatty and sweet foods or a tax on sugary soft drinks to reduce obesity and cholesterol problems, particularly among children. One survey found that 45% of adults would support a 1¢ tax per pound on soft drinks, chips, and butter, with the revenues used to fund health education programs.

In recent years, many communities around the world debated (and some passed) new taxes on sugar-sweetened soft drinks. In 2014, Rosa DeLauro, a Connecticut Congressional representative, proposed a national soda tax. New beverage taxes went into effect in Mexico in 2014 and in Berkeley, California, in 2015. In 2016, the United Kingdom announced it would impose a soft drink tax in 2018. At least 23 states differentially tax soft drinks, candy, chewing gum, and snack foods such as potato chips. Today, many school districts throughout the United States ban soft drink vending machines. This ban discourages consumption, as would an extremely high tax. Britain's largest life insurance firm charges obese people higher premiums for life insurance policies.

New taxes will affect *which foods are produced*, as firms offer new low-fat and low-sugar products, and *how fast foods are produced*, as manufacturers reformulate their products to lower their tax burden. These taxes will also influence *who gets these goods* as consumers, especially children, replace them with less expensive, untaxed products.[1]

1.2 Models

Everything should be made as simple as possible, but not simpler. —Albert Einstein

model
a description of the relationship between two or more economic variables

To *explain* how individuals and firms allocate resources and how market prices are determined, economists use a **model**: a description of the relationship between two or more economic variables. Economists also use models to *predict* how a change in one variable will affect another.

Application

Income Threshold Model and China

According to an *income threshold model*, no one who has an income level below a particular threshold buys a particular consumer durable, such as a refrigerator or car. The theory also holds that almost everyone whose income is above that threshold buys the product.

If this theory is correct, we predict that, as most people's incomes rise above the threshold in emergent economies, consumer durable purchases will increase from near zero to large numbers virtually overnight. This prediction is consistent with evidence from Malaysia, where the income threshold for buying a car is about $4,000.

In China, incomes have risen rapidly and now exceed the threshold levels for many durable goods. As a result, experts correctly predicted that the greatest consumer durable goods sales boom in history would take place there. Anticipating this boom, companies from around the globe greatly increased their investments in durable goods manufacturing plants in China. Annual foreign direct investments went from $916 million a year in 1983 to $127 billion in 2015. In anticipation of this growth potential, even traditional political opponents of the People's Republic—Taiwan, South Korea, and Russia—are investing in China.

One of the most desirable durable goods is a car. Li Rifu, a 46-year-old Chinese farmer and watch repairman, thought that buying a car would improve the odds that his 22- and 24-year-old sons would find girlfriends, marry, and produce

[1]The sources for Applications are available at the back of this book.

grandchildren. Soon after Mr. Li purchased his Geely King Kong car for the equivalent of $9,000, both sons met girlfriends, and his older son got married.

First-time customers buy four-fifths of all new cars sold in China. An influx of first-time buyers was responsible for Chinese car sales increasing by a factor of 15 between 2000 and 2015. By 2010, China became the second largest manufacturer of motor vehicles. By 2014, China was producing more cars than the United States and Japan combined, as well as more than the entire European Union.

Simplifications by Assumption

We stated the income threshold model in words, but we could have presented it using graphs or mathematics. Regardless of how we describe the model, an economic model is a simplification of reality that contains only its most important features. Without simplifications, it is difficult to make predictions, because the real world is too complex to analyze fully.

By analogy, if the manual accompanying your new smart TV has a diagram showing the relationships between all the parts in the TV, the diagram will be overwhelming and useless. In contrast, if it shows a photo of the lights on the front of the machine with labels describing the significance of each light, the manual is useful and informative.

Economists make many *assumptions* to simplify their models.[2] When using the income threshold model to explain car buying in China, we *assume* that factors other than income, such as the color of cars, are irrelevant to the decision to buy cars. Therefore, we ignore the color of Chinese cars in describing the relationship between average income and the number of cars consumers want. If this assumption is correct, by ignoring color, we make our analysis of the auto market simpler without losing important details. If we're wrong and these ignored issues are important, our predictions may be inaccurate.

Throughout this book, we start with strong assumptions to simplify our models. Later, we add complexities. For example, in most of the book, we assume that consumers know the price each firm charges. In many markets, such as the New York Stock Exchange, this assumption is realistic. It is not realistic in other markets, such as the market for used automobiles, in which consumers do not know the prices each firm charges. To devise an accurate model for markets in which consumers have limited information, we add consumer uncertainty about price into the model in Chapter 19.

Testing Theories

Blore's Razor: *When given a choice between two theories, take the one that is funnier*.

Economic *theory* is the development and use of a model to test *hypotheses*, which are predictions about cause and effect. We are interested in models that make clear, testable predictions, such as "If the price rises, the quantity demanded falls." A theory that says "People's behavior depends on their tastes, and their tastes change randomly at random intervals," is not very useful because it does not lead to testable predictions.

[2]An economist, an engineer, and a physicist are stranded on a desert island with a can of beans but no can opener. How should they open the can? The engineer proposes hitting the can with a rock. The physicist suggests building a fire under it to build up pressure and burst the can open. The economist thinks for a while and then says, "*Assume* that we have a can opener . . . ."

An alternative theory.

Economists test theories by checking whether predictions are correct. If a prediction does not come true, they may reject the theory.[3] Economists use a model until evidence refutes it or someone creates a better model.

A good model makes sharp, clear predictions that are consistent with reality. Some very simple models make sharp predictions that are incorrect, and other more complex models make ambiguous predictions—any outcome is possible—which are untestable. The skill in model building is to chart a middle ground.

The purpose of this book is to teach you how to think like an economist in the sense that you can build testable theories using economic models or apply existing models to new situations. Although economists think alike in that they develop and use testable models, they often disagree. One may present a logically consistent argument that prices will go up next quarter. Another, using a different but equally logical theory, may contend that prices will fall. If the economists are reasonable, they agree that pure logic alone cannot resolve their dispute. Indeed, they agree that they'll have to use empirical evidence—facts about the real world—to find out which prediction is correct.

Maximizing Subject to Constraints

In most economic models, a decision maker maximizes an objective subject to a constraint. Consumers maximize their well-being subject to a budget constraint, which says that their resources limit how many goods they can buy. Firms maximize profits subject to technological and other constraints. Governments may try to maximize the welfare of consumers or firms subject to constraints imposed by limited resources and the behavior of consumers and firms. The limit on resources plays a crucial role in these models. Were it not for scarcity, people could consume unlimited amounts of goods and services, and sellers could become rich beyond limit.

The maximizing behavior of individuals and firms determines society's three main allocation decisions: which goods are produced, how they are produced, and who gets them. For example, firms sell diamond-studded pocket combs only if it is profitable to sell them. The firms make and sell these combs only if consumers value the combs at least as much as it costs the firm to produce them. Consumers buy the combs only if they get more pleasure from the combs than they would from other goods they could buy with the same resources.

Positive Versus Normative

Using models of this maximizing behavior sometimes leads to predictions that seem harsh or heartless. For instance, a World Bank economist predicted that if an African

[3] We can use evidence on whether a theory's predictions are correct to *refute* the theory but not to *prove* it. If a model's prediction is inconsistent with what actually happened, the model must be wrong, so we reject it. However, even if the model's prediction is consistent with reality, the model's prediction may be correct for the wrong reason. Hence, we cannot prove that a model is correct—we can only *fail to reject it*.

government used price controls to keep the price of food low during a drought, food shortages would occur and people would starve. The predicted outcome is awful, but the economist was not heartless. The economist was only making a scientific prediction about the relationship between cause and effect: Price controls (cause) lead to food shortages and starvation (effect).

positive statement
a testable hypothesis about cause and effect

Such a scientific prediction is a **positive statement**: a testable hypothesis about cause and effect. "Positive" does not mean that we are *certain* about the truth of our statement—it only indicates that we can *test* the truth of the statement.

If the World Bank economist is correct, should the government control prices? If the government believes the economist's predictions, it knows that the low prices help those consumers who are lucky enough to be able to buy as much food as they want while hurting both the firms that sell food and the people who are unable to buy as much food as they want, some of whom may die. As a result, the government's decision to use price controls or not depends on whether the government cares more about the winners or the losers. In other words, to decide on its policy, the government makes a value judgment.

Instead of making a prediction and then testing it before making a value judgment of whether to use price controls, the government could make a value judgment directly. The value judgment could be based on the belief that "because people *should* have prepared for the drought, the government *should not* try to help them by keeping food prices low." Alternatively, the judgment could be based on the view that "people *should* be protected against price gouging during a drought, so the government *should* use price controls."

normative statement
a conclusion as to whether something is good or bad

These two statements are *not* scientific predictions. Each is a value judgment or **normative statement**: a conclusion as to whether something is good or bad. A normative statement is untestable because a value judgment cannot be refuted by evidence. It is a prescription rather than a prediction. A normative statement concerns what somebody believes *should* happen; a positive statement concerns what *will* happen.

Although people may draw a normative conclusion without first conducting a positive analysis, a policy debate is better informed if positive analyses are conducted first.[4] For instance, if your normative belief is that the government should help the poor, should you vote for a candidate who advocates a higher minimum wage (a law that requires that firms pay wages at or above a specified level)? One who believes in a European-style welfare system (guaranteeing health care, housing, and other basic goods and services)? A politician who wants an end to our current welfare system? Someone who wants to implement a negative income tax (in which the less income a person has, the more the government gives that person)? Or a candidate who favors job training programs? Positive economic analysis can be used to predict whether these programs will benefit poor people but not whether they are good or bad. Using these predictions and your value judgment, you can decide for whom to vote.

Economists' emphasis on positive analysis has implications for what we study and even our use of language. For example, many economists stress that they study people's *wants* rather than their *needs*. Although people need certain minimum levels of food, shelter, and clothing to survive, most people in developed economies have enough money to buy goods well in excess of the minimum levels necessary to maintain life. Consequently, in wealthy countries, calling something a "need" is often a value judgment. You almost certainly have been told by some elder that "you *need*

---

[4] Some economists draw the normative conclusion that, as social scientists, economists *should* restrict ourselves to positive analyses. Others argue that we shouldn't give up our right to make value judgments just like the next person (who happens to be biased, prejudiced, and pigheaded, unlike us).

a college education." That person was probably making a value judgment—"you *should* go to college"—rather than a scientific prediction that you will suffer terrible economic deprivation if you do not go to college. We can't test such value judgments, but we can test a hypothesis such as "One-third of the college-age population *wants* to go to college at current prices."

1.3 Uses of Microeconomic Models

Because microeconomic models *explain* why economic decisions are made and allow us to make *predictions*, they can be very useful for individuals, governments, and firms in making decisions. Throughout this book, we consider examples of how microeconomics aids in actual decision making.

Individuals can use microeconomics to make purchasing and other decisions (Chapters 4 and 5). Consumers' purchasing and investing decisions are affected by inflation and cost of living adjustments (Chapter 5). Whether it pays financially to go to college depends, in part, on interest rates (Chapter 16). Consumers decide for whom to vote based in part on candidates' views on economic issues.

Firms must decide which production methods to use to minimize cost (Chapter 7) and maximize profit (starting with Chapter 8). They may choose a complex pricing scheme or advertise to raise profits (Chapter 12). They select strategies to maximize profit when competing with a small number of other firms (Chapters 13 and 14). Some firms reduce consumer information to raise profits (Chapter 19). Firms use economic principles to structure contracts with other firms (Chapter 20).

Your government's elected and appointed officials use (or could use) economic models in many ways. Recent administrations have placed increased emphasis on economic analysis. Today, economic and environmental impact studies are required before many projects can commence. The President's Council of Economic Advisers and other federal economists analyze and advise national government agencies on the likely economic effects of all major policies.

One major use of microeconomic models by governments is to predict the probable impact of a policy before it is adopted. For example, economists predict the likely impact of a tax on the prices consumers pay and on the tax revenues raised (Chapter 3), whether a price control will create a shortage (Chapter 2), the differential effects of tariffs and quotas on trade (Chapter 9), and the effects of regulation on monopoly price and the quantity sold (Chapter 11).

Summary

1. **Microeconomics: The Allocation of Scarce Resources.** Microeconomics is the study of the allocation of scarce resources. Consumers, firms, and governments must make allocation decisions. The three key trade-offs a society faces are which goods and services to produce, how to produce them, and who gets them. These decisions are interrelated and depend on the prices that consumers and firms face and on government actions. Market prices affect the decisions of individual consumers and firms, and the interaction of the decisions of individual consumers and firms determines market prices. The organization of the market, especially the number of firms in the market and the information consumers and firms have, plays an important role in determining whether the market price is equal to or higher than marginal cost.

2. **Models.** Models based on economic theories are used to predict the future or to answer questions about how some change, such as a tax increase, will affect various sectors of the economy. A good theory is simple to use and makes clear, testable predictions that are not

refuted by evidence. Most microeconomic models are based on maximizing behavior. Economists use models to construct *positive* hypotheses concerning how a cause leads to an effect. These positive questions can be tested. In contrast, *normative* statements, which are value judgments, cannot be tested.

3. **Uses of Microeconomic Models.** Individuals, governments, and firms use microeconomic models and predictions to make decisions. For example, to maximize its profits, a firm needs to know consumers' decision-making criteria, the trade-offs between various ways of producing and marketing its product, government regulations, and other factors. For large companies, beliefs about how a firm's rivals will react to its actions play a critical role in how it forms its business strategies.

2 Supply and Demand

Talk is cheap because supply exceeds demand.

Challenge

Quantities and Prices of Genetically Modified Foods

Countries around the globe are debating whether to permit firms to grow or sell genetically modified (GM) foods, which have their DNA altered through genetic engineering rather than through conventional breeding.[1] The introduction of GM techniques can affect both the quantity of a crop that farmers supply and whether consumers want to buy that crop.

At least 29 countries grow GM food crops, which are mostly herbicide-resistant varieties of corn (maize), soybean, and canola (oilseed rape). Developing countries grow more GM crops than developed countries, though the United States plants 40% of worldwide GM acreage. As of 2015, the largest GM-producing countries are (in order) the United States, Brazil, Argentina, India, Canada, and China. European farmers produce virtually none of the world's GM crops, growing them on only 0.1% of their cultivatable land.

According to some polls, 70% of consumers in Europe object to GM foods. Fears cause some consumers to refuse to buy a GM crop (or the entire crop if consumers cannot distinguish non-GM from GM products). Consumers in other countries, such as the United States, are less concerned about GM foods. However, even in the United States, a 2015 poll found that 57% of U.S. consumers believe that GM foods are generally unsafe to eat, whereas 88% of scientists believe they are generally safe. The U.S. National Academy of Science reported in 2016 that they could find no evidence to support claims that genetically modified organisms are dangerous for either the environment or human health.

As of 2016, 65 nations require labeling of GM foods, including European Union countries, Japan, Australia, Brazil, Russia, China, and the United States. Consumers are unlikely to avoid GM crops if products are unlabeled.

Will the use of GM seeds lead to lower prices and more food sold? What happens to prices and quantities sold if many consumers refuse to buy GM crops? We will use the models in this chapter to answer these questions at the end of the chapter.

To analyze questions concerning the price and quantity responses to the introduction of new products or technologies, new government regulations or taxes, or other events, economists may use the *supply-and-demand model*. When asked, "What is the most important thing you know about economics?" a common reply is, "Supply equals demand." This statement is a shorthand description of one of the simplest yet most powerful models of economics. The supply-and-demand model describes how consumers and suppliers interact to determine the quantity and price of a good or

[1]Sources for Challenges, which appear at the beginning of chapters, and Applications, which appear throughout the chapters, are listed at the end of the book.

service. To use the model, you need to determine three things: buyers' behavior, sellers' behavior, and how they interact.

After reading this chapter, you should be adept enough at using the supply-and-demand model to analyze some of the most important policy questions facing your country today, such as those concerning international trade, minimum wages, and price controls on health care.

After reading that grandiose claim, you may ask, "Is that all there is to economics? Can I become an expert economist that fast?" The answer to both these questions is no, of course not. In addition, you need to learn the limits of this model and what other models to use when this one does not apply. (You must also learn the economists' secret handshake.)

Even with its limitations, the supply-and-demand model is the most widely used economic model. It provides a good description of how competitive markets function. *Competitive markets* are those with many buyers and sellers, such as most agriculture markets, labor markets, and stock and commodity markets. Like all good theories, the supply-and-demand theory can be tested—and possibly shown to be false. But in competitive markets, where it works well, it allows us to make accurate predictions easily.

In this chapter, we examine six main topics

1. **Demand.** The quantity of a good or service that consumers demand depends on price and other factors, such as consumers' incomes and the price of related goods.
2. **Supply.** The quantity of a good or service that firms supply depends on price and other factors, such as the cost of inputs firms use to produce the good or service.
3. **Market Equilibrium.** The interaction between consumers' demand and firms' supply determines the market price and quantity of a good or service that is bought and sold.
4. **Shocking the Equilibrium.** Changes in a factor that affect demand (such as consumers' incomes), supply (such as a rise in the price of inputs), or a new government policy (such as a new tax) alter the market price and quantity of a good.
5. **Effects of Government Interventions.** Government policies may alter the equilibrium and cause the quantity supplied to differ from the quantity demanded.
6. **When to Use the Supply-and-Demand Model.** The supply-and-demand model applies only to competitive markets.

2.1 Demand

Potential consumers decide how much of a good or service to buy on the basis of its price and many other factors, including consumers' tastes, information, prices of other goods, income, and government actions. Before concentrating on the role of price in determining demand, let's look briefly at some of the other factors.

Consumers' *tastes* determine what they buy. Consumers do not purchase foods they dislike, artwork they hate, or clothes they view as unfashionable or uncomfortable. Advertising may influence people's tastes.

Similarly, *information* (or misinformation) about the characteristics of a good affects consumers' decisions. A number of years ago when many consumers were convinced that oatmeal could lower their cholesterol level, they rushed to grocery stores and bought large quantities of oatmeal. (They even ate some of it until they remembered that they couldn't stand how it tastes.)

The *prices of other goods* also affect consumers' purchase decisions. Before deciding to buy Levi's jeans, you might check the prices of other brands. If the price of a close *substitute*—a product that you view as similar or identical to the one you are considering purchasing—is much lower than the price of Levi's jeans, you may buy that brand instead. Similarly, the price of a *complement*—a good that you like to consume at the same time as the product you are considering buying—may affect your decision. If you eat pie only with ice cream, the higher the price of ice cream, the less likely you are to buy pie.

Income plays a major role in determining what and how much to purchase. People who suddenly inherit great wealth may purchase a Rolls-Royce or other luxury items and would probably no longer buy do-it-yourself repair kits.

Government rules and regulations affect purchase decisions. Sales taxes increase the price that a consumer must pay for a good, and government-imposed limits on the use of a good may affect demand. In the nineteenth century, one could buy Bayer heroin, a variety of products containing cocaine, and other drug-related products that most countries ban today. When a city's government bans the use of skateboards on its streets, skateboard sales fall.[2]

Other factors may also affect the demand for specific goods. Consumers are more likely to use a particular smart phone app (application) if their friends use that one. The demand for small, dead evergreen trees is substantially higher in December than in other months.

quantity demanded
the amount of a good that consumers are willing to buy at a given price, holding constant the other factors that influence purchases

demand curve
the *quantity demanded* at each possible price, holding constant the other factors that influence purchases

Although many factors influence demand, economists usually concentrate on how price affects the quantity demanded. The relationship between price and quantity demanded plays a critical role in determining the market price and quantity in a supply-and-demand analysis. To determine how a change in price affects the quantity demanded, economists must hold constant other factors that affect demand, such as income and tastes.

The Demand Curve

The amount of a good that consumers are *willing* to buy at a given price, holding constant other factors that influence purchases, is the **quantity demanded**. The quantity demanded of a good or service can exceed the quantity *actually* sold.

For example, as a promotion, a local store might sell Ghirardelli 4.12 oz. dark chocolate bars for $1 each today only. At that low price, you might want to buy 10 chocolate bars, but because the store has only 5 remaining, you can buy at most 5 chocolate bars at this price. The quantity you demand at this price is 10 chocolate bars—it's the amount you want—even though the amount you actually buy is only 5.

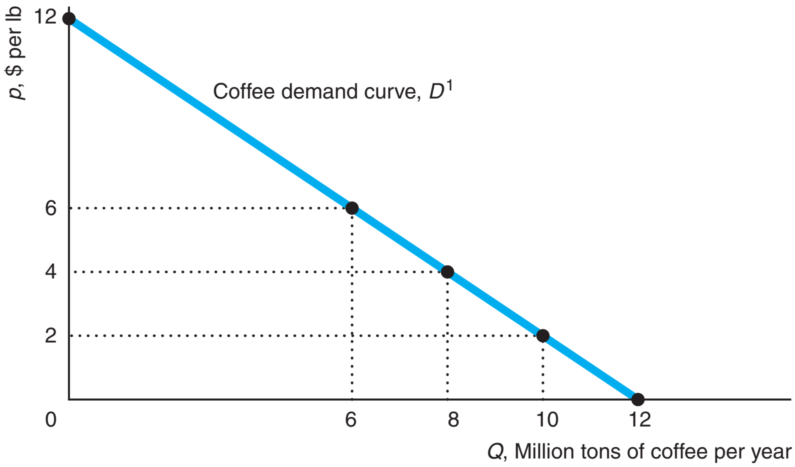
We can show the relationship between price and the quantity demanded graphically. A **demand curve** shows the quantity demanded at each possible price, holding constant other factors that influence purchases. Figure 2.1 shows the estimated annual demand curve, D^1, for green (unroasted) coffee beans.[3] (Although this demand curve is a straight line,

[2]When a Mississippi woman attempted to sell her granddaughter in exchange for $2,000 and a car, state legislators were horrified to discover that they had no law on the books prohibiting the sale of children and quickly passed such a law. (Mac Gordon, "Legislators Make Child-Selling Illegal," *Jackson Free Press*, March 16, 2009.)

[3]Because prices, quantities, and other factors change simultaneously over time, economists use statistical techniques to hold the effects of factors other than the price of the good constant so that they can determine how price affects the quantity demanded (see Appendix 2A at the back of the book). I estimated this model using data from the Food and Agriculture Organization, *Commodity Review and Outlook*; International Coffee Organization, **www.ico.org/new_historical.asp**; International Cocoa Organization, *The World Cocoa Economy: Past and Present* (July 2012); and World Bank, *World Development Indicators*.

Figure 2.1 A Demand Curve

The estimated global demand curve, D^1, for coffee shows the relationship between the annual quantity demanded and the price per lb. The downward slope of the demand curve shows that, holding other factors that influence demand constant, consumers demand a smaller quantity of this good when its price is high and a larger quantity when the price is low. A change in price causes a *movement along the demand curve*. For example, an increase in the price of coffee causes consumers to demand a smaller quantity of coffee.



demand curves may also be smooth curves or wavy lines.) By convention, the vertical axis of the graph measures the price, p, per unit of the good. Here, we measure the price of coffee in dollars per pound (abbreviated "lb"). The horizontal axis measures the quantity, Q, of the good in a *physical measure per period*. Here, we measure the quantity of coffee demanded in millions of tons per year.

The demand curve hits the vertical axis at $12, indicating that the quantity demanded is zero when the price is $12 per lb or higher. The demand curve hits the horizontal quantity axis at 12 million tons per year, which is the quantity of coffee that consumers would want if the price were zero. To find out what quantity is demanded at a price between zero and $4, we pick that price—say, $2—on the vertical axis, draw a horizontal line across until we hit the demand curve, and then draw a vertical line down to the horizontal quantity axis. As the figure shows, the quantity demanded at a price of $2 per lb is 10 million tons per year.

One of the most important things to know about the graph of a demand curve is what it does *not* show. All relevant economic variables that are not explicitly included in the demand curve graph—income, prices of other goods (such as other fruits or vegetables), tastes, information, and so on—are held constant. Thus, the demand curve shows how quantity varies with price but not how quantity varies with income, the price of substitute goods, tastes, information, or other variables. The demand curve is a concise summary of the answer to the question "What happens to the quantity demanded as the price changes, when all other factors are held constant?"

Law of Demand consumers demand more of a good the lower its price, holding constant tastes, the prices of other goods, and other factors that influence consumption

Effect of Prices on the Quantity Demanded Many economists claim that the most important *empirical* finding in economics is the **Law of Demand:** Consumers demand more of a good the lower its price, holding constant other factors that influence the amount they consume. According to the Law of Demand, *demand curves slope downward*, as in Figure 2.1.[4]

A downward-sloping demand curve illustrates that consumers demand a large quantity of this good at a low price and a small quantity at a high price. What happens to the quantity of coffee demanded if the price of coffee drops and all other

[4]Theoretically, a demand curve could slope upward (see Chapter 5); however, available empirical evidence strongly supports the Law of Demand.

variables remain constant? The quantity demanded increases. For example, the demand curve in Figure 2.1 shows that if the price decreases from \$6 to \$4 per lb, the quantity consumers demand increases by 2 million tons, from 6 to 8 million tons.[5] These changes in the quantity demanded in response to changes in price are *movements along the demand curve*.

Effects of Other Factors on Demand If a demand curve measures the effects of price changes when we hold constant all other factors that affect demand, how can we use demand curves to show the effects of a change in one of these other factors, such as household income? One solution is to draw the demand curve in a three-dimensional diagram with the price of coffee on one axis, the income on a second axis, and the quantity of coffee on the third axis. Just thinking about drawing such a diagram probably makes your head hurt. Moreover, what would you do if the demand curve depended on one more factor?

Economists use a simpler approach to show how quantity demanded is affected by a change in a factor other than price. A change in any factor other than the price of the good itself results in a *shift of the demand curve* rather than a *movement along the demand curve*.

If the average household income rises and the price of coffee remains constant, people buy more coffee. Suppose that the average income rises from \$35,000 per year to \$50,000, an increase of \$15,000. Figure 2.2 shows that the higher income causes the coffee demand curve to shift 1.5 units (million tons) to the right from D^1, where the average income is \$35,000, to D^2, where average income is \$50,000.

A change in other factors, such as the prices of *substitutes* and *complements*, may also cause a demand curve to shift. A **substitute** is a good or service that may be consumed instead of another good or service. For many people, tea is a substitute for coffee, so a decrease in the price of tea may cause their demand curve to shift to the left—less coffee is demanded at any given price of coffee. A **complement** is a good or service that is jointly consumed with another good or service. For example, many people drink coffee with sugar. An increase in the price of sugar would cause their demand curves to shift to the left.

Other factors also affect demand curves. For example, if cigarettes become more addictive, the demand curve of existing smokers shifts to the right.[6]

In summary, to properly analyze the effects of a change in a factor on the quantity demanded, we must distinguish between a *movement along a demand curve* and a *shift of a demand curve*. A change in the *price of a good* causes a *movement along a demand curve*. A change in *any other factor besides the price of the good* causes a *shift of the demand curve*.

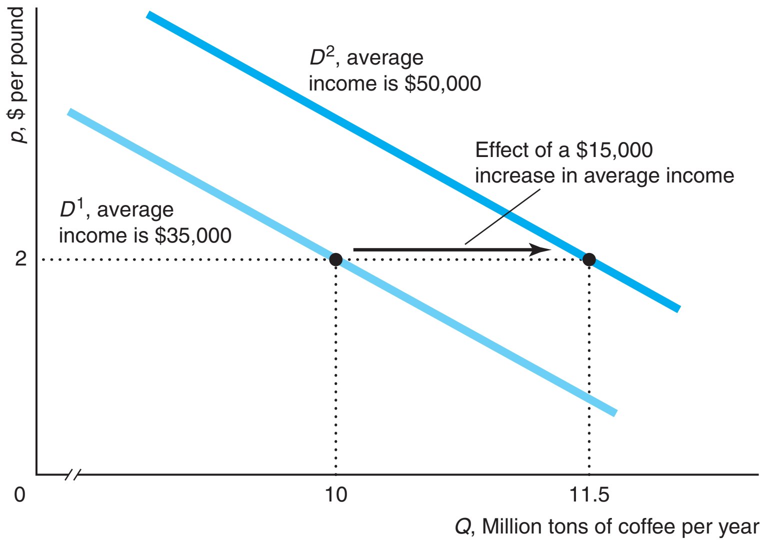
---

[5]Economists typically do not state the relevant physical and period measures unless they are particularly useful. They refer to *quantity* rather than something useful such as "metric tons per year" and *price* rather than "cents per pound." I'll generally follow this convention, usually referring to the price as \$4 (with the "per lb" understood) and the quantity as 8 (with the "million tons per year" understood).

[6]A Harvard School of Public Health study concluded that cigarette manufacturers raised nicotine levels in cigarettes by 11% from 1998 to 2005 to make them more addictive. Gardiner Harris, "Study Showing Boosted Nicotine Levels Spurs Calls for Controls," *San Francisco Chronicle*, January 19, 2007, A-4.

Figure 2.2 A Shift of the Demand Curve

The global coffee demand curve shifts rightward from D^1 to D^2 as average annual household income in high-income countries rises by \$15,000, from \$35,000 to \$50,000. At the higher income, a larger quantity of coffee is demanded at any given price.



Application

Calorie Counting

Information can also affect demand curves. New York City started requiring mandatory posting of calories on menus in chain restaurants in mid-2008. Bollinger, Leslie, and Sorensen (2011) found that New York City's calorie posting requirement caused average calories per transaction at Starbucks to fall by 6% due to reduced consumption of high-calorie foods. They found larger responses to information among wealthier and better-educated consumers and among those who prior to the law consumed relatively more calories.

Some other studies found less of a response to such information. Dingman et al. (2015) posted calorie information on vending machines selling food and sent email to residents of some college residence halls. In other halls, they provided no information. Comparing vending machine sales in the four weeks before and after they provided information, they found no change in behavior from the information. However, U.S. Department of Agriculture studies found that people who already have healthy diets habits—apparently unlike college students—*do* pay attention to this information (Stewart and Mentzer Morrison, 2015).

A U.S. Food and Drug Administration (FDA) rule requires that large restaurant chains include calorie information on their menus. In 2016, the FDA also set new calorie, nutritional, and serving size labeling requirements for packaged food, which becomes mandatory as of July 2018.

substitute
a good or service that may be consumed instead of another good or service

complement
a good or service that is jointly consumed with another good or service

demand function
the relationship between the quantity demanded, price, and other factors that influence purchases

The Demand Function

The **demand function** shows the relationship between the quantity demanded, price, and other factors that influence purchases. Other factors that may influence the quantity demanded include income, substitutes, complements, tastes, and consumer information. Graphically, we illustrate the effect of a change in one of these other relevant factors by shifting the demand curve. We can represent the same information mathematically—information about how price, income, and other variables affect quantity demanded—using a *demand function*. The demand function shows the effect of *all* the relevant factors on the quantity demanded.

We examine a demand function for coffee. The quantity of coffee demanded, Q, varies with the price of coffee, p, the price of sugar, p_s, and consumers' income, Y, so the coffee demand function, D, is

$$Q = D(p, p_s, Y). \tag{2.1}$$

We assume that any other factors that we do not explicitly list in the demand function are irrelevant (such as the price of llamas in Peru) or constant (such as the prices of substitutes and complements, tastes, and consumer information).

Equation 2.1 is a general functional form—it does not specify exactly how Q varies with the explanatory variables, p, p_s, and Y. The estimated demand function that corresponds to the demand curve D in Figure 2.1 and D^1 in Figure 2.2 has a specific (linear) form. Our estimated world demand function for green coffee beans is

$$Q = 8.56 - p - 0.3p_s + 0.1Y, \tag{2.2}$$

where Q is the quantity of coffee demanded in millions of tons per year, p is the price of coffee in dollars per pound (lb), p_s is the price of sugar in dollars per pound, and Y is the average annual household income in high-income countries in thousands of dollars.

When we draw the demand curve in Figures 2.1 and 2.2, we hold p_s and Y at specific values. In these figures, we use the average values of p_s, \$0.20 per lb, and Y, \$35 thousand per year. Substituting those values into Equation 2.2, we can write the quantity demanded as a function of only the price of coffee:

$$\begin{aligned} Q &= 8.56 - p - 0.3p_s + 0.1Y \\ &= 8.56 - p - (0.3 \times 0.2) + (0.1 \times 35) \\ &= 12 - p. \end{aligned} \tag{2.3}$$

The linear demand function in Equation 2.3 corresponds to the straight-line demand curve D in Figure 2.1. The constant term, 12, in Equation 2.3 is the quantity demanded (in millions of tons per year) if the price of coffee is zero. Setting the price equal to zero in Equation 2.3, we find that the quantity demanded is $Q = 12 - (1 \times 0) = 12$. Figure 2.1 shows that $Q = 12$ where D hits the quantity axis—where price is zero.

By plugging any particular value for p into Equation 2.3, we can determine the corresponding quantities. For example, if $p = \$2$, then $Q = 12 - 2 = 10$, as Figure 2.1 shows.

We can also use Equation 2.3 to determine how the quantity demanded varies with a change in price: a movement *along* the demand curve. If the price falls from p_1 to p_2, the change in price, Δp, equals $p_2 - p_1$. (The Δ symbol, the Greek letter delta, means "change in" the variable following the delta, so Δp means "change in price.") If the price of coffee falls from $p_1 = \$4$ to $p_2 = \$2$, then $\Delta p = p_2 - p_1 = \$2 - \$4 = -\$2$. The quantity demanded changes from $Q_1 = 8$ at a price of \$4 to $Q_2 = 10$ at a price of \$2, so $\Delta Q = Q_2 - Q_1 = 10 - 8 = 2$. That is, as price falls by \$2 per pound, the quantity demanded rises by 2 million tons per year.

More generally, the quantity demanded at p_1 is $Q_1 = D(p_1)$, and the quantity demanded at p_2 is $Q_2 = D(p_2)$. The change in the quantity demanded, $\Delta Q = Q_2 - Q_1$, in response to the price change (using Equation 2.3) is

$$\begin{aligned}\Delta Q &= Q_2 - Q_1\\ &= D(p_2) - D(p_1)\\ &= (12 - p_2) - (12 - p_1)\\ &= -(p_2 - p_1)\\ &= -\Delta p.\end{aligned}$$

Thus, the change in the quantity demanded, ΔQ, is -40 times the change in the price, Δp. For example, if $\Delta p = -\$2$, then $\Delta Q = -\Delta p = -(-2) = 2$ million tons per year.

This effect is consistent with the Law of Demand. A \$2 decrease in price causes an increase in quantity demanded of 2 million tons per year. Similarly, raising the price would cause the quantity demanded to fall.

The slope of a demand curve is $\Delta p/\Delta Q$—the "rise" (Δp, the change along the vertical axis) divided by the "run" (ΔQ, the change along the horizontal axis). The slope of the demand curve D^1 in Figures 2.1 and 2.2 is

$$\text{Slope} = \frac{\text{rise}}{\text{run}} = \frac{\Delta p}{\Delta Q} = \frac{\$1 \text{ per lb}}{-1 \text{ million tons per year}} = -\$1 \text{ per million tons per year.}$$

The negative sign of this slope is consistent with the Law of Demand. The slope says that the price rises by \$1 per lb as the quantity demanded falls by 1 million tons per year.

Thus, we can use the demand curve to answer questions about how a change in price affects the quantity demanded and how a change in the quantity demanded affects price. We can also answer these questions using demand functions.

Solved Problem 2.1

MyLab Economics Solved Problem

How much would the price have to fall for consumers to be willing to buy 1 million more tons of coffee per month?

Answer

1. *Express the price that consumers are willing to pay as a function of quantity.* We use algebra to rewrite the demand function as an *inverse demand function*, where price depends on the quantity demanded. Subtracting Q from both sides of Equation 2.3, $Q = 12 - p$, and adding p to both sides, we obtain the inverse demand function:

$$p = 4 - Q \tag{2.4}$$

2. *Use the inverse demand curve to determine how much the price must change for consumers to buy 1 million more tons of coffee per year.* We want the new quantity, Q_2, to equal the original quantity, Q_1, plus one: $Q_2 = Q_1 + 1$. Using the inverse demand function, Equation 2.4, we can determine by how much the price must change:

$$\begin{aligned}\Delta p &= p_2 - p_1\\ &= (1 - Q_2) - (1 - Q_1)\\ &= -(Q_2 - Q_1)\\ &= -\Delta Q.\end{aligned}$$

The change in quantity is $\Delta Q = Q_2 - Q_1 = (Q_1 + 1) - Q_1 = 1$, so the change in price is $\Delta p = -1$. That is, for consumers to demand 1 million more tons of coffee per year, the price must fall by \$1 per pound, which is a *movement along the demand curve.*

Summing Demand Curves

If we know the demand curve for each of two consumers, how do we determine the total demand curve for the two consumers combined? The total quantity demanded *at a given price* is the sum of the quantity each consumer demands at that price.

We can use the demand functions to determine the total demand of several consumers. Suppose that the demand function for Consumer 1 is

$$Q_1 = D^1(p)$$

and the demand function for Consumer 2 is

$$Q_2 = D^2(p).$$

At price p, Consumer 1 demands Q_1 units, Consumer 2 demands Q_2 units, and the total demand of both consumers is the sum of the quantities each demands separately:

$$Q = Q_1 + Q_2 = D^1(p) + D^2(p).$$

We can generalize this approach to look at the total demand for three or more consumers.

It makes sense to add the quantities demanded only when all consumers face the same price. Adding the quantity Consumer 1 demands at one price to the quantity Consumer 2 demands at another price would be like adding apples and oranges.

Application

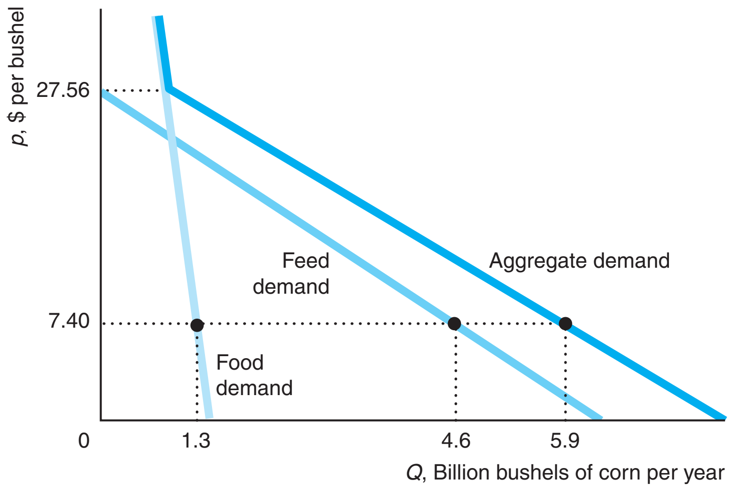
Aggregating Corn Demand Curves

We illustrate how to sum individual demand curves to get an aggregate demand curve using estimated demand curves for corn (McPhail and Babcock, 2012). The figure shows the U.S. feed demand (the use of corn to feed animals) curve, the U.S. food demand curve, and the aggregate demand curve from these two uses.[7]

To derive the sum of the quantity demanded for these two uses at a given price, we add the quantities from the individual demand curves at that price. That is, we add the demand curves horizontally. At a price for corn of \$7.40, the quantity demanded for food is 1.3 billion bushels per year and the quantity demanded for feed is 4.6 billion bushels. Thus, the total quantity demanded at that price is $Q = 1.3 + 4.6 = 5.9$ billion bushels.

When the price of corn exceeds \$27.56 per bushel, farmers stop using corn for animal feed, so the quantity demanded for this use equals zero. Thus, the total demand curve is the same as the food demand curve at prices above \$27.56.

[7]For graphical simplicity, we do not show the other major U.S. demand curves for export, storage, and use in biofuels (ethanol). Thus, this aggregate demand curve is not the total demand curve for corn.



2.2 Supply

Knowing how much consumers want is not enough by itself for us to determine the market price and quantity. We also need to know how much firms want to supply at any given price.

Firms determine how much of a good to supply based on the price of that good and other factors, including the costs of production and government rules and regulations. Usually, we expect firms to supply more at a higher price. Before concentrating on the role of price in determining supply, we'll briefly describe the role of some of the other factors.

Costs of production affect how much firms want to sell of a good. As a firm's cost falls, it is willing to supply more, all else the same. If the firm's cost exceeds what it can earn from selling the good, the firm sells nothing. Thus, factors that affect costs also affect supply. A technological advance that allows a firm to produce a good at lower cost leads the firm to supply more of that good, all else the same.

Government rules and regulations affect how much firms want to sell or whether they may sell a product. Taxes and many government regulations—such as those covering pollution, sanitation, and health insurance—alter the cost of production. Other regulations affect when and how suppliers may sell their products. In some countries, retailers may not sell most goods and services on days of particular religious significance. U.S. city and state governments prohibit the sale of cigarettes and liquor to children.

quantity supplied
the amount of a good that firms *want* to sell at a given price, holding constant other factors that influence firms' supply decisions, such as costs and government actions

supply curve
the *quantity supplied* at each possible price, holding constant the other factors that influence firms' supply decisions

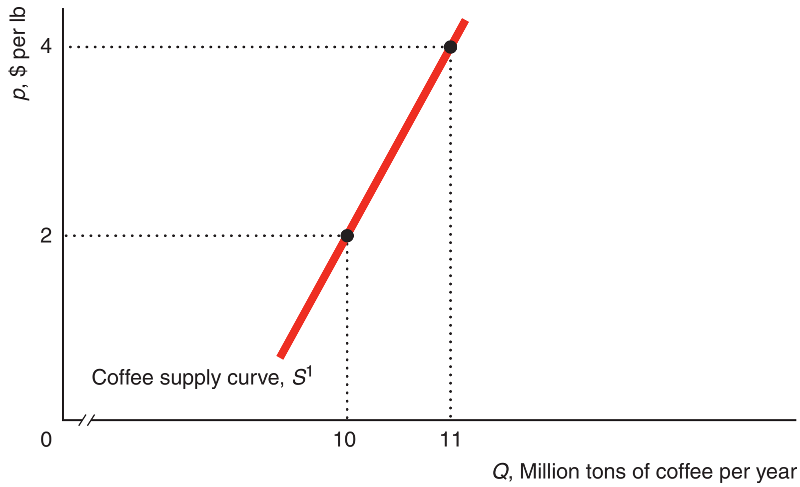
The Supply Curve

The **quantity supplied** is the amount of a good that firms *want* to sell at a given price, holding constant other factors that influence firms' supply decisions, such as costs and government actions. We can show the relationship between price and the quantity supplied graphically. A **supply curve** shows the quantity supplied at each possible price, holding constant the other factors that influence firms' supply decisions.

Figure 2.3 shows the estimated supply curve, S^1, for coffee. As with the demand curve, the price on the vertical axis is measured in dollars per physical unit (dollars per lb), and the quantity on the horizontal axis is measured in physical units per period (millions of tons per year). Because we hold fixed other variables that may

Figure 2.3 A Supply Curve

The estimated global supply curve, S^1, for coffee shows the relationship between the quantity supplied per year and the price per lb, holding constant cost and other factors that influence supply. The upward slope of this supply curve indicates that firms supply more coffee when its price is high and less when the price is low. An increase in the price of coffee causes firms to supply a larger quantity of coffee; any change in price results in a *movement along the supply curve*.



affect the supply, such as costs and government rules, the supply curve concisely answers the question "What happens to the quantity supplied as the price changes, holding all other factors constant?"

Effects of Price on Supply We illustrate how price affects the quantity supplied using the coffee supply curve in Figure 2.3. The supply curve is upward sloping. As the price increases, firms supply more. If the price is \$2 per lb, the quantity supplied by the market is 10 million tons per year. If the price rises to \$4, the quantity supplied rises to 11 million tons. An increase in the price of coffee causes a *movement along the supply curve*—firms supply more coffee.

Although the Law of Demand states that the demand curve slope downward, we have *no* "Law of Supply" that requires the market supply curve to have a particular slope. The market supply curve can be upward sloping, vertical, horizontal, or downward sloping. Many supply curves slope upward, such as the one for coffee. Along such supply curves, the higher the price, the more firms are willing to sell, holding costs and government regulations fixed.

Effects of Other Variables on Supply A change in a factor other than a product's price causes a *shift of the supply curve*. Suppose the price of cocoa (which is a key input in making chocolate) increases by \$3 from \$3 to \$6 per lb. The land on which coffee is grown is also suitable to grow cocoa. When the price of cocoa rises, some coffee farmers switch to producing cocoa. Therefore, when the price of cocoa rises, the amount of coffee produced at any given coffee price falls.

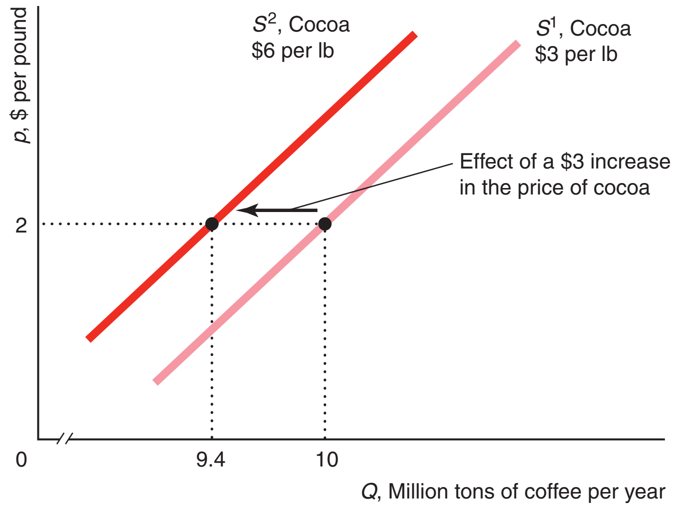
In Figure 2.4, S^1 is the supply curve of coffee before the price of cocoa increases, and S^2 is the supply curve after the price of cocoa rises. The increase in the price of cocoa causes the coffee supply curve to shift to the *left*, from S^1 to S^2.[8] That is, firms want to supply less coffee at any given price than before the increase in the price of cocoa. At a coffee price of \$2 per lb, the quantity of coffee supplied falls from 12 million tons on S^1 to 9.4 million on S^2.

Again, it is important to distinguish between a *movement along a supply curve* and a *shift of the supply curve*. When the coffee price changes, the change in the quantity

[8]Alternatively, we may say that the supply curve shifts *up* because firms will supply a given quantity only at a higher price.

Figure 2.4 A Shift of a Supply Curve

A \$3 per lb increase in the price of cocoa, which farmers can grow instead of coffee, causes the supply curve for coffee to shift left from S^1 to S^2. At a price for coffee of \$2 per lb, the quantity supplied falls from 10 million tons on S^1 to 9.4 million tons on S^2.



supplied reflects a *movement along the supply curve*. When costs, government rules, or other variables that affect supply change, the entire *supply curve shifts*.

The Supply Function

supply function shows the correspondence between the quantity supplied, price, and other factors that influence the number of units offered for sale

The **supply function** shows the relationship between the quantity supplied, price, and other factors that influence the number of units offered for sale. Written generally (without specifying the functional form), the coffee supply function is

$$Q = S(p, p_c), \tag{2.5}$$

where Q is the quantity of coffee supplied, p is the price of coffee, and p_c is the price of cocoa.

The supply function, Equation 2.5, might also incorporate other factors such as wages, transportation costs, and the state of technology. By leaving them out, we are implicitly holding them constant.

Our estimate of the coffee supply function is

$$Q = 9.6 + 0.5p - 0.2p_c, \tag{2.6}$$

where Q is the quantity of coffee in millions of tons per year, p is the price of coffee in dollars per lb, and p_c is the price of cocoa in dollars per lb.

If we fix the cocoa price at \$3 per lb, we can rewrite the supply function in Equation 2.6 as a function solely of the coffee price. Substituting $p_c = \$3$ into Equation 2.6, we find that

$$Q = 9.6 + 0.5p - (0.2 \times 3) = 9 + 0.5p. \tag{2.7}$$

Because we hold fixed other variables that may affect the quantity supplied, such as costs and government rules, this supply function concisely answers the question "What happens to the quantity supplied as the price changes, holding all other factors constant?"

What happens to the quantity supplied if the price of coffee increases by $\Delta p = p_2 - p_1$? As the price increases from p_1 to p_2, the quantity supplied goes from Q_1 to Q_2, so the change in quantity supplied is

$$\Delta Q = Q_2 - Q_1 = (9.6 + 0.5p_2) - (9.6 + 0.5p_1) = 0.5(p_2 - p_1) = 0.5\Delta p.$$

Thus, a \$1 increase in price ($\Delta p = 1$) causes the quantity supplied to increase by $\Delta Q = 0.5$ million tons per year. This change in the quantity of coffee supplied as p increases is a *movement along the supply curve*.

Summing Supply Curves

The total supply curve shows the total quantity produced by all suppliers at each possible price. For example, the total supply of rice in Japan is the sum of the domestic and foreign supply curves of rice.

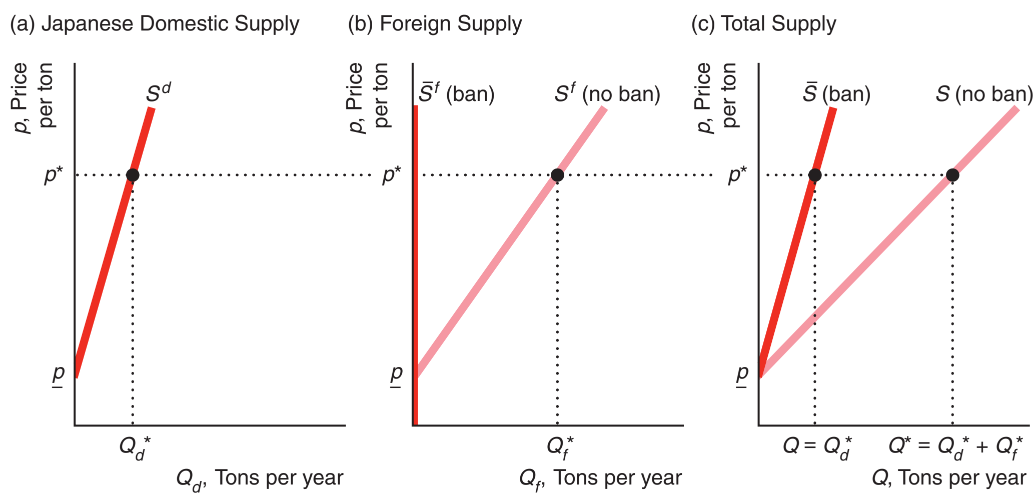
Suppose that the domestic supply curve (panel a) and foreign supply curve (panel b) of rice in Japan are as Figure 2.5 shows. The total supply curve, S in panel c, is the horizontal sum of the Japanese *domestic* supply curve, S^d, and the *foreign* supply curve, S^f. In the figure, the Japanese and foreign supplies are zero at any price equal to or less than $\underline{p}$, so the total supply is zero. At prices above $\underline{p}$, the Japanese and foreign supplies are positive, so the total supply is positive. For example, when price is p^*, the quantity supplied by Japanese firms is Q_d^* (panel a), the quantity supplied by foreign firms is Q_f^* (panel b), and the total quantity supplied is $Q^* = Q_d^* + Q_f^*$ (panel c). Because the total supply curve is the horizontal sum of the domestic and foreign supply curves, the total supply curve is flatter than the other two supply curves.

How Government Import Policies Affect Supply Curves

We can use this approach for deriving the total supply curve to analyze the effect of government policies on the total supply curve. Traditionally, the Japanese government has banned the importation of foreign rice. We want to determine how the ban affects the supply curve in the Japanese market.

Figure 2.5 Total Supply: The Sum of Domestic and Foreign Supply

If foreigners may sell their rice in Japan, the total Japanese supply of rice, S, is the horizontal sum of the domestic Japanese supply, S^d, and the imported foreign supply, S^f. With a ban on foreign imports, the foreign supply curve, $\bar{S}^f$, is zero at every price, so the total supply curve, $\bar{S}$, is the same as the domestic supply curve, S^d.



Without a ban, the foreign supply curve is S^f in panel b of Figure 2.5. A ban on imports eliminates the foreign supply, so the foreign supply curve after the ban is imposed, $\overline{S}^f$, is a vertical line at $Q_f = 0$. The import ban had no effect on the domestic supply curve, S^d, so the supply curve remains the same as in panel a.

Because the foreign supply with a ban, $\overline{S}^f$, is zero at every price, the total supply with a ban, $\overline{S}$ in panel c is the same as the Japanese domestic supply, S^d, at any given price. The total supply curve under the ban lies to the left of the total supply curve without a ban, S. Thus, the effect of the import ban is to rotate the total supply curve toward the vertical axis.

quota
the limit that a government sets on the quantity of a foreign-produced good that may be imported

The limit that a government sets on the quantity that may be imported of a foreign-produced good is called a **quota**. By absolutely banning the importation of rice, the Japanese government set a quota of zero on rice imports. Sometimes governments set positive quotas, $\overline{Q} > 0$. Foreign firms may supply as much as they want, Q_f, as long as they supply no more than the quota: $Q_f \le \overline{Q}$.

We investigate the effect of such a quota in Solved Problem 2.2. In most of the solved problems in this book, you are asked to determine how a *change* in a variable or policy *affects* one or more variables. In this problem, the policy *changes* from no quota to a quota, which *affects* the total supply curve.

Solved Problem 2.2

MyLab Economics Solved Problem

How does the quota, $\overline{Q}$, set by the United States on foreign sugar imports affect the total American supply curve for sugar given the domestic supply curve, S^d in panel a of the graph, and the foreign supply curve, S^f in panel b?

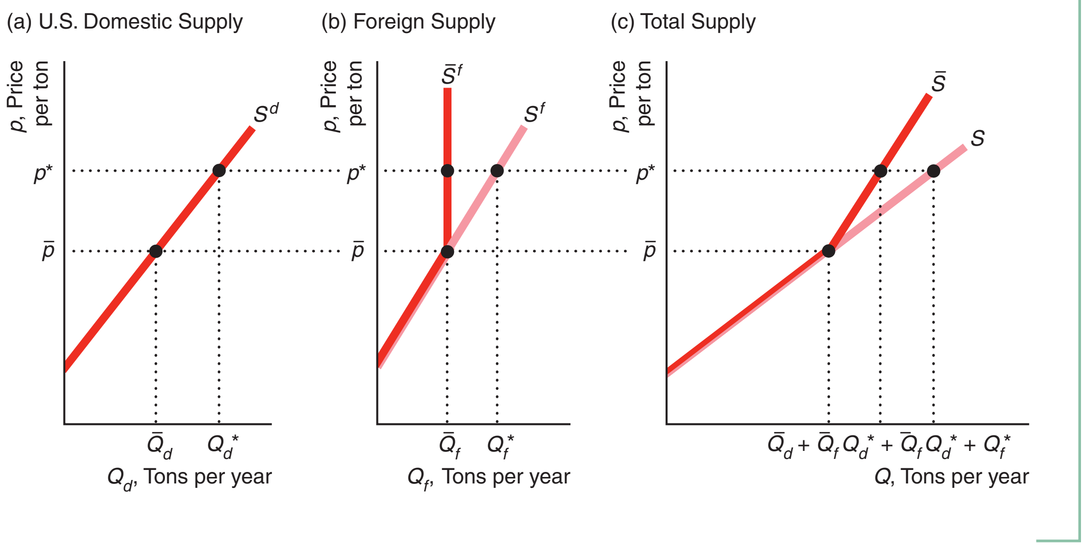
Answer

1. *Determine the American supply curve without the quota.* The *no-quota* total supply curve, S in panel c, is the horizontal sum of the U.S. domestic supply curve, S^d, and the no-quota foreign supply curve, S^f.

2. *Show the effect of the quota on foreign supply.* At prices less than $\overline{p}$, foreign suppliers want to supply quantities less than the quota, $\overline{Q}$. As a result, the foreign supply curve under the quota, $\overline{S}^f$, is the same as the no-quota foreign supply curve, S^f, for prices less than $\overline{p}$. At prices above $\overline{p}$, foreign suppliers want to supply more but are limited to $\overline{Q}$. Thus, the foreign supply curve with a quota, $\overline{S}^f$, is vertical at $\overline{Q}$ for prices above $\overline{p}$.

3. *Determine the American total supply curve with the quota.* The total supply curve with the quota, $\overline{S}$, is the horizontal sum of S^d and $\overline{S}^f$. At any price above $\overline{p}$, the total supply equals the quota plus the domestic supply. For example, at p^*, the domestic supply is Q_d^* and the foreign supply is $\overline{Q}_f$, so the total supply is $Q_d^* + \overline{Q}_f$. Above $\overline{p}$, $\overline{S}$ is the domestic supply curve shifted $\overline{Q}$ units to the right. As a result, the portion of $\overline{S}$ above $\overline{p}$ has the same slope as S^d.

4. *Compare the American total supply curves with and without the quota.* At prices less than or equal to $\overline{p}$, the same quantity is supplied with and without the quota, so $\overline{S}$ is the same as S. At prices above $\overline{p}$, less is supplied with the quota than without one, so $\overline{S}$ is steeper than S, indicating that a given increase in price raises the quantity supplied less with a quota than without one.



2.3 Market Equilibrium

The supply and demand curves determine the price and quantity of goods and services in a market. The demand curve shows the quantity consumers want to buy at various prices, and the supply curve shows the quantity firms want to sell at various prices. Unless the price is set so that consumers want to buy exactly the same amount that suppliers want to sell, either some buyers cannot buy as much as they want or some sellers cannot sell as much as they want.

When all traders are able to buy or sell as much as they want, we say that the market is in **equilibrium**: a situation in which no one wants to change his or her behavior. The *equilibrium price* is the price at which consumers can buy as much as they want and sellers can sell as much as they want. The *equilibrium quantity* is the amount that consumers buy and suppliers sell at the equilibrium price.

equilibrium
a situation in which no one wants to change his or her behavior

Using a Graph to Determine the Equilibrium

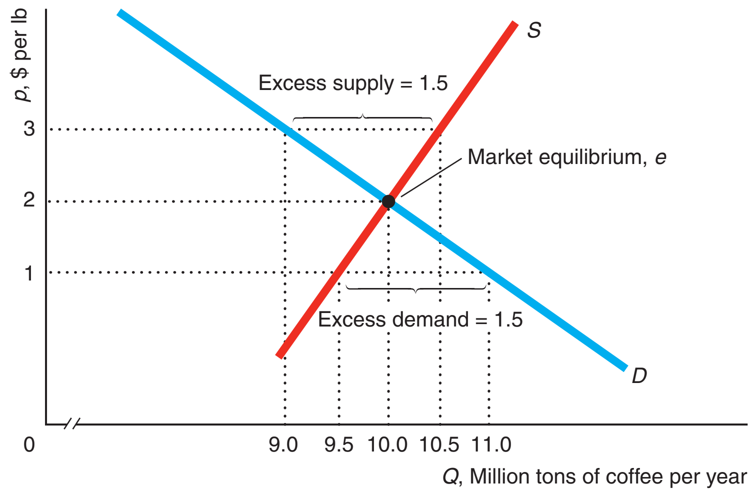
To illustrate how supply and demand curves determine the equilibrium price and quantity, we return to our old friend, the coffee example. Figure 2.6 shows the supply, *S*, and demand, *D*, curves for coffee. The supply and demand curves intersect at point *e*, the market equilibrium. The equilibrium price is $2 per lb, and the equilibrium quantity is 10 million tons per year, which is the quantity firms want to sell and consumers want to buy at the equilibrium price.

Using Math to Determine the Equilibrium

We can determine the equilibrium mathematically, using algebraic representations of the supply and demand curves. We use these two equations to solve for the equilibrium price at which the quantity demanded equals the quantity supplied

Figure 2.6 Market Equilibrium

The intersection of the supply curve, S, and the demand curve, D, for coffee determines the market equilibrium point, e, where the equilibrium price is $p = \$2$ per lb and the equilibrium quantity is $Q = 10$ million tons per year. At a price of $p = \$1$, which is below the equilibrium price, the quantity demanded is 11, but the quantity supplied is only 9.5, so the excess demand is 1.5. At $p = \$3$, a price that exceeds the equilibrium price, the market has an excess supply of 1.5 because the quantity demanded, 9, is less than the quantity supplied, 10.5. With either excess demand or excess supply, market forces drive the price back to the equilibrium price of \$2.



(the equilibrium quantity). The demand curve, Equation 2.3, shows the relationship between the quantity demanded, Q_d, and the price:[9]

$$Q_d = 12 - p.$$

The supply curve, Equation 2.7, tells us the relationship between the quantity supplied, Q_s, and the price:

$$Q_s = 9 + 0.5p.$$

We want to find the equilibrium price, p, at which $Q_d = Q_s = Q$, the equilibrium quantity. Thus, we set the right sides of these two equations equal,

$$9 + 0.5p = 12 - p,$$

and solve for the equilibrium price. Adding p to both sides of this expression and subtracting 9 from both sides, we find that $1.5p = 3$. Dividing both sides of this last expression by 1.5, we learn that the equilibrium price is $p = \$2$.

We can determine the equilibrium quantity by substituting this equilibrium price, $p = \$2$, into either the supply or the demand equation:

$$\begin{aligned} Q_d &= Q_s \\ 12 - (1 \times 2) &= 9 + (0.5 \times 2) \\ 10 &= 10. \end{aligned}$$

Thus, the equilibrium quantity, $Q = Q_d = Q_s$, is 10 million tons per year.

[9]Usually, we use Q to represent both the quantity demanded and the quantity supplied. However, for clarity in this discussion, we use Q_d and Q_s.

Forces That Drive the Market to Equilibrium

A market equilibrium is not just an abstract concept or a theoretical possibility.[10] We observe markets in equilibrium. The ability to buy as much as you want of a good at the market price is indirect evidence that a market is in equilibrium. You can almost always buy as much as you want of milk, ballpoint pens, and many other goods.

Amazingly, a market equilibrium occurs without any explicit coordination between consumers and firms. In a competitive market such as that for agricultural goods, millions of consumers and thousands of firms make their buying and selling decisions independently. Yet each firm can sell as much as it wants, and each consumer can buy as much as he or she wants. It is as though an unseen market force, like an *invisible hand*, directs people to coordinate their activities to achieve market equilibrium.

What really causes the market to be in equilibrium? If the price were not at the equilibrium level, consumers or firms would have an incentive to change their behavior in a way that would drive the price to the equilibrium level.

If the price were initially lower than the equilibrium price, consumers would want to buy more than suppliers would want to sell. For example, if the coffee price were \$1 in Figure 2.6, which is less than the equilibrium price, consumers would demand 11 million tons per year, but firms would be willing to supply only 9.5 million tons. At this price, the market would be in *disequilibrium*, meaning that the quantity demanded would not equal the quantity supplied. The market would have **excess demand**—the amount by which the quantity demanded exceeds the quantity supplied at a specified price—of $11 - 9.5 = 1.5$ million tons per year.

excess demand
the amount by which the *quantity demanded* exceeds the *quantity supplied* at a specified price

Some consumers would be lucky enough to be able to buy coffee at \$1. Other consumers would not find anyone willing to sell them coffee at that price. What could they do? Some frustrated consumers might offer to pay suppliers more than \$1. Alternatively, suppliers, noticing these disappointed consumers, might raise their prices. Such actions by consumers and producers would cause the market price to rise. At higher prices, the quantity that firms want to supply increases and the quantity that consumers want to buy decreases. The upward pressure on the price would continue until it reached the equilibrium price, \$2, where the market has no excess demand.

If, instead, the price is initially above the equilibrium level, suppliers want to sell more than consumers want to buy. For example, at a price for coffee of \$3, suppliers would want to sell 10.5 million tons per year, but consumers would want to buy only 9 million, as Figure 2.6 shows. Thus, at a price of \$3, the market would be in disequilibrium. The market would have **excess supply**—the amount by which the quantity supplied is greater than the quantity demanded at a specified price—of $10.5 - 9 = 1.5$ million tons. Not all firms could sell as much as they wanted. Rather than incur storage costs (and possibly have their unsold coffee spoil), firms might lower their price to attract additional customers. As long as the price remained above the equilibrium price, some firms would have unsold coffee and would want to lower the price further. The price would fall until it reached the equilibrium level, \$2, without excess supply and hence no pressure to lower the price further.[11]

excess supply
the amount by which the *quantity supplied* is greater than the *quantity demanded* at a specified price

[10] MyLab Economics has games (called *experiments*) for your course. These online games allow you to play against the computer. The *Market Experiment* illustrates the operation of the supply-and-demand model, allowing you to participate in a simulated market. To play, go to MyLab Economics Multimedia Library, Single Player Experiment, and set the Chapter field to "All Chapters."

[11] Not all markets reach equilibrium through the independent actions of many buyers or sellers. In institutionalized or formal markets, such as the Chicago Mercantile Exchange—where agricultural commodities, financial instruments, energy, and metals are traded—buyers and sellers meet at a single location (or on a single website). In these markets, certain individuals or firms, sometimes referred to as *market makers*, act to adjust the price and bring the market into equilibrium very quickly.

In summary, at any price other than the equilibrium price, either consumers or suppliers are unable to trade as much as they want. These disappointed buyers or suppliers act to change the price, driving the price to the equilibrium level. The equilibrium price is called the *market clearing price* because it removes from the market all frustrated buyers and sellers: The market has no excess demand or excess supply at the equilibrium price.

2.4 Shocking the Equilibrium

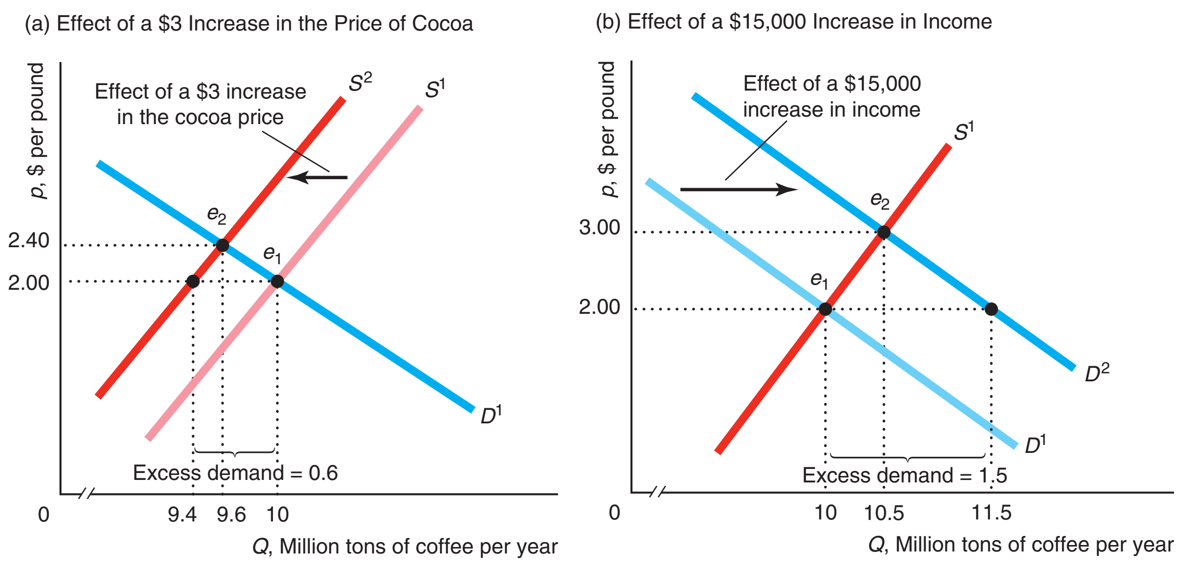
If the variables we hold constant in the demand and supply curves do not change, an equilibrium would persist indefinitely because none of the participants in the market would apply pressure to change the price. The equilibrium changes only if a shock occurs that shifts the demand curve or the supply curve. These curves shift if one of the variables we are holding constant changes. If tastes, income, government policies, or costs of production change, the demand curve or the supply curve or both shift, and the equilibrium changes.

Effects of a Shock to the Supply Curve

Suppose only one variable changes: The price of cocoa increases by \$3 per pound. However, as panel a of Figure 2.7 shows, the increase in the price of cocoa causes the coffee supply curve to shift 0.6 units to the left from S^1 to S^2 at every possible price of

Figure 2.7 Equilibrium Effects of a Shift of a Demand or Supply Curve MyLab Economics Video

(a) A \$3 per pound increase in the price of cocoa causes some farmers to shift from coffee production to cocoa production, reducing the quantity of coffee supplied at every price. The supply curve shifts to the left from S^1 to S^2. At the original equilibrium price of \$2, the market has excess demand of 0.6 million tons per year. Market pressures drive the market equilibrium from e_1 to e_2, where the new equilibrium price is \$2.40. (b) A \$15,000 increase in average annual household income causes the demand curve for coffee to shift to the right from D^1 to D^2. At the original equilibrium, e_1, price, the market has excess demand of 1.5 million tons per year. Market pressures drive up the price to \$3. The new equilibrium is e_2.



Well, there goes the price of lobsters!

coffee. Because the price of cocoa is not included in the demand function, the demand curve, D^1, does not shift.

At the original equilibrium price of coffee, \$2, consumers still want to buy 10 million tons, but suppliers are now willing to supply only 9.4 million tons at that price, so the market has an excess demand of $10 - 9.4 = 0.6$. Market pressure forces the coffee price upward until it reaches the new equilibrium, e_2.

At e_2, the new equilibrium price is \$2.40, and the new equilibrium quantity is 9.6 million tons. Thus, the increase in the price of cocoa causes the equilibrium price of coffee to rise by 40¢ per lb, and the equilibrium quantity to fall by 0.4 million tons. In this case, the increase in the price of cocoa causes a *shift of the supply curve* and a *movement along the demand curve*.

Solved Problem 2.3

Using algebra, determine how the equilibrium price and quantity of coffee change from the initial levels, $p = \$2$ and $Q = 10$, if the price of cocoa increases from its original price of $p_c = \$3$ by \$3 to \$6 per lb.

Answer

1. *Show how the demand and supply functions change due to the increase in the price of cocoa*. Because the demand function does not depend on p_c, it remains unchanged: $Q = 12 - p$ (Equation 2.3). Substituting the new $p_c = \$6$ into the supply function, Equation 2.6, $Q = 9.6 + 0.5p - 0.2p_c$, we find that the new supply function is $Q = 9.6 + 0.5p - (0.2 \times 6) = 8.4 + 0.5p$.

2. *Equate the supply and demand functions to determine the new equilibrium*. The equilibrium price is determined by equating the right sides of these demand and supply functions:

$$12 - p = 8.4 + 0.5p.$$

Solving this equation for p, we find that the equilibrium price is $p = \$2.40$. We calculate the equilibrium quantity by substituting this price into the demand or supply functions: $Q = 12 - 2.40 = 8.4 + (0.5 \times 2.40) = 9.6$ million tons per year.

3. *Show how the equilibrium price and quantity of coffee changes by subtracting the original price and quantity from the new ones*. The change in the equilibrium price is $\Delta p = \$2.40 - \$2 = \$0.40$. The change in the equilibrium quantity is $\Delta Q = 9.6 - 10 = -0.4$ million tons per year. Figure 2.7 illustrates these changes.

Effects of a Shock to the Demand Curve

Now suppose that the price of cocoa stays constant at its original level but annual household income rises from \$35,000 to \$50,000. This change causes the coffee demand curve to shift to the right but does not affect the supply curve, as panel b of Figure 2.7 shows.

At the original equilibrium coffee price, the market has excess demand of 1.5 million tons per year. The excess demand causes upward pressure on the price. In the new equilibrium, e_2, the equilibrium price is \$3 and the equilibrium quantity is 10.5 million tons per year. Here a *shift of the demand curve* results in a *movement along the supply curve*.

In summary, a change in an underlying factor shifts the demand curve or the supply curve, causing the equilibrium to change. To describe the effect of this change, we compare the original equilibrium price and quantity to the new equilibrium values.

2.5 Effects of Government Interventions

A government can affect a market equilibrium in many ways. Sometimes government actions shift the supply curve, the demand curve, or both curves, affecting the equilibrium. However, other government interventions can cause the quantity demanded to differ from the quantity supplied.

Policies That Shift Supply Curves

We concentrate on government policies that affect the supply curve because they are more common than policies that affect the demand curve. We discuss two government supply policies: licensing laws and quotas.

Licensing Laws A government *licensing law* limits the number of firms that may sell goods in a market. For example, many local governments around the world limit the number of taxicabs (see Chapter 9). Governments use zoning laws to limit the number of bars, bookstores, hotel chains, as well as firms in many other markets. In developed countries, early entrants or those people who pass an exam receive licenses. In some developing countries, licenses go to relatives of government officials or to whomever offers those officials the largest bribe.

Application

Occupational Licensing

In the United States, in many occupations, working without a license is illegal. Local, state, or federal governments license more than 800 occupations, including animal masseuse, animal trainers, dietitians and nutritionists, doctors, electricians, embalmers, funeral directors, hairdressers, librarians, nurses, psychologists, real estate brokers, respiratory therapists, salespeople, teachers, tree trimmers, and truck drivers (but not economists).

During the early 1950s, fewer than 5% of U.S. workers were in occupations covered by licensing laws at the state level. Since then, the share of licensed employed workers has grown, reaching nearly 18% by the 1980s, at least 20% in 2000, and 26% in 2015. Licensing is more common in occupations that require extensive education: More than 40% of workers with post-college education are required to have a license compared to only 15% of those with less than a high school education. More than three-fourths of workers in healthcare and technical occupations have licenses.

A worker must pass a test to get a license in some occupations. Frequently, licensed members of the occupation design these tests. By making the exam difficult, current workers can limit entry. For example, only 47% of people taking the California State Bar Examination in July 2015 passed it, although all of them had law degrees. (The national rate for lawyers passing state bar exams in July 2015 was higher, but still only 63%.) To braid hair professionally in South Dakota requires 2,100 hours of education and a cosmetology license; whereas, South Carolina mandates only a six-hour course.

To the degree that testing is objective, licensing may raise the average quality of the workforce. However, its primary effect is to restrict the number of workers in an occupation. To analyze the effects of licensing, one can use a graph similar to panel b of Figure 2.7, where the wage is on the vertical axis and the number of workers per year is on the horizontal axis.

Licensing shifts the occupational supply curve to the left, reducing the equilibrium number of workers and raising the wage. Kleiner and Krueger (2013) found that licensing raises occupational wages by 18%. According to the U.S. Labor Department in 2016, the median worker with a license earned just over $1,000 a week compared to a little under $800 for those without. A license is worth $108 more per week for a worker with a high-school degree, but did not raise earnings for those with bachelor's or advanced degrees. Kleiner (2015) claims that occupational licensing costs consumers $203 billion annually.

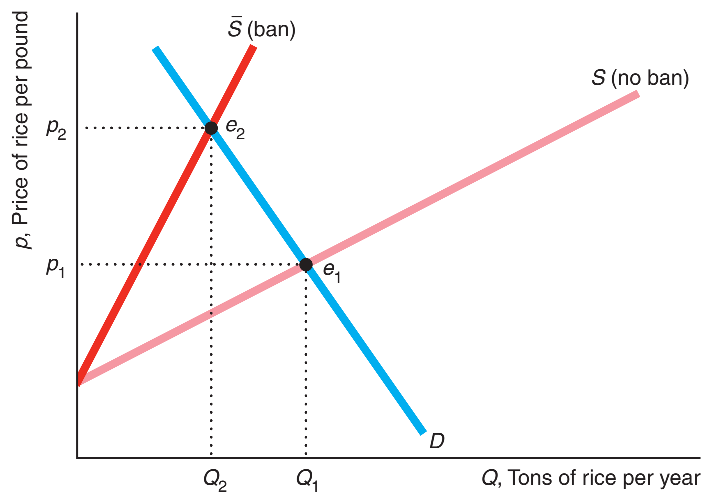
In 2016, the Obama White House announced that it would provide $7.5 million in grants to organizations that would work with states to reduce overly burdensome licensing and make it easier for licensed practitioners to work across state lines.

Quotas Quotas typically limit the amount of a good that firms may sell (rather than the number of firms that sell it). Governments commonly use quotas to limit imports. As we saw earlier, quotas on imports affect the supply curve.

We illustrate the effect of quotas on market equilibrium using our previous example of the ban on rice imports in Japan. The Japanese government's ban (the quota was set to zero) on rice imports raised the price of rice in Japan substantially. Figure 2.8 shows the Japanese demand curve for rice, *D*, and the total supply curve

Figure 2.8 A Ban on Rice Imports Raises the Price in Japan

A ban on rice imports shifts the total supply curve of rice in Japan without a ban, *S*, to $\bar{S}$, which equals the domestic supply alone. As a result, the equilibrium changes from e_1 to e_2. The ban causes the price to rise from p_1 to p_2 and the equilibrium quantity to fall to Q_1 from Q_2.



without a ban, S. The intersection of S and D determines the equilibrium, e_1, if the Japanese government permits rice imports.

What is the effect of a ban on foreign rice on Japanese supply and demand? The ban does not affect the demand curve if Japanese consumers do not care whether they eat domestic or foreign rice. The ban causes the total supply curve to rotate toward the origin from S (total supply is the horizontal sum of domestic and foreign supply) to $\bar{S}$ (total supply equals the domestic supply).

The intersection of $\bar{S}$ and D determines the new equilibrium, e_2, which lies above and to the left of e_1. The ban causes a shift of the supply curve and a movement along the demand curve. It causes the equilibrium quantity to fall from Q_1 to Q_2 and the equilibrium price to increase from p_1 to p_2. Because of a near total ban on imported rice in 2001, the price of rice in Japan was 10.5 times higher than the price in the rest of the world. This markup has decreased in recent years.

A quota of $\bar{Q}$ may have a similar effect to an outright ban; however, a quota may have no effect on the equilibrium if the quota is set so high that it does not limit imports. We investigate this possibility in Solved Problem 2.4.

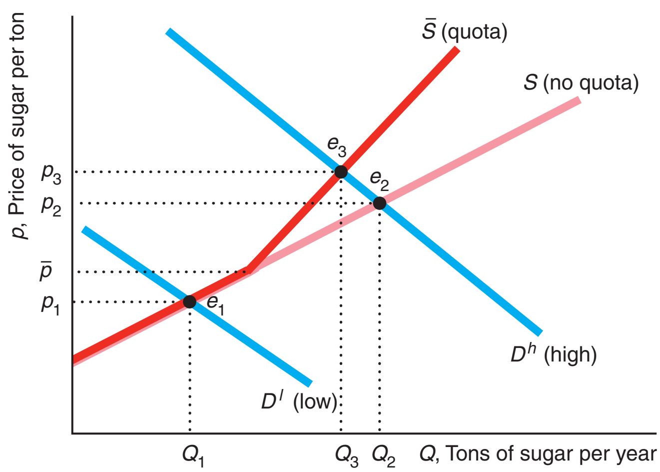
Solved Problem 2.4

MyLab Economics Solved Problem

What is the effect of a United States quota on sugar of $\bar{Q}$ on the equilibrium in the U.S. sugar market? *Hint*: The answer depends on whether the quota *binds* (is low enough to affect the equilibrium).

Answer

1. *Show how a quota, $\bar{Q}$, affects the total supply of sugar in the United States*. The graph reproduces the no-quota total American supply curve of sugar, S, and the total supply curve under the quota, $\bar{S}$ (which we derived in Solved Problem 2.2). At a price below $\bar{p}$, the two supply curves are identical because the quota is not binding: It is greater than the quantity foreign firms want to supply. Above $\bar{p}$, $\bar{S}$ lies to the left of S.



2. *Show the effect of the quota if the original equilibrium quantity is less than the quota so that the quota does not bind*. Suppose that the American demand is relatively *low* at any given price so that the demand curve, D^l, intersects both the supply curves at a price below $\bar{p}$. The equilibria both before and after the quota is imposed are at e_1, where the equilibrium price, p_1, is less than $\bar{p}$. Thus if the demand curve lies near enough to the origin that the quota is not binding, the quota has no effect on the equilibrium.

3. *Show the effect of the quota if the quota binds*. With a relatively *high* demand curve, D^h, the quota affects the equilibrium. The no-quota equilibrium is e_2, where D^h intersects the no-quota total supply curve, S. After the quota is imposed, the equilibrium is e_3, where D^h intersects the total supply curve with the quota, $\bar{S}$. The quota raises the price of sugar in the United States from p_2 to p_3 and reduces the quantity from Q_2 to Q_3.

Comment: Currently, 76% of the sugar Americans consume is produced domestically, while the rest is imported from about 40 countries under a quota system. In a typical year, the U.S. price of sugar is 24% higher than the price in the rest of the world. (This increase in price is applauded by nutritionists who deplore the amount of sugar consumed in the typical U.S. diet.)

Policies That Cause the Quantity Demanded to Differ from the Quantity Supplied

Some government policies do more than merely shift the supply or demand curve. For example, governments may control prices directly, a policy that leads to either excess supply or excess demand if the price the government sets differs from the equilibrium price. We illustrate this result with two types of price control programs: price ceilings and price floors. When the government sets a *price ceiling* at $\bar{p}$, the price at which goods are sold may be no higher than $\bar{p}$. When the government sets a *price floor* at $\underline{p}$, the price at which goods are sold may not fall below $\underline{p}$.

Price Ceilings Price ceilings have no effect if they are set above the equilibrium price that would be observed in the absence of the price controls. If the government says that firms may charge no more than $\bar{p} = \$5$ per gallon of gas and firms are actually charging $p = \$1$, the government's price control policy is irrelevant. However, if the equilibrium price, p, was above the price ceiling $\bar{p}$, the price actually observed in the market would be the price ceiling.

The United States used price controls during both world wars, the Korean War, and in 1971–1973 during the Nixon administration, among other times. The U.S. experience with gasoline illustrates the effects of price controls. In the 1970s, the Organization of Petroleum Exporting Countries (OPEC) reduced supplies of oil (which is converted into gasoline) to Western countries. As a result, the total supply curve for gasoline in the United States—the horizontal sum of domestic and OPEC supply curves—shifted to the left, from S^1 to S^2, in Figure 2.9. Because of this shift, the equilibrium price of gasoline would have risen substantially, from p_1 to p_2. In an attempt to protect consumers by keeping gasoline prices from rising, the U.S. government set price ceilings on gasoline in 1973 and 1979.

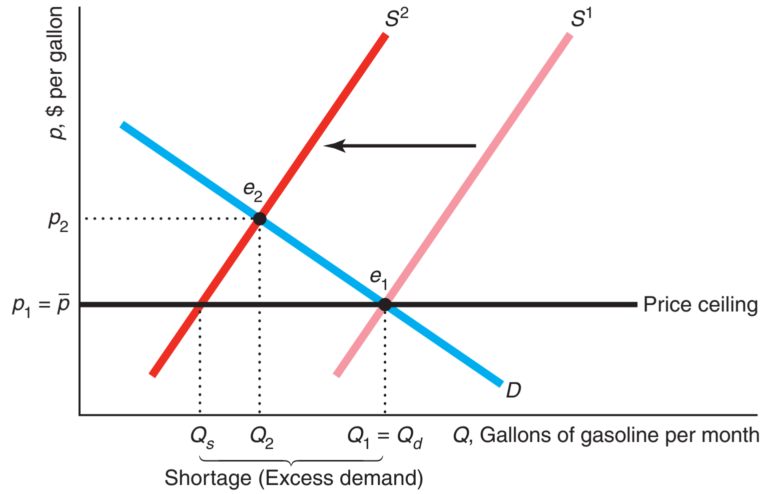
The government told gas stations that they could charge no more than $\bar{p} = p_1$. Figure 2.9 shows the price ceiling as a solid horizontal line extending from the price axis at $\bar{p}$. The price control is binding because $p_2 > \bar{p}$. The observed price is the price ceiling. At $\bar{p}$, consumers *want* to buy $Q_d = Q_1$ gallons of gasoline, which is the equilibrium quantity they bought before OPEC acted. However, firms supply only Q_s gallons at the intersection of the price control line and S^2. As a result of the binding price control, the excess demand is $Q_d - Q_s$.

Were it not for the price controls, market forces would drive up the market price to p_2 and eliminate the excess demand. The government price ceiling prevents this adjustment from occurring. As a result, an enforced price ceiling causes a **shortage**: a persistent excess demand.

shortage
a persistent excess demand

Figure 2.9 Price Ceiling on Gasoline MyLab Economics Video

After a supply shock, the supply curve shifts from S^1 to S^2. Under the government's price control program, gasoline stations may not charge a price above the price ceiling, $\bar{p} = p_1$. At that price, producers are willing to supply only Q_s, which is less than the amount $Q_1 = Q_d$ that consumers want to buy. The result is excessive demand, or a shortage of $Q_d - Q_s$.



At the time of the controls, some government officials argued that the shortages were caused by OPEC cutting off its supply of oil to the United States, but that's not true. Without the price controls, the new equilibrium would be e_2. In this equilibrium, the price, p_2, is much higher than before, p_1; however, no shortage results. Moreover, without controls, the quantity sold, Q_2, is greater than the quantity sold under the control program, Q_s.

With a binding price ceiling, the supply-and-demand model predicts an *equilibrium with a shortage*. In this equilibrium, the quantity demanded does not equal the quantity supplied. The reason that we call this situation an equilibrium, even though a shortage exists, is that no consumers or firms want to act differently, given the law. Without the price controls, consumers facing a shortage would try to get more output by offering to pay more, or firms would raise prices. With effective government price controls, the price cannot rise, so consumers have to live with the shortage.

What happens? Some lucky consumers get to buy Q_s units at the low price of $\bar{p}$. Other potential customers are disappointed: They would like to buy at that price, but they cannot find anyone willing to sell gas to them.

What determines which consumers are lucky enough to find goods to buy at the low price when the government imposes a price control? With enforced price controls, sellers use criteria other than price to allocate the scarce commodity. Firms may supply their friends, long-term customers, or people of a certain race, gender, age, or religion. They may sell their goods on a first-come, first-served basis. Or, they may limit everyone's purchases to only a few gallons.

Another possibility is for firms and customers to evade the price controls. A consumer could go to a gas station owner and say, "Let's not tell anyone, but I'll pay you twice the price the government sets if you'll sell me as much gas as I want." If enough customers and gas station owners behaved that way, no shortage would occur. A study of 92 major U.S. cities during the 1973 gasoline price controls found no gasoline lines in 52 of the cities. However, in cities such as Chicago, Hartford, New York, Portland, and Tucson, potential customers waited in line at the pump for an hour or more.[12] Deacon and Sonstelie (1989) calculated that for every dollar

[12]See MyLab Economics, Chapter 2, "Gas Lines," for a discussion of the effects of the 1973 and 1979 gasoline price controls.

consumers saved during the 1980 gasoline price controls, they lost \$1.16 in waiting time and other factors.

More recently, Hawaii, New York, and New Jersey have imposed gasoline price controls. Hawaii imposed price controls on the wholesale price of gasoline starting in September 2005, but suspended the controls indefinitely in early 2006 due to the public's unhappiness with the law. Following tight supplies of gasoline after Superstorm Sandy hit the East Coast in 2012, both New York and New Jersey enacted price controls for a couple of weeks.

Application

Venezuelan Price Ceilings and Shortages

Venezuela traditionally has been one of the richest countries in Latin America. It is a leading oil producer, and it has many other agricultural and non-agricultural industries.

So why do people start lining up to buy groceries in Venezuela at 4 A.M., when shops open at 8 A.M.? Strict price ceilings on food and other goods create shortages throughout the country.

According to Venezuela's central bank, 28% of products were unavailable in shops in 2014, an all-time high. In 2015, Venezuelans were particularly vexed by condom, birth control pill, and toilet paper shortages. By 2016, the country's economy was melting down. Eighty-seven percent of Venezuelans said that they did not have enough money to buy food. Consumers could not buy Coca-Cola because sugar was not available. Empresas Polar SA, which makes 80% of the beer that Venezuelans consume, ceased operation. Supermarket shelves were often empty, firms transported food under armed guard, and soldiers stood watch over bakeries. Mobs stormed grocery stores, pharmacies, and butcher shops.

One would think that Venezuela should be able to supply its citizens with coffee, which it has produced in abundance for centuries. Indeed, Venezuela exported coffee until 2009. However, since then, it has been importing large amounts of coffee to compensate for a drop in production. Why have farmers and coffee roasters cut production? Due to low retail price ceilings, they would have to produce at a loss.

Because Venezuela regulates the prices of many goods such as gasoline and corn flour and Colombia, its direct neighbor to the west, does not, smuggling occurs. Given that gasoline sold in 2015 for 4¢ a gallon in Venezuela, and the price was 72¢ a gallon in most of Colombia, the temptation to smuggle is great. Venezuela's Táchira state is adjacent to the Colombian border. Its government says that as much as 40% of the food sent to Táchira is smuggled into Colombia. Why sell corn flour at an artificially low price in Venezuela if you can sell it at a higher, market price in Colombia?[13]

Venezuela's populist President Hugo Chávez and his hand-picked successor, Nicolás Maduro, imposed strict price

[13]According to news reports in late 2016, seven border-state governments have stopped enforcing price controls on some basic goods such as food. They are allowing imported food to sell at market prices, which are often 5 to 20 times the regulated prices.

ceilings purportedly to rein in inflation and make the goods more affordable for the poor. Do the ceilings help the poor?

For many Venezuelans, the answer is "No!" As Nery Reyes, a restaurant worker, said, "Venezuela is too rich a country to have this. I'm wasting my day here standing in line to buy one chicken and some rice."[14]

For several years, demonstrators have taken to the streets to protest persistent economic and social problems, including shortages. In 2016, demonstrations average 17 per day across the country. Many people died in violent clashes with the National Guard, including a student beauty queen who had been crowned Miss Tourism for the state of Carabobo. The ultimate irony is that President Nicolás Maduro has advised Venezuelans to consume less to alleviate the shortages.

Price Floors Governments also commonly use price floors. One of the most important examples of a price floor is a minimum wage in a labor market. A minimum wage law forbids employers from paying less than the minimum wage, $\underline{w}$.

Minimum wage laws date from 1894 in New Zealand, 1909 in the United Kingdom, and 1912 in Massachusetts. The Fair Labor Standards Act of 1938 set a federal U.S. minimum wage of 25¢ per hour. Today, the federal minimum wage is \$7.25 an hour, but 31 states, the District of Columbia, and many cities also set a higher minimum wage. For example, Washington State's minimum wage is \$9.47, and Oakland, California's rate is \$12.25. In 2016, the United Kingdom's minimum hourly wage was £6.70 for adult workers. The 2015 statutory monthly minimum wage ranged from the equivalent of 107€ in the Russian Federation to 505€ in Portugal, 1,457.52€ in France, and 1,922.96€ in Luxembourg.[15]

Solved Problem 2.5

Suppose everyone receives the same wage in a labor market. What happens to the equilibrium in this market if the government imposes a binding minimum wage, $\underline{w}$?

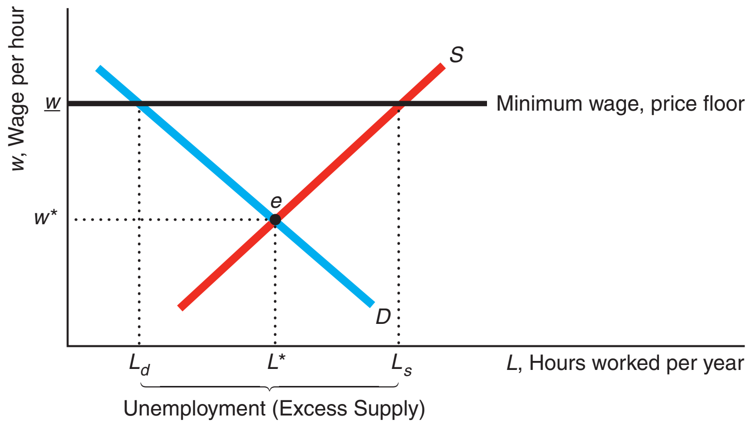
Answer

1. *Show the initial equilibrium before the government imposes the minimum wage.* The figure shows the supply and demand curves for labor services (hours worked). Firms buy hours of labor service—they hire workers. The quantity measure on the horizontal axis is hours worked per year, and the price measure on the vertical axis is the wage per hour. With no government intervention, the intersection of the supply and demand curves determine the market equilibrium at e, where the wage is $\underline{w}^*$ and the number of hours worked is L^*.

2. *Draw a horizontal line at the minimum wage, and show how the market equilibrium changes.* The minimum wage creates a price floor, a horizontal line, at $\underline{w}$. At that wage, the quantity demanded falls to L_d and the quantity supplied rises to L_s. As a result, excess supply or unemployment is $L_s - L_d$. The minimum

[14]William Neuman, "With Venezuelan Food Shortages, Some Blame Price Controls," *New York Times*, April 20, 2012.

[15]The U.S. Department of Labor maintains at its website, **www.dol.gov**, an extensive history of the federal minimum wage law, labor markets, state minimum wage laws, and other information. For European minimum wages, see **www.fedee.com/minwage.html**. See **www.direct.gov.uk** for British rates.



wage prevents market forces from eliminating this excess supply, so it leads to an equilibrium with unemployment.[16]

Comment: The original 1938 U.S. minimum wage law caused massive unemployment in Puerto Rico.[17] Depending on the law's coverage and the type of labor market, the minimum wage may not cause unemployment (see Chapters 10 and 15). Card and Krueger (1995) argued, based on alternatives to the simple supply-and-demand model, that minimum wage laws raise wages in some markets (such as fast foods) without significantly reducing employment. In contrast, Neumark et al. (2014) concluded, based on an extensive review of minimum wage research, that increases in the minimum wage often have negative effects on employment.

Why the Quantity Supplied Need Not Equal the Quantity Demanded

The price ceiling and price floor examples show that the quantity supplied does not necessarily equal the quantity demanded. Many people fail to understand this key point when they blindly parrot:

Common Confusion: Demand must equal supply.

Their claim is true only, in an uninteresting accounting sense, that the amount firms *actually* sell must equal the amount that consumers *actually* purchase. Because we define the quantities supplied and demanded in terms of people's *wants* and not *actual* quantities bought and sold, the statement that "supply equals demand" has content: It is shorthand for a theory, not merely an accounting equality. The supply-and-demand theory says that the intersection of the supply curve and the demand curve determines the equilibrium price and quantity in a market if the government does not intervene.

When the government sets a price ceiling or a price floor, the quantity supplied does not equal the quantity demanded because of the way we define these two concepts. We

[16]The minimum wage could raise the wage enough that total wage payments, wL, rise despite the fall in demand for labor services. If the workers could share the unemployment—everyone works fewer hours than he or she wants—all workers could benefit from the minimum wage.

[17]See MyLab Economics, Chapter 2, "Minimum Wage Law in Puerto Rico."

define the quantity supplied as the amount firms *want to sell* at a given price, holding other factors that affect supply, such as the price of inputs, constant. The quantity demanded is the quantity that consumers *want to buy* at a given price, holding constant other factors that affect demand. When the government regulates price in a market, the quantity that firms want to sell and the quantity that consumers want to buy at a given price do not equal the *actual* quantity that firms sell and consumers buy.

For example, when the government imposes a binding price ceiling on gasoline, the quantity demanded is greater than the quantity supplied. Despite the lack of equality between the quantity supplied and the quantity demanded, the supply-and-demand model is useful in analyzing this market because it predicts the excess demand.

2.6 When to Use the Supply-and-Demand Model

As we've seen, supply-and-demand theory can help us to understand and predict real-world events in many markets. Through Chapter 10, we discuss competitive markets in which the supply-and-demand model is a powerful tool for predicting what will happen to market equilibrium if underlying conditions—tastes, incomes, and prices of inputs—change. The types of markets for which the supply-and-demand model is useful are described at length in these chapters, particularly in Chapters 8 and 9. Briefly, this model is applicable to markets in which:

- **Everyone is a price taker.** Because no consumer or firm is a very large part of the market, no one can affect the market price. Easy entry of firms into the market, which leads to a large number of firms, is usually necessary to ensure that firms are price takers.
- **Firms sell identical products.** Consumers do not prefer one firm's good to another.
- **Everyone has full information about the price and quality of goods.** Consumers know if a firm is charging a price higher than the price others set, and they know if a firm tries to sell them inferior-quality goods.
- **Costs of trading are low.** It is not time-consuming, difficult, or expensive for a buyer to find a seller and make a trade or for a seller to find and trade with a buyer.

Economists call markets with these properties *perfectly competitive markets*.

In a market with many firms and consumers, no single firm or consumer is a large enough part of the market to affect the price. If you stop buying bread or if one of the many thousands of wheat farmers stops selling the wheat used to make the bread, the price of bread will not change. Consumers and firms are *price takers*: They cannot affect the market price.

In contrast, if a market has only one seller of a good or service—a *monopoly* (see Chapter 11)—that seller is a *price setter* and can affect the market price. Because demand curves slope downward, a monopoly can increase the price it receives by reducing the amount of a good it supplies. Firms are also price setters in an *oligopoly*—a market with only a small number of firms—or in markets where they sell differentiated products so that a consumer prefers one product to another (see Chapter 13). In markets with price setters, the market price is usually higher than that predicted by the supply-and-demand model. That doesn't make the model generally wrong. It means only that the supply-and-demand model does not apply to markets with a small number of sellers or buyers. In such markets, we use other models.

If consumers have less information than a firm, the firm can take advantage of consumers by selling them inferior-quality goods or by charging a much higher price than that charged by other firms. In such a market, the observed price is usually higher than that predicted by the supply-and-demand model, the market may not exist at all (consumers and firms cannot reach agreements), or different firms may charge different prices for the same good (see Chapter 19).

The supply-and-demand model is also not entirely appropriate for markets in which it is costly to trade with others because the cost of a buyer finding a seller or of a seller finding a buyer is high. **Transaction costs** are the expenses of finding a trading partner and making a trade for a good or service other than the price paid for that good or service. These costs include the time and money spent to find someone with whom to trade. For example, you may have to pay to place an ad to sell your gray 1999 Honda with 137,000 miles on it. Or, you may have to go to many stores to find one that sells a shirt that fits exactly the way you like, so your transaction costs include transportation costs and your time. The labor cost of filling out a form to place an order is a transaction cost. Other transaction costs include the costs of writing and enforcing a contract, such as the cost of a lawyer's time. Where transaction costs are high, no trades may occur, or if they do occur, individual trades may occur at a variety of prices (see Chapters 12 and 19).

transaction costs
the expenses of finding a trading partner and making a trade for a good or service beyond the price paid for that good or service

Thus, the supply-and-demand model is not appropriate in markets with only one or a few firms (such as electricity), differentiated products (movies), consumers who know less than sellers about quality or price (used cars), or high transaction costs (nuclear turbine engines). Markets in which the supply-and-demand model has proved useful include agriculture, finance, labor, construction, services, wholesale, and retail.

Challenge Solution

Quantities and Prices of Genetically Modified Foods

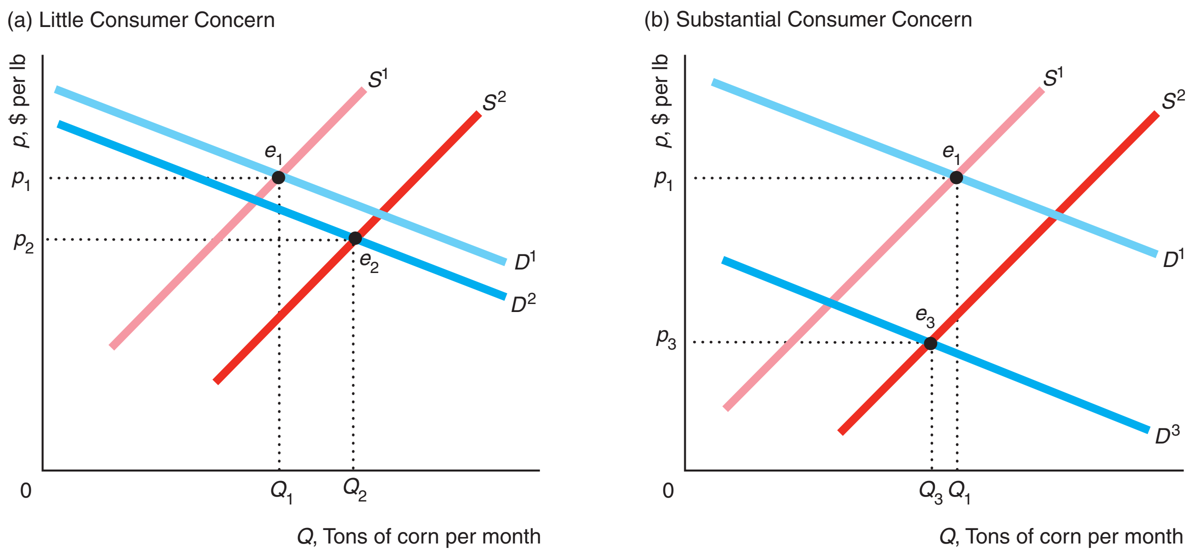
MyLab Economics Solved Problem

We conclude this chapter by returning to the challenge posed at its beginning where we asked about the effects on the price and quantity of a crop, such as corn, from the introduction of GM seeds. The supply curve shifts to the right because GM seeds produce more output than traditional seeds, holding all else constant. If consumers fear GM products, the demand curve for corn shifts to the left. We want to determine how the after-GM equilibrium compares to the before-GM equilibrium. When an event shifts both curves, the qualitative effect on the equilibrium price and quantity may be difficult to predict, even if we know the direction in which each curve shifts. Changes in the equilibrium price and quantity depend on exactly how much the curves shift. In our analysis, we want to take account of the possibility that the demand curve may shift only slightly in some countries where consumers don't mind GM products (or product labels don't show they contain GM products) but substantially in others where many consumers fear GM products.

In the figure, the intersection of the before-GM supply curve, S^1, and the before-GM demand curve, D^1, determines the before-GM equilibrium, e_1, at price p_1 and quantity Q_1. Both panels a and b of the figure show this same equilibrium. After GM seeds are introduced, the supply curve, S^2, shifts to the right of the original supply curve, S^1 in both panels.

Panel a shows the situation if consumers have little concern about GM crops, so that the new demand curve, D^2, lies only slightly to the left of the original demand curve, D^1. In panel b, where consumers are greatly concerned about GM crops, the new demand curve, D^3, lies substantially to the left of D^1.

In panel a, the new equilibrium e_2 is determined by the intersection of S^2 and D^2. In panel b, the new equilibrium e_3 reflects the intersection of S^2 and D^3. The equilibrium price falls from p_1 to p_2 in panel a and to p_3 in panel b. However, the equilibrium quantity rises from Q_1 to Q_2 in panel a, but falls to Q_3 in panel b. That is, the price falls in both cases, but the quantity may rise or fall depending on how much the demand curve shifts. Thus, whether growers in a country decide to adopt GM seeds depends crucially on consumer resistance to these new products.



Summary

1. **Demand.** The quantity of a good or service demanded by consumers depends on their tastes, the price of the good, the price of substitutes and complements, consumers' income, information, government regulations, and other factors. The *Law of Demand*—which is based on observation—says that *demand curves slope downward*. Consumers demand less of the good, the higher the price, holding constant other factors that affect their demand. A change in price causes a *movement along the demand curve*. A change in income, tastes, or another factor that affects demand, other than price, causes a *shift of the demand curve*. To get a total demand curve, we horizontally sum the demand curves of individuals or types of consumers or countries. That is, we add the quantities demanded by each individual at a given price to get the total demand.

2. **Supply.** The quantity of a good or service supplied by firms depends on its price, cost, government regulation, and other factors. The market supply curve need not slope upward but usually does. A change in price causes a *movement along the supply curve*. A change in a government regulation or the price of an input causes a *shift of the supply curve*. The total supply curve is the horizontal sum of the supply curves for individual firms.

3. **Market Equilibrium.** The intersection of the demand curve and the supply curve determines the equilibrium price and quantity in a market. Market forces—actions of consumers and firms—drive the price and quantity to the equilibrium levels if they are initially too low or too high.

4. **Shocking the Equilibrium.** A change in an underlying factor other than price causes a shift of the supply curve or the demand curve, which alters the equilibrium. For example, if consumer income rises, the demand curve for coffee shifts to the right, causing a movement along the supply curve and leading to a new equilibrium at a higher price and quantity. If changes in these underlying factors follow one after the other, a market that adjusts slowly may stay out of equilibrium for an extended period.

5. **Effects of Government Interventions.** Some government policies—such as a ban on imports—cause a shift in the supply or demand curves, which alters the equilibrium. Other government policies—such as price controls or a minimum wage—cause the quantity supplied to be greater or less than the quantity demanded, leading to persistent excesses or shortages.

6. **When to Use the Supply-and-Demand Model.** The supply-and-demand model is a powerful tool

to explain what happens in a market or to make predictions about what will happen if an underlying factor in a market changes. This model, however, is applicable only to markets with many buyers and sellers; identical goods; certainty and full information about price, quantity, quality, incomes, costs, and other market characteristics; and low transaction costs.

Questions

Select questions are available on MyLab Economics*; * = answer appears at the back of this book;* **A** *= algebra problem.*

1. Demand

*1.1 Suppose that the demand function for lamb in Australia is $Q = 63 - 11p + 7p_b + 3p_c + 2Y$, where Q is the quantity in million kilograms (kg) of lamb per year, p is the dollar price per kg (all prices cited are in Australian dollars), p_b is the price of beef per kg, p_c is the price of chicken per kg, and Y is annual per capita income in thousands of Australian dollars. What is the demand curve if we hold p_b, p_c, and Y at their typical values during the period studied: $p_b = 19$, $p_c = 6$, and $Y = 78$? **A**

*1.2 Using the demand function for lamb from Question 1.1, show how the quantity demanded at a given price changes as annual per capita income, Y, increases by AU$200. (*Hint*: See Solved Problem 2.1.) **A**

1.3 Based on the Application "Calorie Counting," show the effect of New York City's calorie posting requirement on Starbucks' demand curve.

1.4 Given an estimated monthly demand function for avocados of $Q = 104 - 40p + 20p_t + 0.01Y$, show how the demand curve shifts as per capita income, Y, increases from $4,000 to $5,000 per month. (*Note*: The price of tomatoes, p_t, is $0.80.) Illustrate this shift in a diagram. **A**

1.5 Given the inverse demand function for lamb (Question 1.1) is $p = 33.64 - 0.09Q$, how much would the price have to rise for consumers to want to buy 2 million kg of lamb less per year? (*Hint*: See Solved Problem 2.1.) **A**

1.6 The food and feed demand curves used in the Application "Aggregating Corn Demand Curves" were estimated by McPhail and Babcock (2012) to be $Q_{food} = 1{,}487 - 22.1p$ and $Q_{feed} = 6{,}247.5 - 226.7p$, respectively. Mathematically derive the total demand curve, which the Application's figure illustrates. (*Hint*: Remember that the demand curve for feed is zero at prices above $27.56.) **A**

*1.7 Suppose that the inverse demand function for movies is $p = 120 - Q_1$ for college students and $p = 120 - 2Q_2$ for other town residents. What is the town's total demand function ($Q = Q_1 + Q_2$ as a function of p)? Use a diagram to illustrate your answer. (*Hint*: See the Application "Aggregating Corn Demand Curves.") **A**

1.8 Duffy-Deno (2003) estimated the demand functions for broadband service are $Q_s = 15.6p^{-0.563}$ for small firms and $Q_l = 16.0p^{-0.296}$ for larger ones, where price is in cents per kilobyte per second and quantity is in millions of kilobytes per second (Kbps). What is the total demand function for all firms? (*Hint*: See the Application "Aggregating Corn Demand Curves.") **A**

1.9 Based on the estimates of Ghose and Han (2014), the demand function for mobile applications at Apple's App Store is $Q_A = 1.4p^{-2}$ and the demand function at Google Play is $1.4p^{-3.7}$, where the quantity is in millions of apps. What is the total demand function for apps? If the price for an app is $1, what is the equilibrium quantity demanded by Apple customers, Google customers, and all customers? (*Hint*: Look at the Application "Aggregating Corn Demand Curves.") **A**

2. Supply

*2.1 Suppose that the supply function for lamb in Australia is $Q = 149 + 8p - 9p_s$, where Q is the quantity in millions of kg of lamb per year, and p and p_s are the prices of lamb and sheep, respectively, in Australian dollars per kg. How does the supply curve change if the price of sheep increases from AU$5 to AU$5.50 per kg? (*Hint*: See Solved Problem 2.1.) **A**

2.2 The estimated supply function for avocados is $Q = 58 + 15p - 20p_f$, where p_f is the price of fertilizer. Determine how much the supply curve for avocados shifts if the price of fertilizer rises from $0.40 to $1.50 per lb. Illustrate this shift in a diagram.

2.3 If the supply curve for fertilizer in Jordan is $Q_j = -2 + 4p$ and the supply curve for fertilizer in the rest of the world is $Q_{row} = 25 + 8p$, what is the world supply curve? **A**

2.4 How would a quota of $\overline{Q} = 6$ million tons set by Iran on foreign wheat imports affect the total Iranian supply curve given the domestic supply curve $Q_s^d = 5p - 15$ and the foreign supply curve $Q_s^f = 2p$, where the quantities are in millions of tons and the price is in rials per ton? How much wheat would domestic and foreign producers supply at a price of IRR5 per ton, both with and without the quota?

3. Market Equilibrium

*3.1 Use a supply-and-demand diagram to explain the statement "Talk is cheap because supply exceeds demand." At what price is this comparison being made?

3.2 Every house in a small town has a well that provides water at no cost. However, if the town wants more than 10,000 gallons a day, it has to buy the extra water from firms located outside of the town. The town currently consumes 9,000 gallons per day.

a. Draw the linear demand curve.

b. The firms' supply curve is linear and starts at the origin. Draw the market supply curve, which includes the supply from the town's wells.

c. Show the equilibrium. What is the equilibrium quantity? What is the equilibrium price? Explain.

3.3 A large number of firms are capable of producing chocolate-covered cockroaches. The linear, upward-sloping supply curve starts on the price axis at \$6 per box. A few hardy consumers are willing to buy this product (possibly to use as gag gifts). Their linear, downward-sloping demand curve hits the price axis at \$4 per box. Draw the supply and demand curves. Does this market have an equilibrium at a positive price and quantity? Explain your answer.

3.4 The demand function for lamb in Australia (see Question 1.1) is $Q_d = 63 - 11p + 7p_b + 3p_c + 2Y$, and the supply function (see Question 2.1) is $Q_s = 149 + 8p - 9p_s$. Solve for the equilibrium price and quantity in terms of the price of beef, $p_b = 19$, the price of chicken, $p_c = 6$, the price of sheep, $p_s = 5$, and annual per capita income, $Y = 78$. All prices are in Australian dollars per kg, quantities in million kg, and income in thousands of Australian dollars.

*3.5 The demand function for a good is $Q = a - bp$, and the supply function is $Q = c + ep$, where a, b, c, and e are positive constants. Solve for the equilibrium price and quantity in terms of these four constants. **A**

*3.6 Suppose the supply function for processing coffee beans from coffee cherries in Mexico is $Q_s = 3.15 + 0.1p - 0.5p_c$ and the demand curve for coffee beans is $Q_d = 4.1 - 0.2p$, where Q_s and Q_d are quantities of coffee beans in thousands of 60-kg bags, p is the price of coffee beans in millions of pesos per thousand 60-kg bags, and $p_c = 0.8$ is the price of coffee cherries in millions of pesos per thousand 60-kg bags. What is the supply curve for coffee beans (that is, supply as a function of only the price of coffee beans)? Solve for the equilibrium price and quantity of coffee beans. **A**

4. Shocking the Equilibrium

*4.1 Use a figure to explain the fisher's comment about the effect of a large catch on the market price in the cartoon about catching lobsters near the beginning of Section 2.4. What is the supply shock?

4.2 Airbnb is a website that lets apartment dwellers list and rent lodging for a few days to out-of-town visitors. Use a figure to show how this new website affects the equilibrium rental price and quantity of apartments in popular tourist cities.

4.3 Suppose that a combination of public health warnings, school programs, a ban on advertising, and warnings on package labelling cut the demand for cigarettes in half (that is, the demand curve shifted left for cigarettes at any price by 50%). Using a supply-and-demand diagram of a competitive market, show the likely effect on equilibrium price and equilibrium quantity. Would you expect equilibrium quantity to change by 50%? Show how the answer depends on the slope of the supply curve.

4.4 The last major outbreak of mad cow disease in the United Kingdom occurred in the mid 1990s. It is estimated that up to 6 million cows were destroyed at that time in an attempt to control the disease. What was the effect of the disease on the price and quantity of beef sold in the United Kingdom?

4.5 Use a supply-and-demand diagram of a competitive market to illustrate the qualitative effects of the following possible shocks in the Italian market for red wine.

a. A new research study shows that red wine helps prevent heart disease, stave off aging, and reduce age-related memory decline.

b. Trade barriers that restricted imports of red wine from countries outside the European Union are eliminated.

c. A recession in Italy causes a decline in per capita income.

d. Genetically engineered wine grapes are created that allow for greater red wine production without increasing cost.

4.6 Increasingly, instead of advertising in newspapers, individuals and firms use websites that offer free or inexpensive classified ads, such as **Classified Ads.com**, **Craigslist.org**, **Realtor.com**, **Jobs.com**, **Monster.com**, and portals like Google and Yahoo!. Using a supply-and-demand model, explain what will happen to the equilibrium levels of newspaper advertising as the use of the Internet grows. Will the growth of the Internet affect the supply curve, the demand curve, or both? Why?

4.7 Ethanol, a fuel, is made from corn. Ethanol production increased 8.5 times from 2000 to 2015 (**www.ethanolrfa.org**). Show the effect of this increased use of corn for producing ethanol on the price of corn and the consumption of corn as food using a supply-and-demand diagram.

4.8 The demand function is $Q = 220 - 2p$, and the supply function is $Q = 20 + 3p - 20r$, where r is the rental cost of capital. Show how the equilibrium price and quantity vary with r? (*Hint*: See Solved Problem 2.3.) **A**

4.9 Using the information in Question 3.6, determine how the equilibrium price and quantity of coffee beans change if the price of coffee cherries falls by 25%. (*Hint*: See Solved Problem 2.3.) **A**

5. Effects of Government Interventions

5.1 The Application "Occupational Licensing" analyzed the effect of exams in licensed occupations given that their only purpose was to shift the supply curve to the left. How would the analysis change if the exam also raised the average ability of people in that occupation, which affects the demand curve?

*5.2 Is it possible that an outright ban on foreign imports will have no effect on the equilibrium price? (*Hint*: Suppose that imports occur only at relatively high prices.)

5.3 In 2015, the European Commission proposed giving individual member states the right to ban imports of genetically modified fruits and vegetables even if those products are still permitted at the European Union level. If a member state acted to ban such imports, what would be the likely effect of that policy on world prices and quantities? Would such a ban help deter the production of genetically modified products? (*Hint*: See Solved Problem 2.4.)

5.4 What is the effect of an import quota $\overline{Q} > 0$ on equilibrium price and quantity? (*Hint*: Carefully show how the total supply curve changes. See Solved Problem 2.4.)

5.5 There has been a great deal of movement of physicians across borders within the European Union, and this migration has raised concerns about patient safety and quality of care. Some member states have responded by tightening certification requirements for foreign-trained physicians. What effect would such a policy have on the equilibrium quantity and price of doctors' services in the country that introduces it? How would doctors trained in that country and the country's consumers be affected? In answering this question, assume that the supply curves for both domestically trained and foreign-trained physicians initially intersect the wage axis at the same point. (*Hint*: See Solved Problem 2.4.)

5.6 Usury laws place a ceiling on interest rates that lenders such as banks can charge borrowers. Low-income households in states with usury laws have significantly lower levels of consumer credit (loans) than comparable households in states without usury laws. Why? (*Hint*: The interest rate is the price of a loan, and the amount of the loan is the quantity measure.)

5.7 The Thai government actively intervenes in markets (Nophakhun Limsamarnphun, "Govt Imposes Price Controls in Response to Complaints," *The Nation*, May 12, 2012).

a. The government increased the daily minimum wage by 40% to Bt300 (300 bahts $\approx$ \$9.63). Show the effect of a higher minimum wage on the number of workers demanded, the supply of workers, and unemployment if the law is applied to the entire labor market.

b. Show how the increase in the minimum wage and higher rental fees at major shopping malls and retail outlets affected the supply curve of ready-to-eat meals. Explain why the equilibrium price of a meal rose to Bt40 from Bt30.

c. In response to complaints from citizens about higher prices of meals, the government imposed price controls on ten popular meals. Show the effect of these price controls in the market for meals.

d. What is the likely effect of the price controls on meals on the labor market?

5.8 Rent control laws are in place in states such as Maharashtra and Karnataka in India. Such price controls establish the price of rental accommodations and limit annual increases in rent up to a maximum amount. Use a supply-and-demand diagram of a competitive market to show the effect of binding rent controls on the equilibrium price and quantity of rental properties.

*5.9 The tsunamis that hit Japan in 2011 and India and Sri Lanka in 2004 were devastating, and their effects were felt for many years afterward. Natural disasters of this type as well as international events often result in severe disruptions to the supply of goods and services and in sharp spikes in product prices. Governments may respond to public outcries against dramatically higher prices by imposing price ceilings to lower product prices or to keep them from rising too high. What effect would such a binding price ceiling have? Who would benefit from this policy, and who would be harmed?

5.10 Suppose the Mexican government decides to control the price of coffee beans to support coffee bean producers in that country. It does so by imposing a binding price floor of 7.5 million pesos per thousand 60-kg bags and by committing to buy any resulting excess supply. Use the information in Question 3.6 to determine the quantities of coffee beans demanded, supplied, and purchased by the government. (*Hint*: See Solved Problem 2.5.) **A**

5.11 Based on the Application "Venezuela Price Ceilings and Shortages," use two figures to show the effects of Venezuela's price control on corn flour in Venezuela and in Colombia's corn flour markets.

6. When to Use the Supply-and-Demand Model

6.1 Are predictions using the supply-and-demand model likely to be reliable in each of the following markets? Why or why not?

a. Apples.

b. Convenience stores.

c. Electronic games (a market with three major firms).

d. Used cars.

7. Challenge

*7.1 Brazil is one of the world's largest exporters of beef and China is a major purchaser of that beef (an estimated 30% of China's beef imports in 2016 came from Brazil). However, in March 2017, China, South Korea, the European Union, and Chile suspended imports of meat products from Brazil as a precautionary measure in response to allegations that meat inspectors and politicians had received bribes to overlook improper meat packing practices and allow sales of tainted food. How would the closing of export markets for a country's beef products together with a fall in domestic sales of beef products and an increase in the domestic equilibrium quantity be reflected in supply-and-demand diagrams of that country's foreign and domestic markets for beef in the short run?

7.2 Question 7.1 asks you to illustrate why an import ban might cause the equilibrium quantity of beef in the domestic market to rise. Depending on how the supply and demand curves shift, could the equilibrium quantity of beef in the domestic market have fallen instead? Would it ever be possible for the equilibrium price in the domestic market to rise?

7.3 When he was the top American administrator in Iraq, L. Paul Bremer III set a rule that upheld Iraqi law: anyone 25 years and older with a "good reputation and character" could own one firearm, including an AK-47 assault rifle. Iraqi citizens quickly began arming themselves. After the bombing of a sacred Shiite shrine in Samarra at the end of February 2006 and the subsequent rise in sectarian violence, the demand for guns increased, resulting in higher prices. The average price of a legal, Russian-made Kalashnikov AK-47 assault rifle jumped from $112 to $290 from February to March 2006, and the price of bullets shot up from 24¢ to 33¢ each (Jeffrey Gettleman, "Sectarian Suspicion in Baghdad Fuels a Seller's Market for Guns," *New York Times*, April 3, 2006). This increase occurred despite the hundreds of thousands of firearms and millions of rounds of ammunition that American troops had been providing to Iraqi security forces, some of which eventually ended up in the hands of private citizens. Use a graph to illustrate why prices rose. Did the price need to rise, or was the rise related to the shapes of and relative shifts in the demand and supply curves?

Applying the Supply-and-Demand Model

3

New Jersey's decision to eliminate the tax on Botox has users elated. At least I think they're elated. I can't really tell.

Challenge

Who Pays the Gasoline Tax?

U.S. consumers and politicians debate endlessly about whether to increase or decrease gasoline taxes, even though U.S. gasoline taxes are very low relative to those in most other industrialized nations. In 2016, the typical American paid a tax of 45¢ per gallon of unleaded gasoline, which includes the federal tax of 18.4¢ and the average state gasoline tax of 26.6¢ per gallon. The comparable tax is $1.12 in Canada, $3.63 in France, and $4.20 in the United Kingdom.

Regularly, at international climate meetings—such as the one in Paris in 2015—government officials, environmentalists, and economists from around the world argue strongly for an increase in the tax on gasoline and other fuels (or, equivalently, on carbon) to retard global warming and improve the air we breathe.

However, whenever gas prices rise suddenly, other politicians call for removing gasoline taxes, at least temporarily. Illinois and Indiana suspended their taxes during an oil price spike in 2000, as did Alaska in 2008. While running for president, Senators John McCain and Hillary Clinton called for a gas tax holiday during the summer of 2008. They wanted Congress to suspend the 18.4¢ per gallon federal gas tax during the traditional high-price summer months to lower gasoline prices. Then-Senator Barack Obama chided them for "pandering," arguing in part that such a suspension would primarily benefit oil firms rather than consumers. A similar debate took place in Britain in 2008. In 2011, Representative Heath Shuler proposed a 45-day federal gasoline tax holiday, as did legislators in Illinois, Indiana, New Hampshire, and New York. In 2013, the Indiana House minority leader proposed a tax holiday.

A critical issue in these debates concerns who pays the tax. Do firms pass the gasoline tax on to consumers in the form of higher prices, or do they absorb the tax themselves? Is the ability of firms to pass a gas tax on to consumers different in the short run (such as during the summer months) than in the long run?

We can use a supply-and-demand analysis to answer such questions. When an underlying factor that affects the demand or supply curve—such as a tax—changes, the equilibrium price and quantity also change. Chapter 2 showed that you can predict the direction of the change—the *qualitative* change—in equilibrium price and quantity even without knowing the exact shape of the supply and demand curves. In most of the examples in Chapter 2, all you needed to know to give a qualitative answer was the direction in which the supply curve or demand curve shifted when an underlying factor changed.

To determine the exact amount the equilibrium quantity and price change—the *quantitative* change—you can use estimated equations for the supply and demand functions, as we demonstrated using the avocado example in Chapter 2. This chapter shows how to use a single number to describe how sensitive the quantity demanded or supplied is to a change in price and how to use these summary numbers to obtain quantitative answers to what-if questions, such as the effects of a tax on the price that consumers pay.

In this chapter, we examine four main topics

1. **How Shapes of Supply and Demand Curves Matter.** The effect of a shock (such as a new tax or an increase in the price of an input) on market equilibrium depends on the shape of supply and demand curves.
2. **Sensitivity of the Quantity Demanded to Price.** The *price elasticity of demand* summarizes the sensitivity of the quantity demanded.
3. **Sensitivity of the Quantity Supplied to Price.** The *price elasticity of supply* summarizes the sensitivity of the quantity supplied.
4. **Effects of a Sales Tax.** How a sales tax increase affects the equilibrium price and quantity of a good and whether the tax falls more heavily on consumers or suppliers depends on the shape of the supply and demand curves.

3.1 How Shapes of Supply and Demand Curves Matter

The shapes of the supply and demand curves determine by how much a shock affects the equilibrium price and quantity. We illustrate the importance of the shape of the demand curve using estimated avocado demand and supply curves.[1] The supply of avocados depends on the price of avocados and the price of fertilizer, a major input in producing avocados. A 55¢ per pound increase in the price of fertilizer causes the avocado supply curve to shift to the left from S^1 to S^2 in panel a of Figure 3.1. The *shift of the supply curve* causes a *movement along the demand curve*, D^1, which is downward sloping. The equilibrium quantity falls from 80 to 72 million pounds (lbs) per month, and the equilibrium price rises 20¢ from \$2.00 to \$2.20 per lb, hurting consumers.

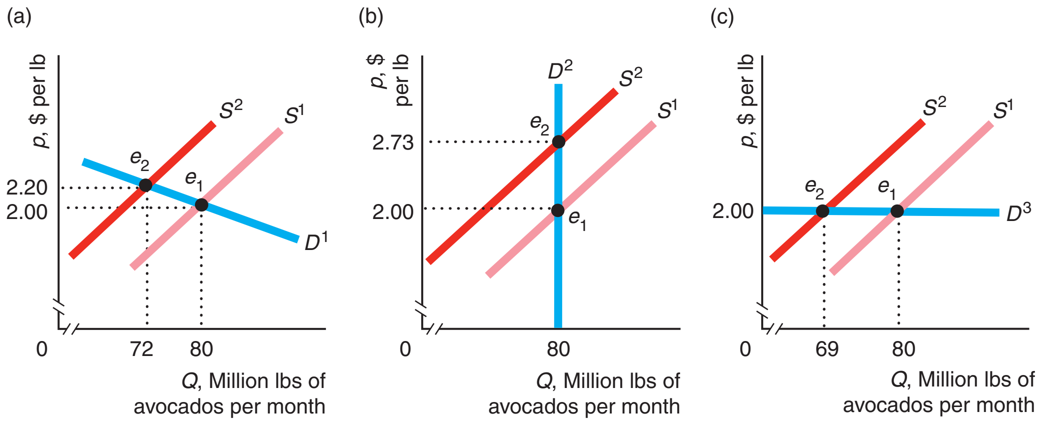
A supply shock would have different effects if the demand curve had a different shape. Suppose that the quantity demanded was not sensitive to a change in the price, so consumers demand the same amount no matter what the price is, as in vertical demand curve D^2 in panel b. A 55¢ increase in the price of fertilizer again shifts the supply curve from S^1 to S^2. Equilibrium quantity does not change, but the price consumers pay rises by 73¢ to \$2.73. Thus, the amount consumers spend rises by more if the demand curve is vertical than if it is downward sloping.

Now suppose that consumers are very sensitive to price, as in the horizontal demand curve, D^3, in panel c. Consumers will buy virtually unlimited quantities of avocados at \$2 per lb (or less), but, if the price rises even slightly, they stop buying avocados. Here an increase in the price of fertilizer has *no* effect on the price consumers pay; however, the equilibrium quantity drops to 69 million lbs per month. Thus, the shape of the demand curve affects the amount by which the equilibrium quantity of avocados falls and the equilibrium price rises in response to an increase in the price of fertilizer.

[1]To obtain our estimated supply and demand curves, we used estimates from Carman (2006), which we updated using more recent data from the California Avocado Commission and supplemented with information from other sources.

Figure 3.1 How the Effect of a Supply Shock Depends on the Shape of the Demand Curve

A 55¢ increase in the price of fertilizer shifts the avocado supply curve to the left from S^1 to S^2. (a) Using the estimated downward-sloping linear demand curve, D^1, the equilibrium price rises from \$2.00 to \$2.20 and the equilibrium quantity falls from 80 to 72. (b) If the demand curve were vertical, D^2, the supply shock would cause price to rise to \$2.73, while quantity would remain unchanged. (c) If the demand curve were horizontal, D^3, the supply shock would not affect price but would cause quantity to fall to 69.



3.2 Sensitivity of the Quantity Demanded to Price

Knowing how much quantity demanded falls as the price increases, holding all else constant, is therefore important in predicting the effect of a shock in a supply-and-demand model. We can determine the reduction in the quantity demanded as the price rises using an accurate drawing of the demand curve or the demand function (the equation that describes the demand curve). However, it is convenient to summarize the relevant information in a number that can be used to answer what-if questions without having to write out an equation or draw a graph. Armed with such a summary number, a firm can predict the effect on the price of its product and its revenue—price times quantity sold—from a shift in the market supply curve.

In this section, we discuss a summary number that describes how much the quantity demanded changes in response to a price increase from a given initial price—the *price elasticity of demand*. In the next section, we discuss a similar number for the supply curve—the *price elasticity of supply*. Then, we show how the government can use these summary numbers for supply and demand to predict the effect of a new sales tax on the equilibrium price, firms' revenues, and tax receipts.

elasticity
the percentage change in a variable in response to a given percentage change in another variable

The most commonly used measure of the sensitivity of one variable, such as the quantity demanded, to a change in another variable, such as price, is an **elasticity**, which is the percentage change in one variable in response to a given percentage change in another variable.

Price Elasticity of Demand

price elasticity of demand
the percentage change in the *quantity demanded* in response to a given percentage change in the price

The **price elasticity of demand** (or in common use, the *elasticity of demand*) is the percentage change in the quantity demanded, Q, in response to a given percentage change in the price, p, at a particular point on the demand curve. The price elasticity of demand (represented by , the Greek letter epsilon) is

$$= \frac{\text{percentage change in quantity demanded}}{\text{percentage change in price}} = \frac{\Delta Q/Q}{\Delta p/p}, \tag{3.1}$$

where the symbol Δ (the Greek letter delta) indicates a change, so ΔQ is the change in the quantity demanded; $\Delta Q/Q$ is the percentage change in the quantity demanded; Δp is the change in price; and $\Delta p/p$ is the percentage change in price.[2] For example, if a 1% increase in the price results in a 3% decrease in the quantity demanded, the elasticity of demand is $= -3\%/1\% = -3$.[3] Thus, the elasticity of demand is a pure number (it has no units of measure).

A negative sign on the elasticity of demand illustrates the Law of Demand: Consumers demand less quantity as the price rises. The elasticity of demand concisely answers the question, "How much does quantity demanded fall in response to a 1% increase in price?" A 1% increase in price leads to an % change in the quantity demanded.

It is often more convenient to calculate the elasticity of demand using an equivalent expression,

$$= \frac{\Delta Q/Q}{\Delta p/p} = \frac{\Delta Q}{\Delta p}\frac{p}{Q}, \tag{3.2}$$

where $\Delta Q/\Delta p$ is the ratio of the change in quantity to the change in price (the inverse of the demand curve's slope).

We can use Equation 3.2 to calculate the elasticity of demand for a linear demand curve, which has a general linear demand function (holding fixed other variables that affect demand) of

$$Q = a - bp. \tag{3.3}$$

In Equation 3.3, a is the quantity demanded when price is zero, $Q = a - (b \times 0) = a$, and $-b$ is the ratio of the fall in quantity to the rise in price, $\Delta Q/\Delta p$.[4] Thus, for a linear demand curve, the elasticity of demand is

$$= \frac{\Delta Q}{\Delta p}\frac{p}{Q} = -b\frac{p}{Q}. \tag{3.4}$$

Solved Problem 3.1

The estimated equation for the linear U.S. corn demand curve is

$$Q = 15.6 - 0.5p, \tag{3.5}$$

where p is the price in dollars per bushel and Q is the quantity demanded in billion bushels per year.[5] What is the elasticity of demand at the point on the demand curve where the price is $p = \$7.20$ per bushel?

---

[2]When we use calculus, we use infinitesimally small changes in price (Δp approaches zero), so we write the elasticity as $(dQ/dp)(p/Q)$. When discussing elasticities, we assume that the change in price is small.

[3]Because demand curves slope downward according to the Law of Demand, the elasticity of demand is a negative number. Realizing that, some economists ignore the negative sign when reporting a demand elasticity. Instead of saying the demand elasticity is -3, they would say that the elasticity is 3 (with the negative sign understood).

[4]As the price increases from p_1 to p_2, the quantity demanded goes from Q_1 to Q_2, so the change in quantity demanded is $\Delta Q = Q_2 - Q_1 = (a - bp_2) - (a - bp_1) = -b(p_2 - p_1) = -b\Delta p$. Thus, $\Delta Q/\Delta p = -b$. (The slope of the demand curve is $\Delta p/\Delta Q = -1/b$.)

[5]This demand curve is a linearized version of the estimated demand curve in Roberts and Schlenker (2013). I have rounded their estimated elasticities slightly for algebraic simplicity.

Answer

Substitute the slope coefficient b, the price, and the quantity values into Equation 3.4. Equation 3.5 is a special case of the general linear demand function Equation 3.3 ($Q = a - bp$), where $a = 15.6$ and $b = 0.5$. Evaluating Equation 3.5 at $p = \$7.20$, we find that the quantity demanded is $Q = 15.6 - (0.5 \times 7.20) = 12$ billion bushels per year. Substituting $b = 0.5$, $p = \$7.20$, and $Q = 12$ into Equation 3.4, we learn that the elasticity of demand at this point on the demand curve is

$$= -b\frac{p}{Q} = -0.5 \times \frac{7.20}{12} = -0.3.$$

Comment: At this point on the demand curve, a 1% increase in the price of corn leads to a 0.3% fall in the quantity of corn demanded. That is, a price increase causes a less than proportionate fall in the quantity of corn demanded.

Application

The Demand Elasticities for Google Play and Apple Apps

The Apple App Store (iOS) had 2 million and Google Play (Android) had 2.2 million apps (mobile applications for smart phones and tablets) as of June 2016. How price sensitive are app consumers? Are Apple aficionados more or less price sensitive than people who use Android devices?

Ghose and Han (2014) estimated the demand elasticity for an app in the Apple App Store is −2.0. That is, a 1% increase in price causes a 2% drop in the demand for an Apple app. Thus, demand is elastic in the Apple App Store. The estimated demand elasticity for an app in Google Play is −3.7, which means that Android app consumers are nearly twice as price sensitive as are Apple app consumers.

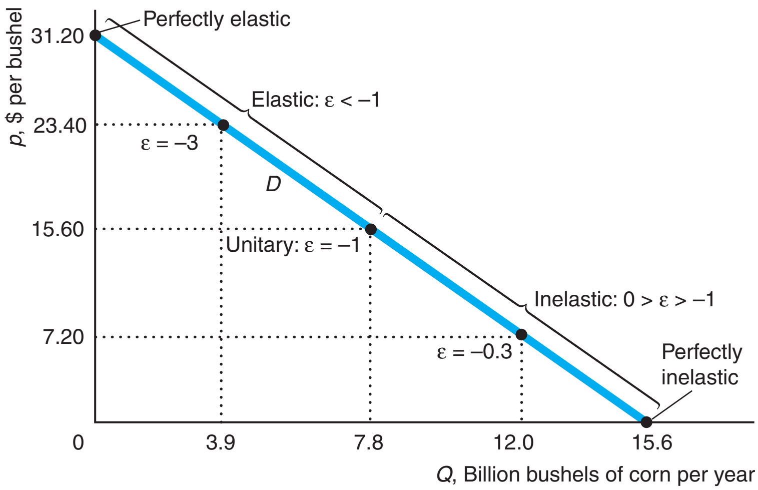
Elasticity Along the Demand Curve

The elasticity of demand varies along most demand curves. The elasticity of demand is different at every point along a downward-sloping linear demand curve, but the elasticities are constant along horizontal and vertical linear demand curves.

Downward-Sloping Linear Demand Curve On strictly downward-sloping linear demand curves—those that are neither vertical nor horizontal—the elasticity of demand, $= -b(p/Q)$, b is constant along the line, but the ratio of p/Q varies, so the elasticity also varies. The higher the price, the more negative the demand elasticity. That is, a 1% increase in price causes a larger percentage fall in quantity near the top (left) of the demand curve than near the bottom (right).

Figure 3.2 shows the linear corn demand curve given by Equation 3.5. Where the demand curve hits the quantity axis ($p = 0$ and $Q = a = 15.6$ billion bushels per year), the elasticity of demand is $= -b(0/a) = 0$, according to Equation 3.4. If the initial price is zero, a 1% increase in price does not raise the price, so quantity does not change. At a point where the elasticity of demand is zero, the demand curve is said to be *perfectly inelastic*. As a physical analogy, if you try to stretch an inelastic steel rod, the length does not change. The change in the price is the force pulling at

Figure 3.2 Elasticity Along the Corn Demand Curve MyLab Economics Video



With a linear demand curve, such as the corn demand curve, the higher the price, the more elastic the demand curve (is larger in absolute value—a larger negative number). The demand curve is perfectly inelastic (= 0) where it hits the horizontal axis, and it is perfectly elastic where it hits the vertical axis. At the midpoint of the demand, it has unitary elasticity (= −1).

demand. If the quantity demanded does not change in response to this pulling, it is perfectly inelastic.

For quantities between the midpoint of the linear demand curve and the lower end, where $Q = a$, the demand elasticity lies between 0 and −1; that is, $0 > \quad > -1$. A point along the demand curve, where the elasticity is between 0 and −1, is *inelastic* (but not perfectly inelastic). Where the demand curve is inelastic, a 1% increase in price leads to a fall in quantity of less than 1%. For example, as we saw in Solved Problem 3.1, at the point where $p = \$7.20$ and $Q = 12$, $\quad = -0.3$, so a 1% increase in price causes quantity to fall by 0.3%. A physical analogy is a piece of rope that does not stretch much—is inelastic—when you pull on it: Changing price has relatively little effect on quantity.

At the midpoint of the linear demand curve, $p = a/(2b) = \$15.60$ and $Q = a/2 = 7.8$, so $\quad = -bp/Q = -b(a/[2b])/(a/2) = -1$. We call such an elasticity of demand a *unitary elasticity*: A 1% increase in price causes a 1% fall in quantity.

At prices higher than at the midpoint of the demand curve, the elasticity of demand is less than negative one, $\quad < -1$. In this range, the demand curve is called *elastic*. A physical analogy is a rubber band that stretches substantially when you pull on it. A 1% increase in price causes quantity to fall by more than 1%. Figure 3.2 shows that the elasticity is −3 where $p = \$23.40$ and $Q = 3.9$: A 1% increase in price causes a 3% drop in quantity.

As the price rises, the elasticity gets more and more negative, approaching negative infinity. Where the demand curve hits the price axis, it is *perfectly elastic*.[6] At the price $a/b = \$31.20$ where $Q = 0$, a 1% decrease in p causes the quantity demanded to become positive, which is an infinite increase in quantity.

The elasticity of demand varies along most demand curves, not just downward-sloping linear ones. Along a special type of demand curve, called a *constant elasticity*

[6]The demand curve hits the price axis at $p = a/b$ and $Q = 0$, so the elasticity is $-bp/0$. As the quantity approaches 0, the elasticity approaches negative infinity, $-\infty$. An intuition for this convention is provided by looking at a sequence, where −1 divided by 1/10 is −10, −1 divided by 1/100 is −100, and so on. The smaller the number we divide by, the more negative is the result, which goes to $-\infty$ in the limit.

demand curve, however, the elasticity is the same at every point along the curve.[7] Two extreme cases of these constant-elasticity demand curves are the strictly vertical and the strictly horizontal linear demand curves.

Horizontal Demand Curve The demand curve that is horizontal at p^* in panel a of Figure 3.3 shows that people are willing to buy as much as firms sell at any price less than or equal to p^*. If the price increases even slightly above p^*, however, demand falls to zero. Thus, a small increase in price causes an infinite drop in quantity, so the demand curve is perfectly elastic at every quantity.

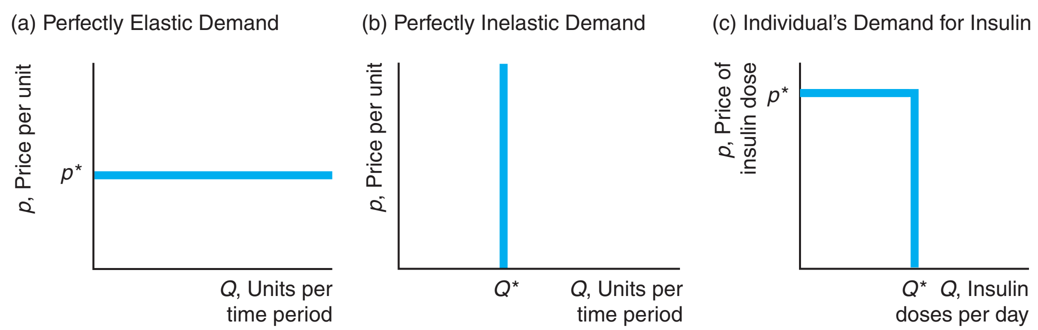
Why would a demand curve be horizontal? One reason is that consumers view this good as identical to another good and do not care which one they buy. Suppose that consumers view Washington apples and Oregon apples as identical. They won't buy Washington apples if these sell for more than apples from Oregon. Similarly, they won't buy Oregon apples if their price is higher than that of Washington apples. If the two prices are equal, consumers do not care which type of apple they buy. Thus, the demand curve for Oregon apples is horizontal at the price of Washington apples.

Vertical Demand Curve A vertical demand curve, panel b in Figure 3.3, is perfectly inelastic everywhere. Such a demand curve is an extreme case of the linear demand curve with an infinite (vertical) slope. If the price goes up, the quantity demanded is unchanged ($\Delta Q/\Delta p = 0$), so the elasticity of demand must be zero: $(\Delta Q/\Delta p)(p/Q) = 0(p/Q) = 0$.

A demand curve is vertical for *essential goods*—goods that people feel they must have and will pay anything to get. Because Jerry is a diabetic, his demand curve for insulin could be vertical at a day's dose, Q^*. More realistically, he may have a demand curve (panel c of Figure 3.3) that is perfectly inelastic only at prices below p^*, the maximum price he can afford to pay. Because he cannot afford to pay more than p^*, he buys nothing at higher prices. As a result, his demand curve is perfectly elastic up to Q^* units at a price of p^*.

Figure 3.3 Vertical and Horizontal Demand Curves

(a) A horizontal demand curve is perfectly elastic at p^*. (b) A vertical demand curve is perfectly inelastic at every price. (c) A diabetic's demand curve for insulin is perfectly inelastic below p^* and perfectly elastic at p^*, which is the maximum price the individual can afford to pay.



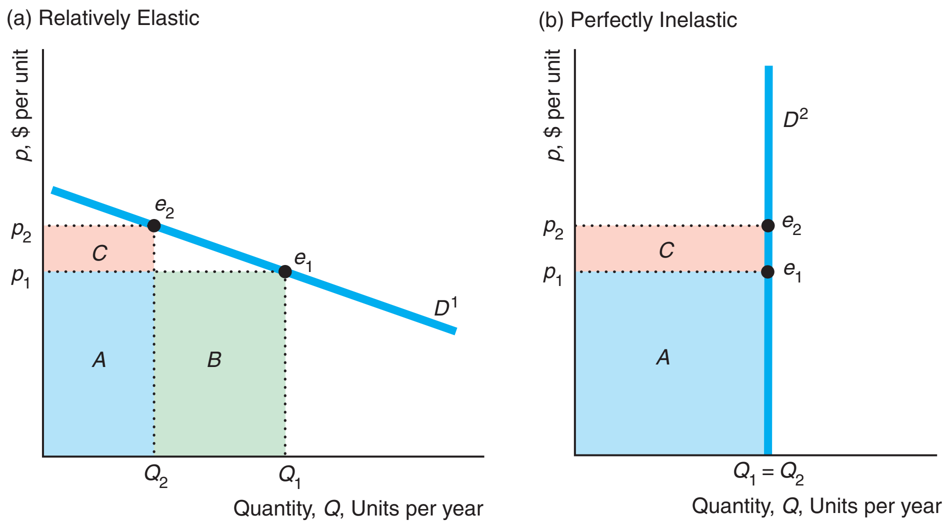
[7]Constant-elasticity demand curves all have the form $Q = Ap$, where A is a positive constant and , a negative constant, is the demand elasticity at every point along these demand curves. See Question 2.6.

Demand Elasticity and Revenue

Any shock that causes the equilibrium price to change affects the industry's *revenue*, which is the price times the market quantity sold. At the initial price p_1 in panel a of Figure 3.4, consumers buy Q_1 units at point e_1 on the demand curve D^1. Thus, the revenue is $R_1 = p_1 \times Q_1$. The revenue equals the area of the rectangle with a height of p_1 and a length of Q_1. At price p_1, the area of the revenue rectangle equals areas $A + B$. If the equilibrium price rises to p_2, so that the quantity demanded falls to Q_2, the new revenue is $R_2 = p_2 \times Q_2$, or area $A + C$. The change in the revenue due to the increase in price is $R_2 - R_1 = (A + C) - (A + B) = C - B$. Whether the revenue rises or falls when the price increases depends on the elasticity of demand, as the next Solved Problem shows.

Figure 3.4 Effect of a Price Change on Revenue

(a) When the price is p_1, consumers buy Q_1 units at e_1 on the demand curve D^1, so revenue is $R_1 = p_1 \times Q_1$, which is area $A + B$. If the price increases to p_2, consumers buy Q_2 units at e_2, so revenue is $R_2 = p_2 \times Q_2$, which is area $A + C$. Thus, the change in revenue is $R_2 - R_1 = (A + C) - (A + B) = C - B$. (b) When the demand curve is inelastic, D^2, the quantity does not change, $Q_1 = Q_2$, as the price rises from p_1 to p_2. Because this figure does not have an area B (as in panel a), revenue increases by area $C = (p_2 - p_1)Q_2$.



Solved Problem 3.2

Does a price increase cause revenue to rise or fall if the demand curve is elastic at the initial price? How does revenue change if the demand curve is inelastic?

Answer

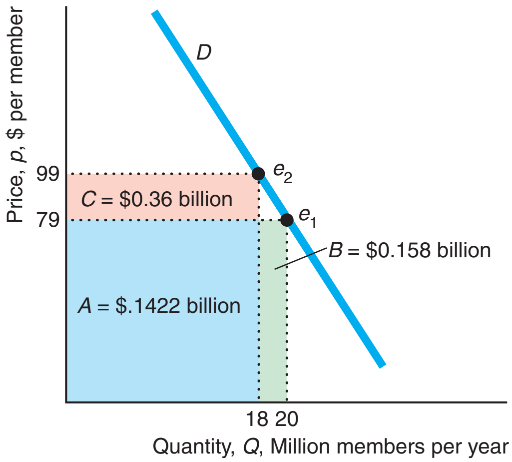
1. *Show that if the demand curve is elastic at the initial price, then area C is relatively small.* Panel a of Figure 3.4 shows a relatively flat demand curve, D^1, which is elastic at the initial price. A small percentage increase in price causes a greater percentage drop in quantity, so that area B is large and area C is small. With such a demand curve, an increase in price causes revenue to fall.[8]

[8] This result is discussed in greater detail using mathematics in Chapter 11.

2. *Consider the extreme case where the demand curve is perfectly inelastic and then generalize to the inelastic case*. In panel b of Figure 3.4, the demand curve D^2 is vertical and hence perfectly inelastic. Consequently, as the price rises from p_1 to p_2, the quantity demanded does not change, $Q_1 = Q_2$, so this figure does not have an area B, unlike panel a. Revenue increases by area $C = (p_2 - p_1)Q_2$. If the demand curve were inelastic (but not perfectly inelastic) at the initial price, a price increase would cause a less than proportional decrease in quantity. Because price rises more than quantity falls, revenue rises: Area B would be relatively thin and have little area; so $C > B$, as the following Application shows.

Application

Amazon Prime



In 2014, Amazon.com considered raising the annual price of its Amazon Prime shipping and streaming-video service from \$79 to \$99 per year, a 25% increase. Such a price increase would be the first in this program's nine-year history.

Amazon was concerned that a price increase would result in many customers dropping the service. The number of customers that it would lose would depend on the price elasticity of demand for Amazon Prime. According to analysts, Amazon expected to lose roughly 10% of its customers, going from 20 million members of its service to 18 million. That is, they estimated that the price elasticity of demand was $-10\%/25\% = -0.4$.

As a result, they predicted that area $C = (99 - 79)18 =$ \$360 million (or \$0.36 billion) and area $B = 79(20 - 18) =$ \$158 million (or \$0.158 billion), so the change in revenue would be positive: $\$0.36 - \$0.158 = \$0.202$ billion. Thus, Amazon raised its price.

Demand Elasticities over Time

The shape of the demand curve depends on the relevant period. Consequently, a short-run elasticity may differ substantially from a long-run elasticity. The duration of the *short run* depends on how long it takes consumers or firms to adjust for a particular good.

Two factors that determine whether short-run demand elasticities are larger or smaller than long-run elasticities are ease of substitution and storage opportunities. Often one can substitute between products in the long run but not in the short run. Consumers can easily store some goods, such as cans of peas, but not others, such as gasoline.

When U.S. gasoline prices fell sharply from \$3.75 per gallon in July 2014 to \$1.93 in February 2016, most consumers did not greatly alter the amount of gasoline that they demanded in the short run. Someone who drove 27 miles to and from work every day in a recently purchased Ford Focus did not suddenly start using more gasoline. However, if gas prices were to remain low in the long run, the share of consumers buying large cars would rise, many people would take jobs farther from home, and more families would take driving vacations.

Liddle (2012) estimated the gasoline demand elasticities across many countries and found that the short-run elasticity for gasoline was -0.16 and the long-run elasticity was -0.43. Thus, a 1% increase in price lowers the quantity demanded by only 0.16% in the short run but by more than twice as much, 0.43%, in the long run.

For goods that can be stored easily or are long-lived, short-run demand curves may be more elastic than long-run curves. If frozen orange juice goes on sale this week at your local supermarket, you may buy large quantities and store the extra in your freezer. As a result, you may respond more to a short-run price change for frozen orange juice than to a comparable long-run price change. Prince (2008) estimated that the demand curve for computers is more elastic in the short run, −2.74, than in the long run, −2.17. People buy computers when the price is relatively low and then keep them for years.

Because demand elasticities differ over time, the effect on revenue of a price increase also differs over time. For example, because the demand curve for gasoline is more inelastic in the short run than in the long run, a given increase in price raises revenue by more in the short run than in the long run.

Other Demand Elasticities

We refer to the price elasticity of demand as *the* elasticity of demand. However, other types of demand elasticities show how changes in relevant variables, other than the good's own price affect the quantity demanded. Two such demand elasticities are the income elasticity of demand and the cross-price elasticity of demand.

Income Elasticity As consumers' income increases, the demand curve shifts. If the demand curve shifts to the right, consumers demand a larger quantity at any given price. If instead the demand curve shifts to the left, consumers demand a smaller quantity at any given price.

income elasticity of demand
the percentage change in the *quantity demanded* in response to a given percentage change in income

We measure the sensitivity of the quantity demanded at a given price to changes in income using an elasticity. The **income elasticity of demand** (or *income elasticity*) is the percentage change in the quantity demanded in response to a given percentage change in income, Y. The income elasticity of demand is

$$\xi = \frac{\text{percentage change in quantity demanded}}{\text{percentage change in income}} = \frac{\Delta Q/Q}{\Delta Y/Y} = \frac{\Delta Q}{\Delta Y}\frac{Y}{Q},$$

where ξ is the Greek letter xi. If quantity demanded increases as income rises, the income elasticity of demand is positive. If the quantity does not change as income rises, the income elasticity is zero. Finally, if the quantity demanded falls as income rises, the income elasticity is negative.

We can calculate the income elasticity for avocados using the estimated demand function

$$Q = 104 - 40p + 20p_t + 0.01Y, \tag{3.6}$$

where we measure quantity in millions of lbs per month; the avocado price, p, and the tomato price, p_t, in dollars per lb; and average monthly income in dollars. Because the change in quantity as income changes is $\Delta Q/\Delta Y = 0.01$,[9] we can write the income elasticity as

$$\xi = \frac{\Delta Q}{\Delta Y}\frac{Y}{Q} = 0.01\frac{Y}{Q}.$$

At the competitive equilibrium, quantity is $Q = 80$ and income is $Y = 4{,}000$, so the income elasticity is $\xi = 0.01 \times (4{,}000/80) = 0.5$. The positive income elasticity

[9]At income Y_1, the quantity demanded is $Q_1 = 104 - 40p + 20p_t + 0.01Y_1$. At income Y_2, $Q_2 = 104 - 40p + 20p_t + 0.01Y_2$. Thus, $\Delta Q = Q_2 - Q_1 = 0.01(Y_2 - Y_1) = 0.01(\Delta Y)$, so $\Delta Q/\Delta Y = 0.01$.

shows that an increase in income causes the avocado demand curve to shift to the right. Holding the price of avocados constant at \$2 per lb, a 1% increase in income causes the demand curve for avocados to shift to the right by 0.4 ($= \xi \times 80 \times 0.01$) million lbs, which is about 0.5% of the equilibrium quantity.

Income elasticities play an important role in our analysis of consumer behavior in Chapter 5. Typically, goods that society views as necessities, such as food, have income elasticities between zero and one. Goods that society considers luxuries generally have income elasticities greater than one.

cross-price elasticity of demand
the percentage change in the *quantity demanded* in response to a given percentage change in the price of another good

Cross-Price Elasticity The **cross-price elasticity of demand** is the percentage change in the quantity demanded in response to a given percentage change in the price of another good, p_o. The cross-price elasticity is

$$\frac{\text{percentage change in quantity demanded}}{\text{percentage change in price of another good}} = \frac{\Delta Q/Q}{\Delta p_o/p_o} = \frac{\Delta Q}{\Delta p_o}\frac{p_o}{Q}.$$

When the cross-price elasticity is negative, the goods are *complements* (Chapter 2). If the cross-price elasticity is negative, people buy less of the good when the price of the other good increases: The demand curve shifts to the left. For example, if people like cream in their coffee, as the price of cream rises, they consume less coffee, so the cross-price elasticity of the quantity of coffee with respect to the price of cream is negative.

If the cross-price elasticity is positive, the goods are *substitutes* (Chapter 2). As the price of the other good increases, people buy more of this good. For example, the quantity demanded of avocados increases when the price of tomatoes, p_t, rises. From Equation 3.6, we know that $\Delta Q/\Delta p_t = 20$. As a result, the cross-price elasticity between the price of tomatoes and the quantity of avocados is

$$\frac{\Delta Q}{\Delta p_t}\frac{p_t}{Q} = 20\frac{p_t}{Q}.$$

At the equilibrium, $Q = 80$ million lbs per year and $p_t = \$0.80$ per lb, so the cross-price elasticity is $20 \times (0.8/80) = 0.2$. That is, as the price of tomatoes rises by 1%, the quantity of avocados demanded rises by one-fifth of 1%.

Taking account of cross-price elasticities is important in making business and policy decisions. For example, General Motors wants to know how much a change in the price of a Toyota affects the demand for its Chevy. Society wants to know if taxing soft drinks will substantially increase the demand for milk.

Application

Anti-Smoking Policies May Reduce Drunk Driving

Anyone who has seen bar scenes in old movies will know that smoking and drinking frequently go together: they're complements. Recent research has examined this relationship more rigorously, showing clearly that cigarettes and alcohol are complements. When the price of cigarettes rises or when smoking becomes more difficult (due to smoking bans, for example) people smoke less *and* drink less. The cross-elasticity of demand for alcohol with respect to cigarette prices is substantial. Krauss et al. (2014) found that a 1% increase in the price of cigarettes gives rise to almost a 1% decline in alcohol consumption. That is, the cross-price elasticity between alcoholic beverages and cigarettes is almost -1.

This complementarity between smoking and drinking implies that tobacco taxes (which raise the price of cigarettes) and restrictions on smoking have two health benefits. First, the tax reduces the negative health effects of smoking. Second, the tax reduces the negative effects of excessive drinking, including the very serious health risks created by drunk driving.

3.3 Sensitivity of the Quantity Supplied to Price

To answer many what-if questions, we need information about the sensitivity of the quantity supplied to changes in price. For example, to determine how a sales tax will affect market price, a government needs to know the sensitivity to price of both the quantity supplied and the quantity demanded.

Elasticity of Supply

price elasticity of supply the percentage change in the *quantity supplied* in response to a given percentage change in the price

Just as we can use the elasticity of demand to summarize information about the shape of a demand curve, we can use the elasticity of supply to summarize information about the supply curve. The **price elasticity of supply** (or *elasticity of supply*) is the percentage change in the quantity supplied in response to a given percentage change in the price. The price elasticity of supply (represented by η, the Greek letter eta) is

$$\eta = \frac{\text{percentage change in quantity supplied}}{\text{percentage change in price}} = \frac{\Delta Q/Q}{\Delta p/p} = \frac{\Delta Q}{\Delta p}\frac{p}{Q}, \tag{3.7}$$

where Q is the *quantity supplied*. If $\eta = 2$, a 1% increase in price leads to a 2% increase in the quantity supplied.

The definition of the elasticity of supply, Equation 3.7, is very similar to the definition of the elasticity of demand, Equation 3.1. The key distinction is that the elasticity of supply describes the movement along the *supply* curve as price changes, whereas the elasticity of demand describes the movement along the *demand* curve as price changes. That is, in the numerator, supply elasticity depends on the percentage change in the *quantity supplied*, whereas demand elasticity depends on the percentage change in the *quantity demanded*.

If the supply curve is upward sloping, $\Delta p/\Delta Q > 0$, the supply elasticity is positive: $\eta > 0$. If the supply curve slopes downward, the supply elasticity is negative: $\eta < 0$.

We use the terms *inelastic* and *elastic* to describe *upward-sloping* supply curves, just as we did for demand curves. If $\eta = 0$, we say that the supply curve is *perfectly inelastic*: The supply does not change as price rises. If $0 < \eta < 1$, the supply curve is *inelastic* (but not perfectly inelastic): A 1% increase in price causes a less than 1% rise in the quantity supplied. If $\eta = 1$, the supply curve has a *unitary elasticity*: A 1% increase in price causes a 1% increase in quantity. If $\eta > 1$, the supply curve is *elastic*. If η is infinite, the supply curve is *perfectly elastic*.

The supply function of a linear supply curve is

$$Q = g + hp, \tag{3.8}$$

where g and h are constants. By the same reasoning as before, $\Delta Q = h\Delta p$, so $h = \Delta Q/\Delta p$ shows the change in the quantity supplied as price changes. Thus, the elasticity of supply for a linear supply function is

$$\eta = \frac{\Delta Q}{\Delta p}\frac{p}{Q} = h\frac{p}{Q}. \tag{3.9}$$

We can illustrate this calculation using the U.S. supply function for corn (based on Roberts and Schlenker, 2013),

$$Q = 10.2 + 0.25p, \tag{3.10}$$

where Q is the quantity of corn supplied in billion bushels per year and p is the price of corn in dollars per bushel. Equation 3.10 is a special case of the general supply function Equation 3.8, $Q = g + hp$, where $g = 10.2$ and $h = 0.25$.

Thus, if the price is \$7.20 per bushel and the quantity is 12 billion bushels per year, we substitute $h = 0.25$, $p = 7.20$, and $Q = 12$ into Equation 3.9 to find the elasticity of supply at this point on the supply curve:

$$\eta = h\frac{p}{Q} = 0.25 \times \frac{7.20}{12} = 0.15.$$

At this point on the supply curve, a 1% increase in the price of corn leads to a 0.15% rise in the quantity of corn demanded, which is only about a seventh as much. That is, the supply curve is inelastic at this point.

Elasticity Along the Supply Curve

The elasticity of supply varies along most linear supply curves. Only *constant elasticity of supply curves* have the same elasticity at every point along the curve.[10] Vertical and horizontal supply curves are two extreme examples of both constant elasticity of supply curves and linear supply curves.

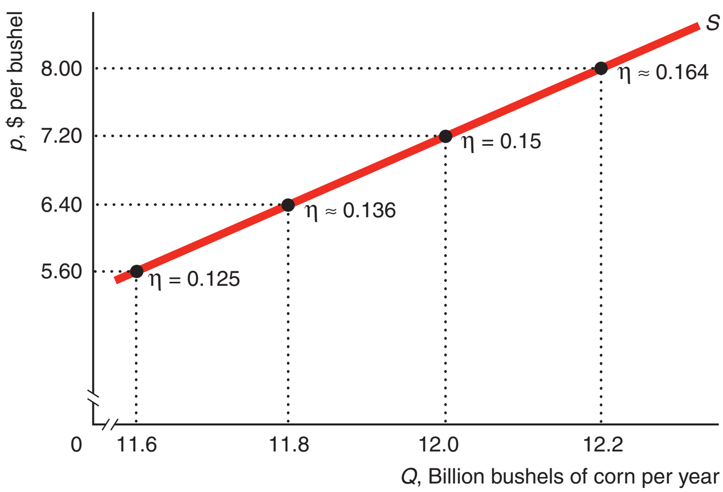
The elasticity of supply for a linear supply function is $\eta = h(p/Q)$, Equation 3.9. If $h > 0$, as the ratio p/Q rises, the supply elasticity rises. As the corn supply function in Equation 3.10 shows, $h = 0.25$ is positive for corn, so the elasticity of supply, $\eta = 0.25(p/Q)$, increases as p/Q rises. The corn supply curve in Figure 3.5 is inelastic at each point shown. It is 0.125 at $p = \$5.60$, 0.15 at $p = \$7.20$, and about 0.164 at $p = \$8$.

A supply curve that is vertical at a quantity, Q^*, is *perfectly inelastic*. Regardless of the price, firms supply Q^*. An example of inelastic supply is a perishable item such as fresh fruit. If suppliers do not sell the perishable good, it quickly becomes worthless. Thus, the seller accepts any market price for the good.

A supply curve that is horizontal at a price, p^*, is *perfectly elastic*. Firms supply as much as the market wants—a potentially unlimited amount—if the price is p^* or above. Firms supply nothing at a price below p^*, which does not cover their cost of production.

Figure 3.5 Elasticity Along the Corn Supply Curve

The elasticity of supply, η, varies along the corn supply curve. The higher the price, the larger is the supply elasticity.



[10]Constant elasticity of supply curves are of the form $Q = Bp^{\eta}$, where B is a constant and η is the constant elasticity of supply at every point along the curve.

Supply Elasticities over Time

Supply curves may have different elasticities in the short run than in the long run. If a manufacturing firm wants to increase production in the short run, it can do so by hiring workers to operate its machines around the clock. The fixed size of the firm's manufacturing plant and the number of machines it has limits how much the firm can expand its output. However, in the long run, the firm can build another plant and buy more equipment. Thus, we would expect this firm's long-run supply elasticity to be greater than its short-run elasticity. For example, Clemens and Gottlieb (2014) estimated that the health care supply elasticity is twice as elastic in the long run, 1.4, as in the short run, 0.7. It takes time to hire skilled medical care providers.

Application

Oil Drilling in the Arctic National Wildlife Refuge

We can use information about supply and demand elasticities to answer an important public policy question: Would selling oil from the Arctic National Wildlife Refuge substantially affect the price of oil? The Arctic National Wildlife Refuge, established in 1980, is the largest of Alaska's 16 national wildlife refuges, covers 20 million acres, and is believed to contain large deposits of petroleum (about the amount consumed in the United States in 2005). For decades, a debate has raged over whether the owners of Arctic National Wildlife Refuge—the citizens of the United States—should keep it undeveloped or permit oil drilling.[11]

In the simplest form of this complex debate, President Barack Obama has sided with environmentalists who stress that drilling would harm the wildlife refuge and pollute the environment. On the other side, the Republican Governors Association (in 2012), former President George W. Bush, and other drilling proponents argue that extracting this oil would substantially reduce the price of petroleum as well as decrease U.S. dependence on foreign oil. Recent large fluctuations in the price of gasoline and unrest in the Middle East have heightened this intense debate.

The effect of selling Arctic National Wildlife Refuge oil on the world price of oil is a key element of this dispute. We can combine oil production information with supply and demand elasticities to make a "back of the envelope" estimate of the price effects.

Analysts dispute how much Arctic National Wildlife Refuge can produce. The Department of Energy's Energy Information Service predicts that production from Arctic National Wildlife Refuge would average about 800,000 barrels per day. That production would be less than 1% of the worldwide oil production, which averaged about 94 million barrels per day in 2016.

A report by the U.S. Department of Energy predicted that Arctic National Wildlife Refuge drilling could lower the price of oil by about 1%. Severin Borenstein, an economist who directed the University of California Energy Institute, concluded that drilling in the Arctic National Wildlife Refuge would reduce oil prices by no more than a few percentage

[11] I am grateful to Robert Whaples, who wrote an earlier version of this analysis. In the following discussion, we assume for simplicity that the oil market is competitive, and we use current values of prices and quantities even though drilling in Arctic National Wildlife Refuge could not take place for at least a decade.

points, and therefore would "never noticeably affect gasoline prices." In the following Solved Problem, we make our own calculation of the price effect of drilling in the Arctic National Wildlife Refuge.

Solved Problem 3.3

MyLab Economics Solved Problem

What would be the effect of Arctic National Wildlife Refuge production on the world price of oil given that $\quad = -0.25$, $\eta = 0.25$ (Baumeister and Peersman, 2013), the pre-Arctic National Wildlife Refuge daily world production of oil is $Q_1 = 94$ million barrels per day, the pre-Arctic National Wildlife Refuge world price is $p_1 = \$50$ per barrel, and daily Arctic National Wildlife Refuge production is 0.8 million barrels per day?[12] We assume that the supply and demand curves are linear and that the introduction of Arctic National Wildlife Refuge oil would cause a parallel shift in the world supply curve to the right by 0.8 million barrels per day.

Answer

1. *Determine the long-run linear demand function that is consistent with pre-Arctic National Wildlife Refuge world output and price.* The general formula for a linear demand curve is $Q = a - bp$, where a is the quantity when $p = 0$ (where the demand curve hits the horizontal axis) and $b = \Delta Q/\Delta p$. At the original equilibrium, e_1 in the figure, $p_1 = \$50$ and $Q_1 = 94$ million barrels per day, so the elasticity of demand is $\quad = (\Delta Q/\Delta p)(p_1/Q_1) = -b(50/94) = -0.25$. Using algebra, we find that $b = 0.25(94/50) = 0.47$, so the demand function is $Q = a - 0.47p$. At e_1, the quantity demanded is $Q = 94 = a - (0.47 \times 50)$. Using algebra, we find that $a = 94 + (0.47 \times 50) = 117.5$. Thus, the demand function is $Q = 117.5 - 0.47p$.

2. *Determine the long-run linear supply function that is consistent with pre-Arctic National Wildlife Refuge world output and price.* The general formula for a linear supply curve is $Q = c + dp$, where c is the quantity when $p = 0$ and $d = \Delta Q/\Delta p$. Where S^1 intercepts D at the original equilibrium, e_1, the elasticity of supply is $\eta = (\Delta Q/\Delta p)(p_1/Q_1) = d(50/94) = 0.25$. Solving this equation, we find that $d = 0.25(94/50) = 0.47$, so the supply function is $Q = c + 0.47p$. Evaluating this equation at e_1, $Q = 94 = b + (0.47 \times 50)$. Solving for b, we find that $b = 94 - (0.47 \times 50) = 70.5$. Thus, the supply function is $Q = 70.5 + 0.47p$.

3. *Determine the post-Arctic National Wildlife Refuge linear supply function.* The oil pumped from the refuge would cause a parallel shift in the supply curve, moving S^1 to the right by 0.8 to S^2. That is, the slope remains the same, but the intercept on the quantity axis increases by 0.8. Thus, the supply function for S^2 is $Q = 71.3 + 0.47p$.

4. *Use the demand curve and the post-Arctic National Wildlife Refuge supply function to calculate the new equilibrium price and quantity.* The intersection of S^2 and D determines the new equilibrium, e_2. Setting the right side of the demand function equal to the right side of the post-Arctic National Wildlife Refuge supply function, we obtain an expression for the post-Arctic National Wildlife Refuge price, p_2:

$$117.5 - 0.47p_2 = 71.3 + 0.47p_2.$$

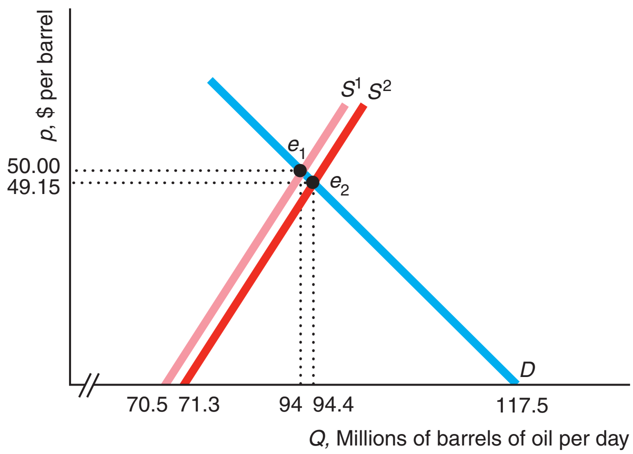
---

[12]This price is for June 2016, and it was the average price for 2015. From 2007 through 2015, the price of a barrel of oil fluctuated between about \$30 and \$140. The calculated percentage change in the price in Solved Problem 2.5 is not sensitive to the initial price of oil used in the calculation.

We can solve this expression for the new equilibrium price: $p_2 \approx$ \$49.15. That is, the price drops about \$0.85, or 1.7%. If we substitute this new price into either the demand curve or the post-Arctic National Wildlife Refuge supply curve, we find that the new equilibrium quantity is 94.4 million barrels per day. That is, equilibrium output rises by 0.4 million barrels per day (0.43%), which is only a little more than half of the predicted daily refuge supply, because other suppliers will decrease their output slightly in response to the lower price.

Comment: Our calculation—showing that selling Arctic National Wildlife Refuge oil would cause only a small drop in the world oil price—would not change substantially if our estimates of the elasticities of supply and demand were moderately larger or smaller or if the equilibrium price of oil were higher or lower. The main reason for this result is that the refuge output would be a very small portion of worldwide supply—the new supply curve would lie only slightly to the right of the initial supply curve. Thus, drilling in the Arctic National Wildlife Refuge alone cannot insulate the American market from international events that roil the oil market.

In contrast, a new war in the Persian Gulf could shift the worldwide supply curve to the left by up to 24 million barrels per day (the amount of oil produced in the Persian Gulf), or 30 times the Arctic National Wildlife Refuge's potential production. Such a shock would cause the price of oil to soar whether we drill in the Arctic National Wildlife Refuge or not.



3.4 Effects of a Sales Tax

Before voting for a new sales tax, legislators want to predict the effect of the tax on prices, quantities, and tax revenues. If the new tax will produce a large price increase, legislators who vote for the tax may lose their jobs in the next election. Voters' ire is likely to be even greater if the tax fails to raise significant tax revenues.

Governments use two types of sales taxes: ad valorem and specific taxes. Economists call the most common sales tax an *ad valorem* tax, while real people call it *the* sales tax. For every dollar the consumer spends, the government keeps a fraction, v, which is the ad valorem tax rate. For example, Japan's national ad valorem sales tax is 8%. If a Japanese consumer buys a Nintendo Wii for ¥40,000,[13] the government

[13]The symbol for Japan's currency, the yen, is ¥. Roughly, ¥111 = \$1.

collects $v \times ¥40{,}000 = 8\% \times ¥40{,}000 = ¥3{,}200$ in taxes, and the seller receives $(1 - v) \times ¥40{,}000 = 92\% \times ¥40{,}000 = ¥36{,}800$.[14]

The other type of sales tax is a *specific* or *unit* tax: The government collects a specified dollar amount, t, per unit of output. For example, the federal government collects $t = 18.4¢$ on each gallon of gas sold in the United States.

In this section, we examine four questions about the effects of a sales tax:

1. What effect does a specific sales tax have on equilibrium price and quantity as well as on tax revenue?
2. Are the equilibrium price and quantity dependent on whether the government collects a specific tax from suppliers or their customers?
3. Is it true, as many people claim, that producers *pass along* to consumers any taxes collected from producers? That is, do consumers pay the entire tax?
4. Do comparable ad valorem and specific taxes have equivalent effects on equilibrium prices and quantities and on tax revenue?

The shapes of the supply and demand curves determine how much a tax affects the equilibrium price and quantity and how much of the tax consumers pay. Knowing only the elasticities of supply and demand, which summarize the shapes of these curves, we can make accurate predictions about the effects of a new tax and determine how much of the tax is paid by consumers.

Effects of a Specific Tax on the Equilibrium

We illustrate the answer to the first question—regarding the equilibrium effects of a specific tax—using our estimated corn supply and demand curves. We start by assuming that the government collects a new specific tax of $t = \$2.40$ per bushel of corn from firms (farmers). If a customer pays a price of p to a firm, the government takes t, so the firm keeps $p - t$.

Thus, at every possible price paid by customers, firms are willing to supply less than when they received the full amount that customers paid. Before the tax, firms were willing to supply 11.6 billion bushels of corn per year at a price of \$5.60 per bushel, as the pre-tax supply curve S^1 in panel a of Figure 3.6 shows. After the tax, if customers pay \$5.60, firms receive only \$3.20 (=\$5.60 − \$2.40), so they are not willing to supply 11.6 billion bushels. For firms to be willing to supply that quantity, consumers must pay \$8.00 so that firms receive \$5.60 (=\$8.00 − \$2.40) after the tax. By this reasoning, the after-tax supply curve, S^2, is $t = \$2.40$ above the original supply curve S^1 at every quantity.

We can compare the pre-tax and post-tax equilibria to identify the effect of the tax. In panel a, the intersection of the pre-tax corn supply curve, S^1, and the corn demand curve D^1 determines the pre-tax equilibrium, e_1. The equilibrium price is $p_1 = \$7.20$, and the equilibrium quantity is $Q_1 = 12$.

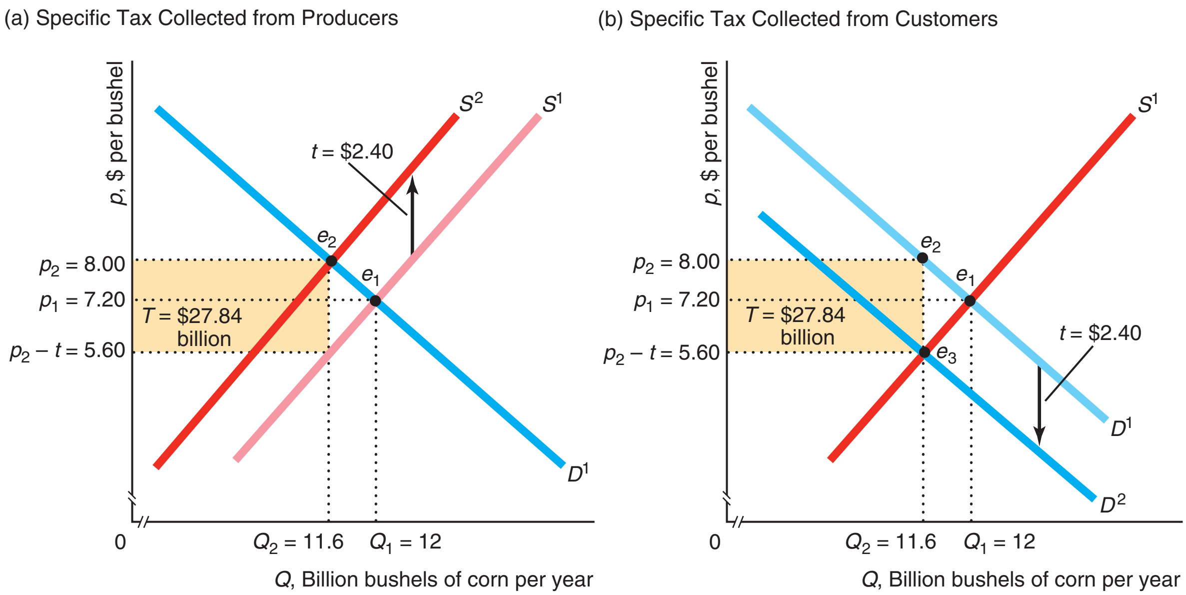
The tax shifts the supply curve to S^2, so the after-tax equilibrium is e_2, where customers pay $p_2 = \$8$, farmers receive $p_2 - \$2.40 = \$5.60$, and $Q_2 = 11.6$. Thus, the tax causes the equilibrium price that customers pay to increase

[14]For specificity, we assume that the price firms receive is $p = (1 - v)p^*$, where p^* is the price consumers pay and v is the ad valorem tax rate on the price consumers pay. However, many governments set the ad valorem sales tax, V, as an amount added to the price sellers charge, so consumers pay $p^* = (1 + V)p$. By setting v and V appropriately, the taxes are equivalent. Here $p = p^*/(1 + V)$, so $(1 - v) = 1/(1 + V)$. For example, if $V = \frac{1}{3}$, then $v = \frac{1}{4}$.

Figure 3.6 Equilibrium Effects of a Specific Tax MyLab Economics Video

(a) The specific tax of $t = \$2.40$ per bushel of corn collected from producers shifts the pre-tax corn supply curve, S^1, up to the post-tax supply curve, S^2. The tax causes the equilibrium to shift from e_1 (determined by the intersection of S^1 and D^1) to e_2 (intersection of S^2 with D^1). The equilibrium price—the price consumers pay—increases from $p_1 = \$7.20$ to $p_2 = \$8.00$. The government collects tax revenues of $T = tQ_2 = \$27.84$ billion per year.

(b) The specific tax collected from customers shifts the demand curve down by $t = \$2.40$ from D^1 to D^2. The intersection of D^2 and S^1 determines the new equilibrium e_3 and the new price that firms receive, $p_2 - t = \$5.60$. Corresponding to this point is e_2 on D^1, which shows the price that consumers pay, $p_2 = \$8.00$, and is the same point as when the tax is applied to suppliers in panel a.



($\Delta p = p_2 - p_1 = \$8 - \$7.20 = 80¢$) and the equilibrium quantity to fall ($\Delta Q = Q_2 - Q_1 = 11.6 - 12 = -0.4$).

Although the customers and producers are worse off because of the tax, the government acquires new tax revenue of $T = tQ = \$2.40$ per bushel $\times$ 11.6 billion bushels per year = \$27.84 billion per year. The length of the shaded rectangle in Figure 3.6 is $Q_2 = 11.6$ billion per year, and its height is $t = \$2.40$ per bushel, so the area of the rectangle equals the tax revenue. Thus, the answer to our first question is that *a specific tax causes the equilibrium price customers pay to rise, the equilibrium quantity to fall, and tax revenue to rise.*

The Equilibrium Is the Same No Matter Whom the Government Taxes

Our second question is, “Are the equilibrium price or quantity dependent on whether the specific tax is collected from suppliers or their customers?” We can use our supply-and-demand model to answer this question, showing that the equilibrium is the same regardless of whether the government collects the tax from suppliers or their customers.

If a customer pays a firm p for a bushel of corn, and the government collects a specific tax t from the customer, the total the customer pays is $p + t$. Suppose that customers bought a quantity Q at a price p^* before the tax. After the tax, they are willing to continue to buy Q only if the price falls to $p^* - t$, so that the after-tax price, $p^* - t + t$, remains at p^*. Consequently, the demand curve, from the perspective of firms, shifts down by $t = \$2.40$ from D^1 to D^2 in panel b of Figure 3.6.

The intersection of D^2 and the supply curve S^1 determines e_3. At e_3, the quantity is $Q_2 = 11.6$ and the price received by producers is $p_2 - t = \$5.60$. Directly above e_3 is the after-tax equilibrium e_2 on D^1, which shows the price that consumers pay is $p_2 = \$8$.

Comparing the two panels in Figure 3.6, we see that the after-tax equilibrium e_2 is the same regardless of whether the government imposes the tax on consumers or sellers. Similarly, the tax revenue that the government collects, $T = \$27.84$ billion, remains the same. Consequently, regardless of whether sellers or buyers pay the tax to the government, you can solve tax problems by shifting the supply curve or by shifting the demand curve.

Solved Problem 3.4

Show mathematically the effects on the equilibrium price and quantity of corn from a specific tax of $t = \$2.40$ collected from suppliers, as illustrated in panel a of Figure 3.6.

Answer

1. *Show how the tax shifts the supply curve.* Given the tax t, farmers receive only $p - t$ if consumers pay p. Consequently, their supply function changes from Equation 3.10, $Q = 10.2 + 0.25p$, to $Q = 10.2 + 0.25(p - t) = 10.2 + 0.25(p - 2.4) = 9.6 + 0.25p$. As panel a of Figure 3.6 shows, the after-tax supply curve, S^2, shifts up from S^1 by t as a result.

2. *Determine the after-tax equilibrium price by equating the after-tax supply function and the original demand function.* Because the tax does not affect the demand function Equation 3.5, $Q = 15.6 - 0.5p$, we equate the right sides of the demand function and the after-tax supply function: $15.6 - 0.5p = 9.6 + 0.25p$. Solving for the equilibrium price, we learn that $p = 8$.

3. *Determine the after-tax equilibrium quantity by substituting the equilibrium price in either the demand function or the after-tax supply function.* Using the demand function, we find that the equilibrium quantity is $Q = 15.6 - (0.5 \times 8) = 11.6$.

Firms and Customers Share the Burden of the Tax

Our third question concerns whether customers bear the entire burden of a tax, as many politicians and news stories assert.

> **Common Confusion:** Businesses pass any sales tax along to consumers, so that the entire burden of the tax falls on consumers.

This claim is not generally true, as we now show. We start by determining the share of the tax that consumers bear and then show how that share depends on the elasticities of supply and demand.

incidence of a tax on consumers
the share of the tax that consumers pay

Tax Incidence If the government sets a new specific tax of t, it changes the tax from 0 to t: $\Delta t = t - 0 = t$. The **incidence of a tax on consumers** is the share of the tax that consumers pay. The incidence of the tax that falls on consumers is $\Delta p/\Delta t$, the amount by which the price to consumers rises as a fraction of the amount the tax increases.

In the corn example in Figure 3.6, a $\Delta t = \$2.40$ increase in the specific tax causes customers to pay $\Delta p = 80¢$ more per bushel than they would if no tax were assessed. Thus, customers bear one-third of the incidence of the corn tax:

$$\frac{\Delta p}{\Delta t} = \frac{\$0.80}{\$2.40} = \frac{1}{3}.$$

The change in the price that firms receive is $(p_2 - t) - p_1 = (\$8 - \$2.40) - \$7.20 = \$5.60 - \$7.20 = -\$1.60$. That is, they receive \$1.60 less per bushel than they would in the absence of the tax. The incidence of the tax on farmers—the amount by which the price to them falls, divided by the tax—is $\$1.60/\$2.40 = \frac{2}{3}$. The sum of the share of the tax on customers, $\frac{1}{3}$, and that on firms, $\frac{2}{3}$, equals the entire tax effect, 1. Equivalently, the price increase to customers minus the price decrease to farmers equals the tax: $\$0.80 - (-\$1.60) = \$2.40 = t$.

Tax Effects Depend on Elasticities The tax incidence on customers depends on the elasticities of supply and demand. In response to a change in the tax of Δt, the price customers pay increases by

$$\Delta p = \left(\frac{\eta}{\eta - \quad}\right)\Delta t, \tag{3.11}$$

where   is the demand elasticity and η is the supply elasticity at the equilibrium (this equation is derived in Appendix 3A). The demand elasticity for corn is $\quad = -0.3$, and the supply elasticity is $\eta = 0.15$, so a change in the tax of $\Delta t = \$2.40$ causes the price customers pay to rise by

$$\Delta p = \left(\frac{\eta}{\eta - \quad}\right)\Delta t = \frac{0.15}{0.15 - [-0.3]} \times \$2.40 = 80¢,$$

as Figure 3.6 shows.

By dividing both sides of Equation 3.11 by Δt, we learn that the incidence of the tax that consumers pay is

$$\frac{\Delta p}{\Delta t} = \frac{\eta}{\eta - \quad}. \tag{3.12}$$

Thus, the incidence of the corn tax on consumers is

$$\frac{0.15}{0.15 - (-0.3)} = \frac{1}{3}.$$

For a given supply elasticity, the more elastic demand is, the less the equilibrium price rises when a tax is imposed. For example, if the corn supply elasticity remains 0.15, but the demand elasticity doubles to -0.6, the incidence on consumers would fall to $0.15/(0.15 - [-0.6]) = 0.2$, and the price customers pay would rise by only $[0.15/(0.15 - [-0.6])] \times \$2.40 = 48¢$ instead of 80¢.

Similarly, for a given demand elasticity, the greater the supply elasticity, the larger the increase in the equilibrium price customers pay when a tax is imposed. In the corn example, if the demand elasticity remains -0.3 but the supply elasticity doubles to 0.3, the incidence would rise to $0.3/(0.3 - [-0.3]) = 0.5$, and the price to customers would increase by $[0.3/(0.3 - [-0.3])] \times \$2.40 = \$1.20$ instead of 80¢.

Application

Taxes to Prevent Obesity

Many governments use taxes to discourage people from buying “sin” goods, such as cigarettes and alcohol. Recently, many governments have added sugar and fat to this list of sin goods. They tax sugar and fats to slow the worldwide rise in obesity rates. For a while, Hungary imposed a “hamburger tax” and a “chip tax” on fatty foods as well as a tax on soda and alcohol, with the proceeds going to health care. Finland, France, Mexico, and Norway tax products with refined sugar. The European Union, Romania, and South Africa have debated such taxes. The United Kingdom’s tax on sugar-sweetened beverages goes into effect in 2018. In the United States, at least 25 states tax soft drinks, candy, chewing gum, and snack foods such as potato chips.

A number of nutritionists have argued that cheap corn is a major cause of obesity. Farmers use corn for animal feed, so cheap corn lowers the price of meat and encourages greater meat (and fat) consumption. Food processors use corn to produce high-fructose corn syrup, the main sweetener in soft drinks, fruit drinks, and many foods, such as peanut butter and spaghetti sauce.

To discourage the consumption of meat, soft drinks, and other fattening foods, a number of nutritionists and others have called for a tax on corn or corn syrup.[15] Would a tax on corn significantly lower the consumption of corn? The answer depends on the corn demand and supply elasticities.

In Figure 3.6, we consider a gigantic specific tax of \$2.40 per bushel, which is one-third of the equilibrium price. However, this large tax only reduces the equilibrium quantity from 12 to 11.6 billion bushels per year, just $3\frac{1}{3}\%$. It is probably politically infeasible to impose such a large tax. And even with such a hefty tax, the reduction in corn production would be small. Thus, taxing corn isn’t an effective way to attack the obesity problem. However, the government may like this tax because it would raise gigantic amounts of tax revenue.

Why would a corn tax have such a small effect on quantity? Given the supply and demand elasticities, because the consumer tax incidence is $\frac{1}{3}$, only a third of the tax, 80¢, is passed on to customers, which is one-ninth of the equilibrium price.

Even if the entire gigantic tax were passed through to consumers so that the price to consumers would increase by \$2.40/\$7.20 $= \frac{1}{3}$, the quantity demanded would fall by only $\frac{1}{3} \quad = \frac{1}{3} \times (-0.3) \approx 10\%$, because the demand curve is inelastic at the equilibrium. We now discuss a supply elasticity that leads to consumers bearing the entire tax.

Solved Problem 3.5

If the supply curve is perfectly elastic and the demand curve is downward sloping, what is the effect of a \$1 specific tax collected from producers on equilibrium price and quantity, and what is the incidence on consumers? Why?

[15]Corn farmers receive large subsidies, which are negative taxes. Thus, rather than taxing corn, it would be more straightforward for the government to reduce these subsidies.

Answer

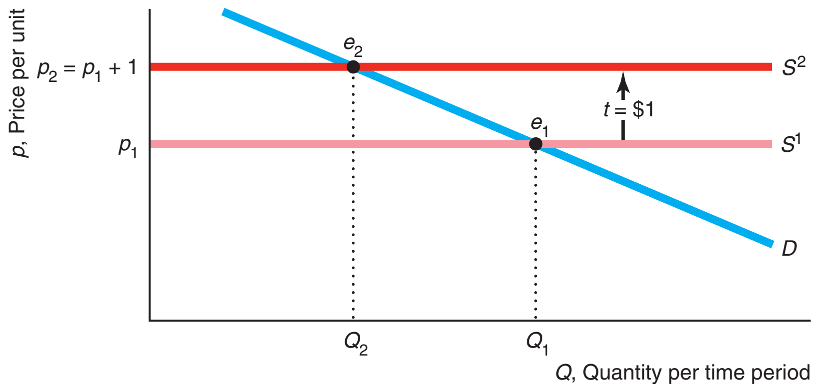
1. *Determine the equilibrium in the absence of a tax*. Before the tax, the perfectly elastic supply curve, S^1 in the graph, is horizontal at p_1. In the figure, the downward-sloping linear demand curve, D, intersects S^1 at the pre-tax equilibrium, e_1, where the price is p_1 and the quantity is Q_1.

2. *Show how the tax shifts the supply curve and determine the new equilibrium*. A specific tax of $1 shifts the pre-tax supply curve, S^1, upward by $1 to S^2, which is horizontal at $p_1 + 1$. The intersection of D and S^2 determines the after-tax equilibrium, e_2, where the price consumers pay is $p_2 = p_1 + 1$, the price firms receive is $p_2 - 1 = p_1$, and the quantity is Q_2.

3. *Compare the before- and after-tax equilibria*. The specific tax causes the equilibrium quantity to fall from Q_1 to Q_2, the price firms receive to remain at p_1, and the equilibrium price consumers pay to rise from p_1 to $p_2 = p_1 + 1$. The entire incidence of the tax falls on consumers:

$$\frac{\Delta p}{\Delta t} = \frac{p_2 - p_1}{\Delta t} = \frac{\$1}{\$1} = 1.$$

4. *Explain why*. Consumers must absorb the entire tax because firms will not supply the good at a price that is any lower than they received before the tax, p_1. Thus, the price must rise enough that the price suppliers receive after the tax is unchanged. As customers do not want to consume as much at a higher price, the equilibrium quantity falls.



Ad Valorem and Specific Taxes Have Similar Effects

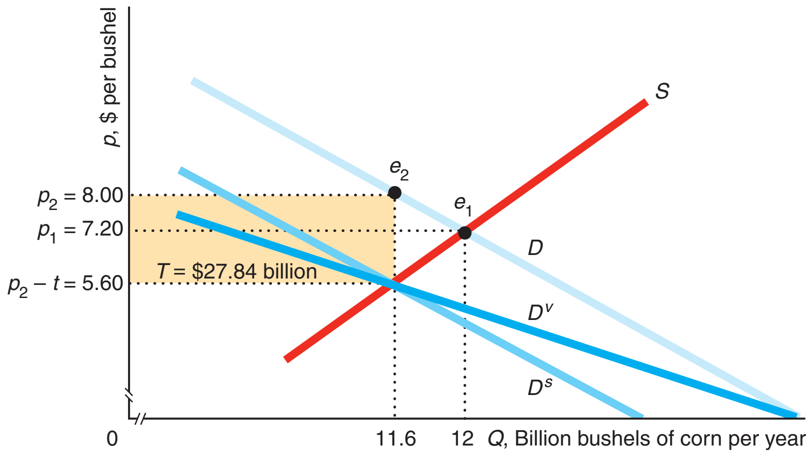
Our fourth question asks whether an ad valorem tax has the same effect on equilibrium price and quantity as a comparable specific tax. In contrast to specific sales taxes, governments levy ad valorem taxes on a wide variety of goods. Most states apply an ad valorem sales tax to almost all goods and services, exempting only a few staples such as food and medicine.

Suppose that the government imposes an ad valorem tax of v, instead of a specific tax, on the price that consumers pay for corn. We already know that the equilibrium price is $8 with a specific tax of $2.40 per bushel. At that price, an ad valorem tax of $v = \$2.40/\$8 = 30\%$ raises the same amount of tax per unit as a $2.40 specific tax.

It is usually easiest to analyze the effects of an ad valorem tax by shifting the demand curve. Figure 3.7 shows how a specific tax and an ad valorem tax shift the

Figure 3.7 Comparison of an Ad Valorem and a Specific Tax

Without a tax, the demand curve is D and the supply curve is S. An ad valorem tax of $v = 30\%$ shifts the demand curve facing firms to D^v. The gap between D and D^a, the per-unit tax, is larger at higher prices. In contrast, the demand curve facing firms given a specific tax of \$2.40 per bushel, D^s, is parallel to D. The after-tax equilibrium, e_2, and the tax revenue, T, are the same with both of these taxes.



corn demand curve. The specific tax shifts the original, pre-tax demand curve, D, down to D^s, which is parallel to the original curve. The ad valorem tax rotates the demand curve to D^a. At any given price p, the gap between D and D^a is vp, which is greater at high prices than at low prices. The gap is \$2.40 (=0.3 × \$8) per unit when the price is \$8, but only \$1.20 (=0.3 × \$4) when the price is \$4.

Imposing an ad valorem tax causes the after-tax equilibrium quantity, 11.6, to fall below the pre-tax quantity, 12, and the post-tax price, p_2, to rise above the pre-tax price, p_1. The tax collected per unit of output is $t = vp_2$. The incidence of the tax that falls on consumers is the change in price, $\Delta p = (p_2 - p_1)$, divided by the change in the per-unit tax, $\Delta t = vp_2 - 0$, collected: $\Delta p/(vp_2)$. Usually, buyers and sellers share the incidence of an ad valorem tax. Because the ad valorem tax of $v = 30\%$ has exactly the same impact on the equilibrium corn price and raises the same amount of tax per unit as the \$2.40 specific tax, the incidence is the same for both types of taxes. (As with specific taxes, the incidence of the ad valorem tax depends on the elasticities of supply and demand, but we'll spare you going through that in detail.)

Solved Problem 3.6

If the short-run supply curve for fresh fruit is perfectly inelastic and the demand curve is a downward-sloping straight line, what is the effect of an ad valorem tax on equilibrium price and quantity, and what is the incidence on consumers? Why?

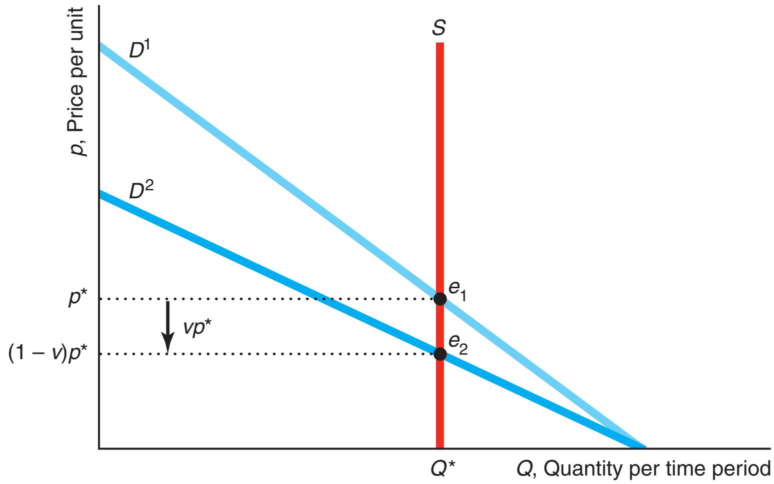
Answer

1. *Determine the before-tax equilibrium.* The perfectly inelastic supply curve, S, is vertical at Q^* in the graph. The pre-tax demand curve, D^1, intersects S at e_1, where the equilibrium price to both consumers and producers is p^* and the equilibrium quantity is Q^*.

2. *Show how the tax shifts the demand curve, and determine the after-tax equilibrium.* When the government imposes an ad valorem tax with a rate of v, the demand curve as seen by the firms rotates down to D^2, where the gap between the two demand curves is vp^*. The intersection of S and D^2 determines the after-tax equilibrium, e_2. The equilibrium quantity remains unchanged at Q^*. Consumers continue to pay p^*. The government collects vp^* per unit, so firms receive less, $(1 - v)p^*$, than the p^* they received before the tax.

3. *Determine the incidence of the tax on consumers.* The consumers continue to pay the same price, so $\Delta p = 0$ when the tax increases by vp^* (from 0), and the incidence of the tax that falls on consumers is $0/(vp^*) = 0$.

4. *Explain why the incidence of the tax falls entirely on firms.* Firms absorb the entire tax because they supply the same amount of fruit, Q^*, no matter what tax the government sets. If firms were to raise the price, consumers would buy less fruit and suppliers would be stuck with the essentially worthless excess quantity, which would spoil quickly. Thus, because suppliers prefer to sell their produce at a positive price rather than a zero price, they absorb any tax-induced drop in price.



Subsidies

A *subsidy* is a negative tax. The government takes money from firms or consumers using a tax, but gives money using a subsidy. Governments often give subsidies to firms to encourage the production of specific goods and services such as certain crops, health care, motion pictures, and clean energy.

Our analysis of the effects of taxes also applies to subsidies. Because a subsidy is a negative tax, a subsidy has the opposite effect on the equilibrium as does a tax.

For example, suppose that in the corn market the original supply curve in panel a of Figure 3.6 was S^2 and that the original equilibrium was e_2. A specific subsidy of \$2.40 per bushel would shift the supply curve down to S^1, so that the post-subsidy equilibrium would be e_1. Thus, the subsidy would *lower* the equilibrium price and *increase* the quantity.

Application

Subsidizing Ethanol

For thirty years, the U.S. government subsidized ethanol directly and indirectly with the goal of replacing 15% of U.S. gasoline use with this biofuel. The government eliminated the explicit ethanol subsidy in 2012. (However, as of 2016, the government continues to subsidize corn, the main input, and requires that gas stations sell a gasoline-ethanol mix, which greatly increases the demand for ethanol.[16])

[16]Henry Ford designed the first Model T in 1908 to run on ethanol, gasoline, or a combination of the two.

In 2011, the last year of the ethanol subsidy, the subsidy cost the government \$6 billion. According to a 2010 Rice University study, the government spent \$4 billion in 2008 to replace about 2% of the U.S. gasoline supply with ethanol, at a cost of about \$1.95 per gallon on top of the gasoline retail price. The combined ethanol and corn subsidies amounted to about \$2.59 per gallon of ethanol.

What was the subsidy's incidence on ethanol consumers? That is, how much of the subsidy went to purchasers of ethanol? Because a subsidy is a negative tax, we need to change the sign of consumer incidence formula, Equation 3.12, for a subsidy, s, rather than for a tax. That is, the consumer incidence is $\Delta p/\Delta s = \eta/(\quad - \eta)$.

According to McPhail and Babcock (2012), the supply elasticity of ethanol, η, is about 0.13, and the demand elasticity, , is about -2.1. Thus, at the equilibrium, the supply curve is relatively inelastic (nearly the opposite of the situation in Solved Problem 3.5, where the supply curve was perfectly elastic), and the demand curve is relatively elastic. Using Equation 3.12, the consumer incidence was $\eta/(\quad - \eta) = 0.13/(-2.1 - 0.13]) \approx -0.06$. In other words, almost none (6%) of the subsidy went to consumers in the form of a lower price—producers captured almost the entire subsidy. A detailed empirical study, Bielen, et al. (2016), found that consumers and corn farmers received a negligible amount of the benefits from the ethanol subsidy, with virtually all of the benefits going to ethanol producers and gasoline blenders.

Challenge Solution

Who Pays the Gasoline Tax?

MyLab Economics Solved Problem

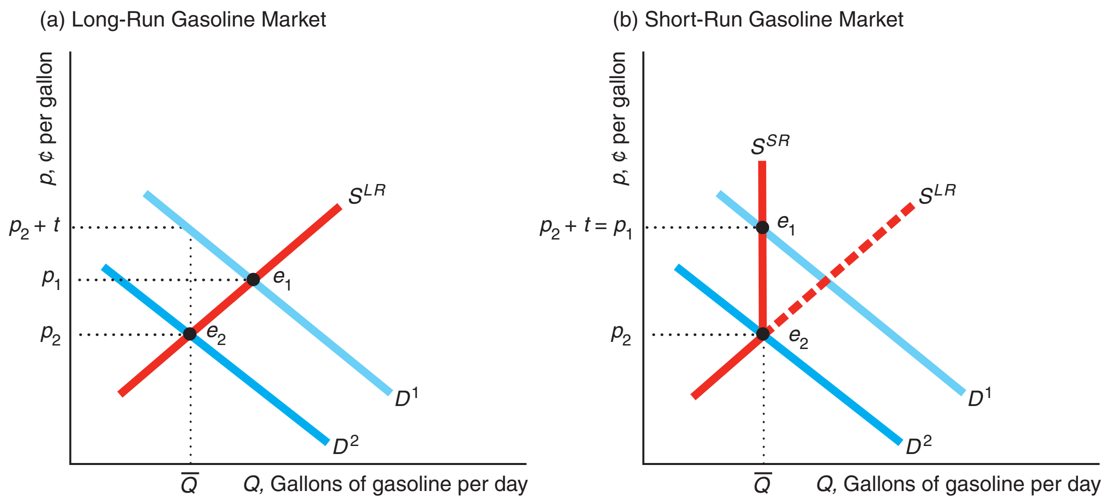
What is the long-run incidence of the federal gasoline tax on consumers? What is the short-run incidence if the government suspends the tax during summer months when gasoline prices are typically higher?

The short-run tax incidence differs from the long-run incidence, because the long-run supply curve differs substantially from the short-run curve. The long-run supply curve is upward sloping, as in our typical figure. However, the U.S. short-run supply curve is very close to vertical. The U.S. refinery capacity has fallen slightly over the last three decades. In 2016, the 140 U.S. refineries could process a maximum of 18.1 million barrels of crude oil per day, compared to 1981 when 324 refineries could process 18.6 million barrels per day. Refineries operate at almost full capacity during the summer, when the gasoline demand curve shifts to the right because families take long car trips during their vacations. Consequently, refineries cannot increase their output in the short run, and the supply curve for gasoline is nearly vertical at the maximum capacity, $\overline{Q}$. That is, even if the price of gasoline rises, producers sell no more gasoline than $\overline{Q}$.

From empirical studies, we know that the U.S. federal gasoline specific tax of $t = 18.4¢$ per gallon is shared roughly equally between gasoline companies and consumers in the long run. However, because the short-run supply curve is less elastic than the long-run supply curve, more of the tax will fall on gasoline firms in the short run (similar to Solved Problem 3.6). By the same reasoning, if the tax is suspended in the short run, more of the benefit will go to the firms than in the long run.

We contrast the long-run and short-run effects of a gasoline tax in the figure. In both panels, the specific gasoline tax, t, collected from consumers (for simplicity) causes the before-tax demand curve D^1 to shift down by t to the after-tax demand curve D^2.

In the long run in panel a, imposing the tax causes the equilibrium to shift from e_1 (intersection of D^1 and S^{LR}) to e_2 (intersection of D^2 with S^{LR}). The price that firms receive falls from p_1 to p_2, and the consumers' price goes from p_1 to $p_2 + t$. As the figure illustrates, consumers and firms share the tax roughly equally in the long run.



In the short run in panel b, the upward-sloping short-run supply curve becomes vertical at full capacity, $\overline{Q}$. The short-run equilibrium shifts from e_1 (intersection of D^1 and S^{SR}) to e_2 (intersection of D^2 with S^{SR}), so the price that consumers pay is the same before the tax, p_1, and after the tax, $p_2 + t$. The price that gasoline firms receive falls by the full amount of the tax. Thus, the gasoline firms absorb the tax in the short run but share half of it with consumers in the long run. As a result, then-Senator Obama's prediction that suspending the gas tax during the summer would primarily benefit firms and not consumers was correct.

Summary

1. **How Shapes of Supply and Demand Curves Matter.** The degree to which a shock (such as a price increase of a factor) shifts the supply curve and affects the equilibrium price and quantity depends on the shape of the demand curve. Similarly, the degree to which a shock (such as a price increase of a substitute) shifts the demand curve and affects the equilibrium depends on the shape of the supply curve.

2. **Sensitivity of the Quantity Demanded to Price.** The price elasticity of demand (or elasticity of demand), , summarizes the shape of a demand curve at a particular point. The elasticity of demand is the percentage change in the quantity demanded in response to a given percentage change in price. For example, a 1% increase in price causes the quantity demanded to fall by %. Because demand curves slope downward according to the Law of Demand, the elasticity of demand is always negative.

   The demand curve is perfectly inelastic if $= 0$, inelastic if $0 > \; > -1$, unitary elastic if $= -1$, elastic if < -1, and perfectly elastic when approaches negative infinity. A vertical demand curve is perfectly inelastic at every price. A horizontal demand curve is perfectly elastic. The income elasticity of demand is the percentage change in the quantity demanded in response to a given percentage change in income. The cross-price elasticity of demand is the percentage change in the quantity demanded of one good when the price of a related good increases by a given percentage.

   Where consumers can substitute between goods more readily in the long run, long-run demand curves are more elastic than short-run demand curves. However, if goods can be stored easily, short-run demand curves are more elastic than long-run curves.

3. **Sensitivity of the Quantity Supplied to Price.** The price elasticity of supply (or elasticity of supply), η, is the percentage change in the quantity supplied in response to a given percentage change in price. The elasticity of supply is positive if the supply curve has an upward slope. A vertical supply curve is perfectly inelastic. A horizontal supply curve is perfectly elastic. If producers can increase output at lower extra cost in the long run than in the short run, the long-run elasticity of supply is greater than the short-run elasticity.

4. **Effects of a Sales Tax.** The two common types of sales taxes are ad valorem taxes, by which the government collects a fixed percent of the price paid per unit, and specific taxes, by which the government collects a fixed amount of money per unit sold. Both types of sales taxes typically raise the equilibrium price and lower the equilibrium quantity. Both usually raise the price consumers pay and lower the price suppliers receive, so consumers do not bear the full burden or incidence of the tax. The effect of a tax on equilibrium quantities, prices, and tax revenue is unaffected by whether the tax is collected from consumers or producers. The tax incidence that falls on consumers depends on the supply and demand elasticities. Equivalent ad valorem and specific taxes have the same effect on the equilibrium. Subsidies are negative taxes, so they have the opposite effect on the equilibrium of taxes.

Questions

Select questions are available on MyLab Economics*; * = answer appears at the back of this book;* **A** *= algebra problem;* **C** *= calculus problem.*

1. How Shapes of Supply and Demand Curves Matter

1.1 Using graphs similar to those in Figure 3.1, illustrate how the effect of a demand shock depends on the shape of the supply curve. Consider supply curves that are horizontal, linear upward sloping, linear downward sloping, and vertical.

1.2 Argentina has a strong football tradition. Its national football team has appeared in five World Cup finals, and among several accolades, it has won the Copa América 14 times. The Argentine Olympic team won Olympic football tournaments in Athens (2004) and in Beijing (2008). When the national team is playing in its home stadium, many people travel to Buenos Aires to see the games, and hotel prices can rise substantially. Use a supply-and-demand diagram to explain the sharp rise in hotel room prices. In doing so, consider carefully the shape of the supply curve and what happens when all the hotel rooms are filled.

1.3 Six out of ten teens no longer use watches to tell time—they've turned to cell phones and iPods. Sales of inexpensive watches dropped 12% from 2004 to 2005, and sales of teen favorite, Fossil, Inc., fell 19%. Sales dropped 9% in 2009, during the Great Recession. However, sales rose by 9% in both 2010 and 2011 and continued to rise or stay fairly level through 2015. During all of these ups and downs, the price of inexpensive watches has not changed substantially. What can you conclude about the shape of the supply curve? Illustrate these events using a graph.

2. Sensitivity of the Quantity Demanded to Price

2.1 In a commentary piece on the rising cost of health insurance ("Healthy, Wealthy, and Wise," *Wall Street Journal,* May 4, 2004, A20), economists John Cogan, Glenn Hubbard, and Daniel Kessler state, "Each percentage-point rise in health-insurance costs increases the number of uninsured by 300,000 people." Assuming that their claim is correct, demonstrate that the price elasticity of demand for health insurance depends on the number of people who are insured. What is the price elasticity if 200 million people are insured? What is the price elasticity if 220 million people are insured? (*Hint*: See Solved Problem 3.1.) **A**

*2.2 According to Duffy-Deno (2003), when the price of broadband access capacity (the amount of information one can send over an Internet connection) increases 10%, commercial customers buy about 3.8% less capacity. What is the elasticity of demand for broadband access capacity for firms? Is demand at the current price inelastic? **A**

2.3 Gillen and Hasheminia (2013) estimate that the elasticity of demand for air travel is −0.17 for people traveling alone and −3.09 for couples. Are these demand elasticities elastic or inelastic? Which type of traveler has more demand? Why do you think these elasticities differ in this way?

2.4 Wong et al. (2013) estimated that the own-price elasticities of demand for lamb, beef, and chicken in Australia are −0.71, −0.41, and −0.28, respectively. Are the demands for these products elastic or inelastic? Which product has the least elastic demand?

2.5 What section of a straight-line demand curve is elastic?

*2.6 Use calculus to prove that the elasticity of demand is a constant everywhere along the demand curve whose demand function is $Q = Ap$. **C**

2.7 Duffy-Deno (2003) estimated that the demand function for broadband service was $Q_s = 15.6p^{-0.563}$ for small firms and $Q_l = 16.0p^{-0.296}$ for larger ones. These two demand functions cross. What can you say about the elasticities of demand on the two demand curves at the point where they cross? What can you say about the elasticities of demand more generally (at other prices)? (*Hint*: The question about the crossing point may be a red herring. Explain why.) **C**

2.8 Suppose that the demand curve for wheat in each country is inelastic up to some "choke" price p^*—a price so high that nothing is bought—so that the demand curve is vertical at Q^* at prices below p^* and horizontal at p^*. If p^* and Q^* vary across countries, what does the world's demand curve look like? Discuss how the elasticity of demand varies with price along the world's demand curve.

2.9 The British Broadcasting Corporation (BBC) conducts an annual "Price of Football" study. The 2013 version suggested that an 11% increase in the price of football tickets resulted in a drop of about 5% in average annual attendance at matches. Based on these percentages, what is the price elasticity of demand for British football? In percentage terms, by how much did the spending of fans, which equals the revenue (price times quantity) of the clubs in the Football League, change? (*Hint*: See Solved Problem 3.2.)

2.10 Ghose and Han (2014) found that the elasticity of demand for Google Play apps is −3.7. (See the Application, "The Demand Elasticities for Google Play and Apple Apps.") This elasticity applies to a small college town where approximately 1,000 apps per month are sold. If price rises by 5%, what would be the effect on quantity demanded? Would revenue rise or fall? What is the percentage change in revenue (= price × quantity)? (*Hint*: See Solved Problem 3.2.)

2.11 If the demand elasticity is −1 at the initial equilibrium and price increases by 1%, by how much does revenue change? (*Hint*: See Solved Problem 3.2.)

2.12 Konovalova and Vidishcheva (2013) report substantial negative (less than −1) cross-price elasticities of demand for travel to the United Kingdom and France by both Spanish and Italian tourists. If the city tax on hotel stays in France falls, describe how this would affect the number of Spanish tourists visiting France and the United Kingdom. How would the market for tourism in the United Kingdom be affected?

2.13 The coconut oil demand function (Buschena and Perloff, 1991) is

$$Q = 1{,}200 - 9.5p + 16.2p_p + 0.2Y,$$

where Q is the quantity of coconut oil demanded in thousands of metric tons per year, p is the price of coconut oil in cents per pound, p_p is the price of palm oil in cents per pound, and Y is the income of consumers. Assume that p is initially 45¢ per pound, p_p is 31¢ per pound, and Q is 1,275 thousand metric tons per year. Calculate the income elasticity of demand for coconut oil. (If you do not have all the numbers necessary to calculate numerical answers, write your answers in terms of variables.) **A**

*2.14 Using the coconut oil demand function from Question 2.13, calculate the price and cross-price elasticities of demand for coconut oil. **A**

2.15 In the figure in the Application "Amazon Prime," what do the areas A, $A + B$, $A + C$, and $B - C$ represent?

2.16 Suppose a seller of textile cloth in Egypt wants to raise the price of its product from $p_1 = 30$ pounds per meter to $p_2 = 36$ pounds per meter. If it is presently selling $Q_1 = 2{,}000$ meters per month and the price elasticity of demand for the product equals −0.7, would the firm's revenues rise or fall? Use a diagram of a linear demand curve to explain why. Label the various revenue areas on the diagram appropriately.

2.17 Henri Matisse was an innovative and famous French artist of the early 20th century, perhaps best known for his expressive use of color in paintings. Some of his paintings have sold for over $20 million. The supply curve for his art has been vertical, perfectly inelastic supply since his death in 1954. Suppose the price elasticity of demand for Matisse's paintings by wealthy buyers is −0.2. If 10% of the paintings are either destroyed or lost, by how much will the equilibrium price for his paintings be affected? **A**

3. Sensitivity of the Quantity Supplied to Price

*3.1 The linear supply function is Equation 3.8, $Q = g + hp$. Derive a formula for the elasticity of supply in terms of p (and not Q). Now write a formula entirely in terms of Q. **A**

3.2 Use calculus to derive the elasticity of supply if the supply function is $Q = Bp^{0.5}$. **C**

3.3 The wine-making industries of the Niagara Peninsula in Ontario, Canada, produce award-winning wines that are attracting global attention. However, it takes at least three years for new vines to produce grapes and five to six years for grapevines to begin producing a consistent, heavy crop of grapes. Discuss the relative sizes of the short-run and long-run elasticities of supply for wine from the Niagara region and some factors (land, environmental, legal) that might affect it.

3.4 While examining the impact of immigration into the United Kingdom on wages over the period 1992–2014, Nickell and Saleheen (2015) find that a 10% rise in immigrants working in semi/unskilled services—for example, in care homes, bars, shops, restaurants, and cleaning—led to a 1.88% fall in average wages. Use these results in a supply-and-demand diagram to illustrate the effects of immigration in the market for unskilled labor in the United Kingdom. In doing so, what inference can be drawn from the Nickell and Saleheen (2015) findings about the relative elasticities of demand and supply?

3.5 Solved Problem 3.3 claims that a new war in the Persian Gulf could shift the world supply curve to the left by up to 24 million barrels per day, causing the world price of oil to soar regardless of whether we drill in the Arctic National Wildlife Refuge. Use the same type of analysis as in the Solved Problem to calculate how much such a shock would cause the price to rise with and without the Arctic National Wildlife Refuge production. **A**

3.6 Australian government legislation defines fuel containing more than 1% ethanol as an ethanol-blended fuel, but it limits the amount of ethanol in petrol to 10%. Suppose the demand curve for ethanol in fuel is $Q_d = 0.75 - 0.9p$ and the supply curve is $Q_s = 0.1 + 0.3p$, where p is the price in Australian dollars per liter and the quantities are in billions of liters. How would a government-mandated cut in the amount of ethanol in petrol, reducing it to 5%, affect the price of ethanol? (*Hint*: See Solved Problem 3.3.)

4. Effects of a Sales Tax

4.1 Dan has a much higher elasticity of demand for fish than most other people. Is the incidence of a tax on fish greater for him than for other people?

4.2 An empirical study by Callison and Kaestner (2012) suggests that a 100% cigarette tax would be required to decrease adult smoking by as much as 5%. What does this result imply about the shapes of the supply and demand curves (assuming that the cigarette market is competitive)?

4.3 Governments often use a sales tax to raise *tax revenue*, which is the tax per unit times the quantity sold. All else the same, will a specific tax raise more tax revenue if the demand curve is inelastic or elastic at the original price?

4.4 To make home ownership in India more affordable, the government subsidizes interest rate payments on loans acquired by lower-income, first-time homebuyers. What effect does a per-unit subsidy for homebuyers have on the equilibrium price and quantity in the housing market? What is the incidence of the subsidy on buyers? (*Hint*: See the Application "Subsidizing Ethanol.")

4.5 Thermal power generation in Pakistan remains heavily subsidized by the government. These subsidies lower the price consumers pay to substantially less than the cost of producing electricity. What effect would a per-unit subsidy, s, have on the equilibrium price and quantity? What would be the incidence of the subsidy on consumers and producers? (*Hint*: See the Application "Subsidizing Ethanol.")

*4.6 Use math to show that, as the supply curve at the equilibrium becomes nearly perfectly elastic, the entire incidence of the tax falls on consumers. **A**

4.7 Green et al. (2005) estimate that the demand elasticity is −0.47 and the long-run supply elasticity is 12.0 for almonds. The corresponding elasticities are −0.68 and 0.73 for cotton and −0.26 and 0.64 for processing tomatoes. If the government were to apply a specific tax to each of these commodities, what incidence would fall on consumers? **A**

4.8 A constant elasticity supply curve, $Q = Bp^{\eta}$, intersects a constant elasticity demand curve, $Q = Ap$, where A, B, η, and are constants. What is the incidence of a \$1 specific tax? Does your answer depend on where the supply curve intersects the demand curve? Why? **A**

4.9 The United Kingdom has a drinking problem. British per-capita consumption of alcohol rose 19% between 1980 and 2007, compared with a 13% decline in other developed countries. Worried about excessive drinking among young people, the British government increased the tax on beer by 42% from 2008 to 2012. Under what conditions will this specific tax substantially reduce the equilibrium quantity of alcohol? Answer in terms of the elasticities of the demand and supply curves.

4.10 What is the effect of a \$1 specific tax on equilibrium price and quantity if demand is perfectly inelastic? What is the incidence on consumers? Explain. (*Hint*: See Solved Problems 3.5 and 3.6.)

4.11 What is the effect of a \$1 specific tax on equilibrium price and quantity if demand is perfectly elastic? What is the incidence on consumers? Explain. (*Hint*: See Solved Problems 3.5 and 3.6.)

4.12 What is the effect of a \$1 specific tax on equilibrium price and quantity if supply is perfectly inelastic? What is the incidence on consumers? Explain. (*Hint*: See Solved Problems 3.5 and 3.6.)

4.13 What is the effect of a \$1 specific tax on equilibrium price and quantity if demand is perfectly elastic and supply is perfectly inelastic? What is the incidence on consumers? Explain. (*Hint*: See Solved Problems 3.5 and 3.6.)

*4.14 Do you care whether a 15¢ tax per gallon of milk is collected from milk producers or from consumers at the store? Why?

4.15 If the inverse demand function is $p = a - bQ$ and the inverse supply function is $p = c + dQ$, show that the incidence of a specific tax of t per unit falling on consumers is $b/(b + d) = \eta/(\eta - \quad)$. **A**

4.16 Algebraically solve for the after-tax equilibrium price and quantity in the corn market if the government collects a specific tax of $t = \$2.40$ from customers, as panel b of Figure 3.6 illustrates. (*Hint*: See Solved Problem 3.4.) **A**

4.17 On July 1, 1965, the U.S. government eliminated its ad valorem taxes on many goods and services. By comparing the prices from before and after this change, we can determine how much the price fell in response to the tax's elimination. When the tax was in place, the tax per unit on a good that sold for p was αp. If the price fell by αp when the tax was eliminated, consumers must have been bearing the full incidence of the tax. The entire amount of the tax cut was passed on to consumers for all commodities and services Brownlee and Perry (1967) studied. Taxes had been collected at the retail level (except for motion picture admissions and club dues) and excise taxes had been imposed at the manufacturer level for most commodities, including face powder, sterling silverware, wristwatches, and handbags. List the conditions (in terms of the elasticities or shapes of supply or demand curves) that are consistent with consumers bearing the full incidence of the taxes. Use graphs to illustrate your answer.

*4.18 Essentially none of the savings from removing the federal ad valorem tax were passed on to consumers for motion picture admissions and club dues (Brownlee and Perry, 1967; see Question 4.17). List the conditions (in terms of the elasticities or shapes of supply or demand curves) that are consistent with the incidence of the taxes falling entirely on firms. Use graphs to illustrate your answer.

4.19 For a tax on sugar or fat to reduce the consumption of fattening foods and drinks by a very large amount, what must the demand and supply elasticities be? (*Hint*: See the Application "Taxes to Prevent Obesity.")

5. Challenge

5.1 The Challenge Solution says that a gas tax is roughly equally shared by consumers and firms in the long run. If so, what can you say about the elasticities of supply and demand? If in the short run the supply curve is nearly vertical, what (if anything) can you infer about the demand elasticity from observing the effect of a tax on the change in price and quantity?

5.2 Natural disasters such as earthquakes significantly reduce tourism to the area directly affected and to surrounding areas for a time. Nishimura et al. (2013) investigate the extent of the reduction in tourism, how long tourism is affected, and the areas that are affected for five recent earthquake disasters in Japan. They find that the number of tourists return to their original trends within a year. Assume that a natural disaster renders 50% of hotel rooms unusable and also causes demand to fall by 50%. Use a supply-and-demand diagram to show how this would affect the tourism demand and supply curves and the effects on equilibrium price and quantity. What would happen to the equilibrium price if demand fell by more than 50%? (*Hint*: What is the shape of the supply curve of hotel rooms when the city is at full capacity?)

Consumer Choice

4

If this is coffee, please bring me some tea; but if this is tea, please bring me some coffee.

Challenge

Why Americans Buy More Ebooks Than Do Germans

Are you reading this text electronically? Ebooks are appearing everywhere in the English-speaking world. Thanks to the popularity of the Kindle, iPad, and other ebook readers, ebooks account for about 27% of the U.S. and U.K. book markets, but only 9% in the German market according to 2016 reports.

Why are ebooks more successful in the United States than in Germany? Jürgen Harth of the German Publishers and Booksellers Association attributed the difference to tastes: what he called a "cultural issue." More than others, Germans love printed books—after all, a German invented the modern printing press. As Harth said, "On just about every corner there's a bookshop. That's the big difference between Germany and the United States."

An alternative explanation concerns government regulations and taxes that affect prices in Germany. Even if Germans and Americans have the same tastes, Americans are more likely to buy ebooks because they are less expensive than printed books in the United States. However, ebooks are more expensive than printed books in Germany. Unlike in the United States, where publishers and booksellers are free to set prices, Germany regulates book prices. To protect small booksellers, Germany's fixed-price system requires all booksellers to charge the same price for new printed books. In addition, although ebooks can sell for slightly lower prices, they are subject to a 19% tax rather than to the 7% tax that applies to printed books as of 2016.[1] So do differences in tastes account for why Germans and Americans read different types of books, or can taxes and price differences explain their behavior?

Microeconomics provides powerful insights into the myriad questions and choices facing consumers. In addition to the ebook question, we can address questions such as: How can we use information about consumers' allocations of their budgets across various goods in the past to predict how a price change will affect their demands for goods today? Are consumers better off receiving cash or a comparable amount in food stamps? Why do young people buy relatively more alcohol and less marijuana when they turn 21?

To answer questions about individual decision making, we need a model of individual behavior. Our model of consumer behavior is based on the following premises:

- Individual *tastes* or *preferences* determine the amount of pleasure people derive from the goods and services they consume.

[1] The United Kingdom has a similar difference in tax rates.

- Consumers face *constraints* or limits on their choices.
- Consumers *maximize* their well-being or pleasure from consumption, subject to the constraints they face.

Consumers spend their money on the bundle of products that give them the most pleasure. If you like music and don't have much of a sweet tooth, you spend a lot of your money on concerts and iTune songs and relatively little on candy.[2] By contrast, your chocoholic friend with a tin ear may spend a great deal on Hershey's Kisses and very little on music.

All consumers must choose which goods to buy because limits on wealth prevent them from buying everything that catches their fancy. In addition, government rules restrict what they may buy: Young consumers can't buy alcohol or cigarettes legally, and people of all ages are prohibited from buying crack and other "recreational" drugs. Therefore, consumers buy the goods that give them the most pleasure, subject to the constraints that they cannot spend more money than they have and that they cannot spend it in ways that the government prevents.

In economic analyses designed to explain behavior (positive analysis—see Chapter 1) rather than judge it (normative statements), economists assume that *the consumer is the boss*. If your brother gets pleasure from smoking, economists don't argue with him that it is bad for him any more than they'd tell your sister, who likes reading Stephen King, that she should read Adam Smith's *The Wealth of Nations* instead.[3]

Accepting each consumer's tastes is not the same as condoning the resulting behaviors. Economists want to predict behavior. They want to know, for example, whether your brother will smoke more next year if the price of cigarettes decreases 10%. The prediction is unlikely to be correct if economists say, "He shouldn't smoke; therefore, we predict he'll stop smoking next year." A prediction based on your brother's actual tastes is more likely to be correct: "Given that he likes cigarettes, he is likely to smoke more of them next year if the price falls."

In this chapter, we examine five main topics

1. **Preferences.** We use three properties of preferences to predict which combinations, or *bundle*, of goods an individual prefers to other combinations.
2. **Utility.** Economists summarize a consumer's preferences using a *utility* function, which assigns a numerical value to each possible bundle of goods, reflecting the consumer's relative ranking of these bundles.
3. **Budget Constraint.** Prices, income, and government restrictions limit a consumer's ability to make purchases by determining the rate at which a consumer can trade one good for another.
4. **Constrained Consumer Choice.** Consumers maximize their pleasure from consuming various possible bundles of goods given their income, which limits the amount of goods they can purchase.
5. **Behavioral Economics.** Experiments indicate that people sometimes deviate from rational, maximizing behavior.

---

[2]Remember that microeconomics is the study of trade-offs: Should you save your money or buy that Superman *Action Comics* Number 1 you've always wanted? Indeed, an anagram for *microeconomics* is *income or comics*.

[3]As the ancient Romans put it, "De gustibus non est disputandum"—there is no disputing about (accounting for) tastes. Or, as Joan Crawford's character said in the movie *Grand Hotel* (1932), "Have caviar if you like, but it tastes like herring to me."

4.1 Preferences

I have forced myself to contradict myself in order to avoid conforming to my own taste. —Marcel Duchamp, Dada artist

We start our analysis of consumer behavior by examining consumer preferences. Using three basic assumptions, we can make many predictions about preferences. Once we know about consumers' preferences, we can add information about the constraints consumers face so that we can answer questions, such as the one posed in the Challenge at the beginning of this chapter, or derive demand curves , as is done in the next chapter.

As a consumer, you choose among many goods. Should you have ice cream or cake for dessert? Should you spend more money on a large apartment or rent a single room and use the savings to pay for trips and concerts? In short, you allocate your money to buy a *bundle* (*market basket* or combination) of goods.

How do consumers choose the bundles of goods they buy? One possibility is that consumers behave randomly and blindly choose one good or another without any thought. However, consumers appear to make systematic choices. For example, most consumers buy very similar items each time they visit a grocery store.

To explain consumer behavior, economists *assume* that consumers have a set of tastes or preferences that they use to guide them in choosing between goods. These tastes may differ substantially among individuals.[4] Let's start by specifying the underlying assumptions in the economist's model of consumer behavior.

Mackenzie and Chase are Civilization's greatest threat.

Properties of Consumer Preferences

A consumer chooses between bundles of goods by ranking them as to the pleasure the consumer gets from consuming each. We summarize a consumer's ranking with *preference relation* symbols: weakly prefers, $\succsim$, strictly prefers, $\succ$, and indifferent between, $\sim$. If the consumer likes Bundle a at least as much as Bundle b, we say that the consumer *weakly prefers a* to b, which we write as $a \succsim b$.

Given this weak preference relation, we can derive two other relations. If the consumer weakly prefers Bundle a to b, $a \succsim b$, but the consumer does not weakly prefer b to a, then we say that the consumer *strictly prefers a* to b—would definitely choose a rather than b if given a choice—which we write as $a \succ b$.

If the consumer weakly prefers a to b and weakly prefers b to a—that is $a \succsim b$ and $b \succsim a$—then we say that the consumer is *indifferent* between the bundles, or likes the two bundles equally, which we write as $a \sim b$.

We make three assumptions about the properties of consumers' preferences. For brevity, we refer to these properties as *completeness*, *transitivity*, and *more is better*.

---

[4]Of Americans younger than 35, half the women but only a quarter of the men have tattoos. Harper's Index, *Harper's Magazine*, August 2014.

Completeness The completeness property holds that, when facing a choice between any two bundles of goods, Bundles a and b, a consumer can rank them so that one and only one of the following relationships is true: $a \succsim b$, $b \succsim a$, or both relationships hold so that $a \sim b$. The completeness property rules out the possibility that the consumer cannot rank the bundles.

Transitivity It would be very difficult to predict behavior if consumers' rankings of bundles were not logically consistent. The transitivity property eliminates the possibility of certain types of illogical behavior. According to this property, a consumer's preferences over bundles is consistent in the sense that, if the consumer *weakly prefers* a to b, $a \succsim b$, and weakly prefers b to c, $b \succsim c$, then the consumer also weakly prefers a to c, $a \succsim c$.

If your sister told you that she preferred a scoop of ice cream to a piece of cake, a piece of cake to a candy bar, and a candy bar to a scoop of ice cream, you'd probably think she'd lost her mind. At the very least, you wouldn't know which dessert to serve her.

If completeness and transitivity hold, then the preference relation $\succsim$ is said to be *rational*. That is, the consumer has well-defined preferences between any pair of alternatives.

good
a commodity for which more is preferred to less, at least at some levels of consumption

bad
something for which less is preferred to more, such as pollution

More Is Better The more-is-better property states that, all else the same, more of a commodity is better than less of it.[5] Indeed, economists define a **good** as a commodity for which more is preferred to less, at least at some levels of consumption. In contrast, a **bad** is something for which less is preferred to more, such as pollution. We now concentrate on goods (except in Chapter 18).

Although the completeness and transitivity properties are crucial to the analysis that follows, the more-is-better property is included to simplify the analysis—our most important results would follow even without this property.

So why do economists assume that the more-is-better property holds? The most compelling reason is that it appears to be true for most people.[6] Another reason is that if consumers can freely dispose of excess goods, a consumer can be no worse off with extra goods. (We examine a third reason later in the chapter: Consumers buy goods only when the more-is-better condition holds.)

Application

You Can't Have Too Much Money

"Having more money doesn't make you happier. I have 50 million dollars but I'm just as happy as when I had 48 million." —Arnold Schwarzenegger.

Surprisingly, studies of people in many nations find mixed results regarding whether richer people are happier on average than poorer people (Helliwell et al., 2012;

[5]*Jargon alert*: Economists call this property *nonsatiation* or *monotonicity*.

[6]When teaching microeconomics to Wharton MBAs, I told them about a cousin of mine who had just joined a commune in Oregon. His worldly possessions consisted of a tent, a Franklin stove, enough food to live on, and a few clothes. He said that he didn't need any other goods—that he was *satiated*. A few years later, one of these students bumped into me on the street and said, "Professor, I don't remember your name or much of anything you taught me in your course, but I can't stop thinking about your cousin. Is it really true that he doesn't want *anything* else? His very existence is a repudiation of my whole way of life." Actually, my cousin had given up his ascetic life and was engaged in telemarketing, but I, for noble pedagogical reasons, responded, "Of course he still lives that way—you can't expect everyone to have the tastes of an MBA."

Easterlin, 2015; Easterlin, 2016). Do people become satiated? Can people be so rich that they can buy everything they want and additional income does not increase their feelings of well-being? Using data from many countries, Stevenson and Wolfers (2013) found no evidence of a satiation point beyond which wealthier countries or wealthier individuals have no further increases in subjective well-being. Moreover, they found a clear positive relationship between average levels of self-reported feelings of happiness or satisfaction and income per capita within and across countries.

Less scientific, but perhaps more compelling, is a 2005 survey of wealthy U.S. citizens who were asked, "How much wealth do you need to live comfortably?" On average, those with a net worth of over $1 million said that they needed $2.4 million to live comfortably, those with at least $5 million in net worth said that they need $10.4 million, and those with at least $10 million wanted $18.1 million. Apparently, many people never have enough.

Preference Maps

Surprisingly, with just the completeness, transitivity, and more-is-better properties, we can tell a lot about a consumer's preferences. One of the simplest ways to summarize information about a consumer's preferences is to create a graphical interpretation—a map—of them. For graphical simplicity, we concentrate throughout this chapter on choices between only two goods, but the model can handle any number of goods.

Each semester, Lisa, who lives for fast food, decides how many pizzas and burritos to eat. Panel a of Figure 4.1 shows the various bundles of pizzas and burritos she might consume, with (individual-size) pizzas per semester on the horizontal axis and burritos per semester on the vertical axis.

At Bundle *e*, for example, Lisa consumes 25 pizzas and 15 burritos per semester. The more-is-better property says that Lisa prefers all the bundles that lie above and to the right (area *A*) of Bundle *e* because they contain at least as much or more of both pizzas and burritos as Bundle *e*. Thus, she prefers Bundle *f* (30 pizzas and 20 burritos), which lies in area *A*, to Bundle *e*.

By using the more-is-better property, Lisa prefers Bundle *e* to all the bundles that lie in area *B*, below and to the left of Bundle *e*, such as Bundle *d* (15 pizzas and 10 burritos). All the bundles in area *B* contain fewer pizzas or fewer burritos, or fewer of both, than does Bundle *e*.

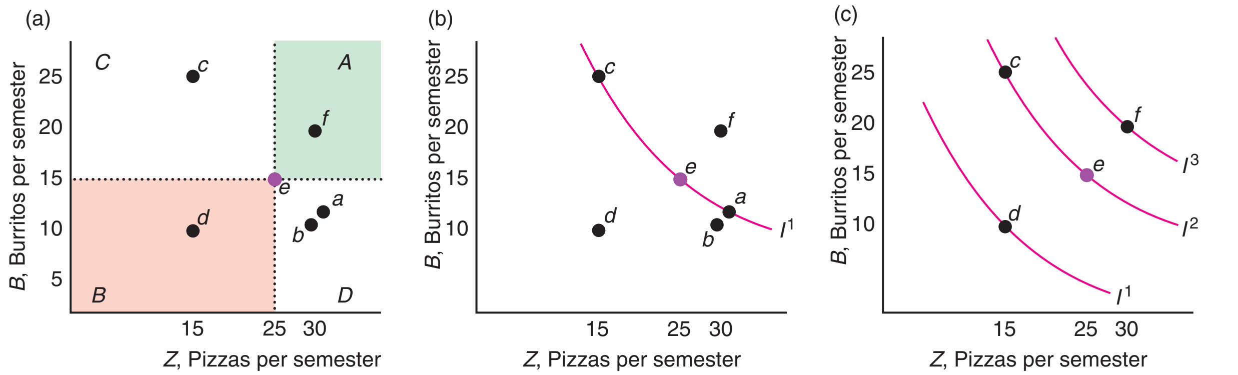
From panel a, we do not know whether Lisa prefers Bundle *e* to bundles such as *b* (30 pizzas and 10 burritos) in area *D*, which is the region below and to the right of Bundle *e*, or Bundle *c* (15 pizzas and 25 burritos) in area C, which is the region above and to the left of Bundle *e*. We can't use the more-is-better property to determine which bundle is preferred because each of these bundles contains more of one good and less of the other than *e* does. To be able to state with certainty whether Lisa prefers particular bundles in areas *C* or *D* to Bundle *e*, we have to know more about her specific tastes for pizza and burritos.

indifference curve
the set of all bundles of goods that a consumer views as being equally desirable

Indifference Curves Suppose we asked Lisa to identify all the bundles that gave her the same amount of pleasure as consuming Bundle *e*. Using her answers, we draw curve *I* in panel b of Figure 4.1 through all bundles she likes as much as *e*. Curve *I* is an **indifference curve**: the set of all bundles of goods that a consumer views as being equally desirable.

Figure 4.1 Bundles of Pizzas and Burritos Lisa Might Consume

Pizzas per semester are on the horizontal axis, and burritos per semester are on the vertical axis. (a) Lisa prefers more to less, so she prefers Bundle *e* to any bundle in area *B*, including *d*. Similarly, she prefers any bundle in area *A*, including *f*, to *e*. (b) The indifference curve, I^1,shows a set of bundles (including *c*, *e*, and *a*) among which she is indifferent. (c) The three indifference curves, I^1, I^2, and I^3, are part of Lisa's preference map, which summarizes her preferences.



Indifference curve *I* includes Bundles *c*, *e*, and *a*, so Lisa is indifferent about consuming Bundles *c*, *e*, and *a*. From this indifference curve, we also know that Lisa prefers *e* (25 pizzas and 15 burritos) to *b* (30 pizzas and 10 burritos). How do we know that? Bundle *b* lies below and to the left of Bundle *a*, so Lisa strictly prefers Bundle *a* to Bundle *b*, $a \succ b$, by the more-is-better property. Both Bundle *e* and Bundle *a* are on indifference curve *I*, so Lisa is indifferent between Bundle *e* and Bundle *a*, $e \sim a$. Because Lisa is indifferent between *e* and *a* and she prefers *a* to *b*, she must prefer *e* to *b*, $e \succ b$, by transitivity.

indifference map
a complete set of indifference curves that summarize a consumer's tastes or preferences

If we asked Lisa many, many questions about whether she prefers one bundle to another, we could draw an entire set of indifference curves through every possible bundle of burritos and pizzas.[7] Lisa's preferences are summarized in an **indifference map** or *preference map*, which is a complete set of indifference curves that summarize a consumer's tastes. We call it a *map* because it uses the same principle as a topographical or contour map, in which each line shows all points with the same height or elevation. With an indifference map, each line shows points (combinations of goods) with the same utility or well-being. Panel c of Figure 4.1 shows three of Lisa's indifference curves: I^1, I^2, and I^3.

We assume that indifference curves are continuous—have no gaps—as the figure shows. The indifference curves are parallel in the figure, but they need not be. We can demonstrate that all indifference curve maps must have the following four properties:

1. Bundles on indifference curves farther from the origin are preferred to those on indifference curves closer to the origin.
2. An indifference curve goes through every possible bundle.

[7]For example, by questioning people about which goods they would choose, Rousseas and Hart (1951) constructed indifference curves between eggs and bacon and MacCrimmon and Toda (1969) constructed indifference curves between French pastries and money (which can be used to buy all other goods).

3. Indifference curves cannot cross.
4. Indifference curves slope downward.

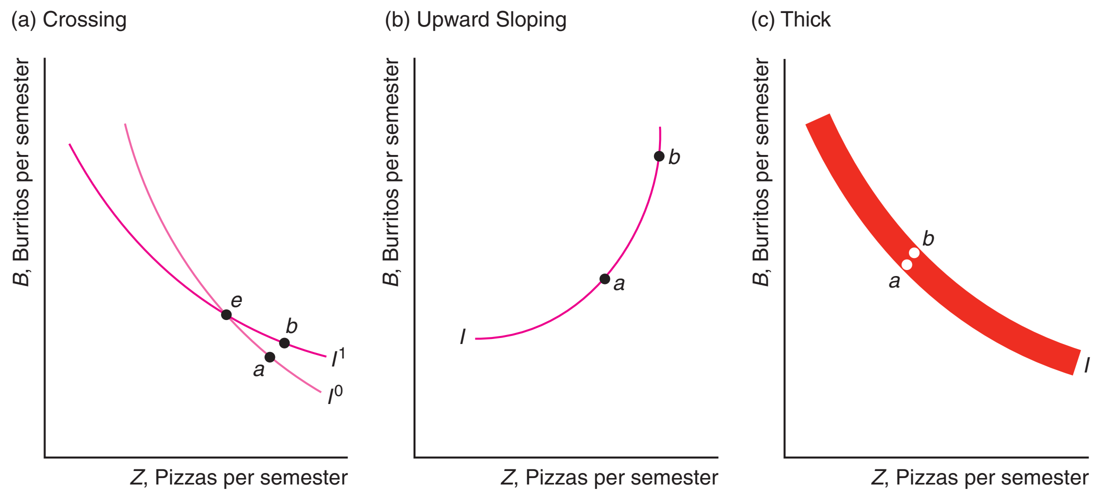
First, we show that bundles on indifference curves farther from the origin are preferred to those on indifference curves closer to the origin. By the more-is-better property, Lisa prefers Bundle *f* to Bundle *e* in panel c of Figure 4.1. She is indifferent among Bundle *f* and all the other bundles on indifference curve I^3, just as she is indifferent among all the bundles on indifference curve I^2, such as Bundles *c* and *e*. By the transitivity property, she prefers Bundle *f* to Bundle *e*, which she likes as much as Bundle *c*, so she also prefers Bundle *f* to Bundle *c*. By this type of reasoning, she prefers all bundles on I^3 to all bundles on I^2.

Second, we show that an indifference curve goes through every possible bundle using the completeness property: The consumer can compare any bundle to another. Compared to a given bundle, a consumer views some bundles as better, some as providing equal pleasure, and others as inferior. Connecting bundles that give the same pleasure as the given bundle produces an indifference curve that includes the given bundle.

Third, we show that indifference curves cannot cross: A given bundle cannot be on two indifference curves. Suppose that two indifference curves crossed at Bundle *e* as in panel a of Figure 4.2. Because Bundles *e* and *a* lie on the same indifference curve I^0, Lisa is indifferent between *e* and *a*. Similarly, she is indifferent between *e* and *b* because both are on I^1. By transitivity, if Lisa is indifferent between *e* and *a*, and she

Figure 4.2 Impossible Indifference Curves

(a) Suppose that the indifference curves cross at Bundle *e*. Lisa is indifferent between *e* and *a* on indifference curve I^0 and between *e* and *b* on I^1. If Lisa is indifferent between *e* and *a*, and she is indifferent between *e* and *b*, she must be indifferent between *a* and *b* by transitivity. But *b* has more of both pizzas and burritos than *a*, so she *must* prefer *a* to *b*. Because of this contradiction, indifference curves cannot cross. (b) Suppose that indifference curve *I* slopes upward. The consumer is indifferent between *b* and *a*, because they lie on *I*, but prefers *b* to *a* by the more-is-better assumption. Because of this contradiction, indifference curves cannot be upward sloping. (c) Suppose that indifference curve *I* is thick enough to contain both *a* and *b*. The consumer is indifferent between *a* and *b*, because both are on *I*, but prefers *b* to *a* by the more-is-better assumption, because *b* lies above and to the right of *a*. Because of this contradiction, indifference curves cannot be thick.



is indifferent between e and b, she must be indifferent between a and b. But that's impossible! Bundle b is above and to the right of Bundle a, so Lisa *must* prefer one bundle to the other. Indeed, by the more-is-better property, she prefers b to a because b is above and to the right of a. Thus, because preferences are transitive and more is better than less, indifference curves cannot cross.

Fourth, we show that indifference curves must be downward sloping. Suppose that an indifference curve sloped upward, as in panel b of Figure 4.2. The consumer is indifferent between Bundles a and b because both lie on the same indifference curve, I. But the consumer prefers b to a by the more-is-better property: Bundle a lies below and to the left of Bundle b. Because of this contradiction—the consumer cannot be indifferent between a and b and strictly prefer b to a—indifference curves cannot be upward sloping. For example, if Lisa views pizza and burritos as goods, she can't be indifferent between a bundle of one pizza and one burrito and another bundle with six of each.

Solved Problem 4.1

Can indifference curves be thick?

Answer

Draw an indifference curve that is at least two bundles thick, and show that a preference property is violated. Panel c of Figure 4.2 shows a thick indifference curve, I, with two bundles, a and b, identified. Bundle b lies above and to the right of a: Bundle b has more of both burritos and pizza. Thus, by the more-is-better property, Bundle b must be strictly preferred to Bundle a, $b \succ a$. But the consumer must be indifferent between a and b, $a \sim b$, because both bundles are on the same indifference curve. Because both relationships between a and b cannot be true, we have a contradiction. Consequently, indifference curves cannot be thick.

Comment: We illustrate this point by drawing indifference curves with very thin lines in our figures.

We are out of tickets for Swan Lake. Do you want tickets for Wrestlemania?

Willingness to Substitute Between Goods Lisa is willing to make some trades between goods. The downward slope of her indifference curves shows that Lisa is willing to give up some burritos for more pizza or vice versa. She is indifferent between Bundles a and b on her indifference curve I in panel a of Figure 4.3. If she initially has Bundle a (eight burritos and three pizzas), she could get to Bundle b (five burritos and four pizzas) by trading three burritos for one more pizza. She is indifferent whether she makes this trade or not.

Lisa's willingness to trade one good for another is measured by her **marginal rate of substitution** (***MRS***): the maximum amount of one good a consumer will sacrifice to obtain one more unit of another good. The marginal rate of substitution refers to the trade-off (rate of substitution) of burritos for a marginal (small additional or incremental) change in the number of pizzas. Lisa's marginal rate of substitution of burritos for pizza is

$$MRS = \frac{\Delta B}{\Delta Z},$$

marginal rate of substitution (*MRS*)
the maximum amount of one good a consumer will sacrifice to obtain one more unit of another good

where ΔZ is the number of pizzas Lisa will give up to get ΔB, more burritos, or vice versa, and pizza (Z) is on the horizontal axis. *The marginal rate of substitution is the slope of the indifference curve.*[8]

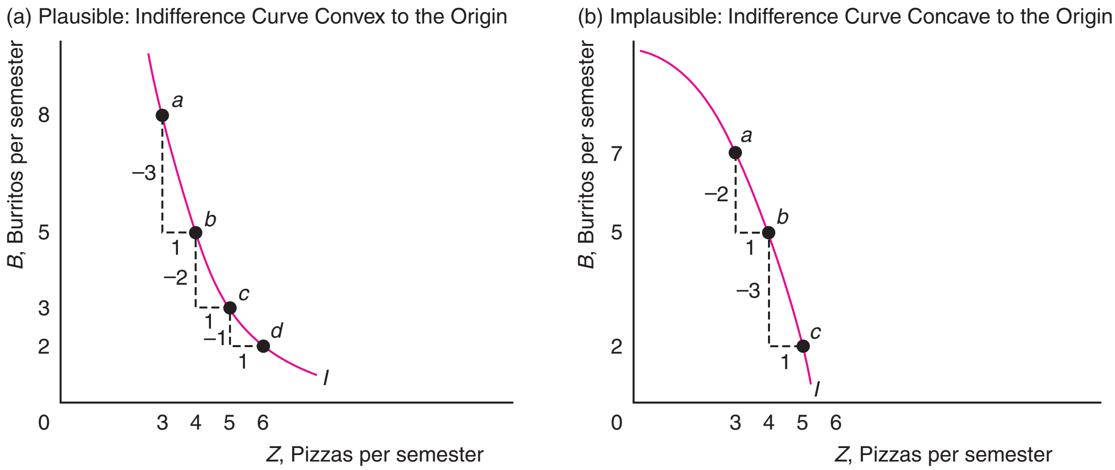
Moving from Bundle *a* to Bundle *b* in panel a of Figure 4.3, Lisa will give up three burritos, $\Delta B = -3$, to obtain one more pizza, $\Delta Z = 1$, so her marginal rate of substitution is $-3/1 = -3$. That is, the slope of the indifference curve is -3. The negative sign shows that Lisa is willing to give up some of one good to get more of the other: Her indifference curve slopes downward.

Curvature of Indifference Curves Must an indifference curve, such as *I* in panel a of Figure 4.3, be *convex* to the origin (that is, must the middle of the curve be closer to the origin than if the indifference curve were a straight line)? An indifference curve doesn't have to be convex, but casual observation suggests that most people's indifference curves are convex. When people have a lot of one good, they are willing to give up a relatively large amount of it to get a good of which they have relatively little. However, after that first trade, they are willing to give up less of the first good to get the same amount more of the second good.

Lisa is willing to give up three burritos to obtain one more pizza if she has Bundle *a* in panel a of Figure 4.3. If she has Bundle *b*, she is willing to trade only two burritos for a pizza. With Bundle *c*, she is even less willing to trade; she will

Figure 4.3 Marginal Rate of Substitution

(a) At Bundle *a*, Lisa is willing to give up three burritos for one more pizza; at *b*, she is willing to give up only two burritos to obtain another pizza. That is, the relatively more burritos she has, the more she is willing to trade for another pizza. (b) An indifference curve of this shape is unlikely to be observed. Lisa would be willing to give up more burritos to get one more pizza, the fewer the burritos she has. Moving from Bundle *c* to *b*, she will trade one pizza for three burritos, whereas moving from *b* to *a*, she will trade one pizza for two burritos, even though she now has relatively more burritos than pizzas.



[8]The *slope* is *the rise over the run*: how much we move along the vertical axis (rise) as we move along the horizontal axis (run). Technically, by the marginal rate of substitution, we mean the slope at a particular bundle. That is, we want to know what the slope is as ΔZ gets very small. In calculus terms, the relevant slope is a derivative. See Appendix 4A.

give up only one burrito for another pizza. This willingness to trade fewer burritos for one more pizza as we move down and to the right along the indifference curve reflects a *diminishing marginal rate of substitution*: The marginal rate of substitution approaches zero as we move down and to the right along an indifference curve. That is, the indifference curve becomes flatter (less sloped) as we move down and to the right.

It is hard to imagine Lisa's indifference curves as *concave*, as in panel b of Figure 4.3, rather than *convex*, as in panel a. If her indifference curve is concave, Lisa is willing to give up more burritos to get one more pizza, the fewer the burritos she has. In panel b, she trades one pizza for three burritos moving from Bundle *c* to *b*, and she trades one pizza for only two burritos moving from *b* to *a*, even though her ratio of burritos to pizza is greater. Though it is difficult to imagine concave indifference curves, two extreme versions of downward-sloping, convex indifference curves are plausible: straight-line or right-angle indifference curves.

perfect substitutes
goods that a consumer is completely indifferent as to which to consume

One extreme case is **perfect substitutes**: goods that a consumer is completely indifferent as to which to consume. Because Bill cannot taste any difference between Coca-Cola and Pepsi-Cola, he views them as perfect substitutes: He is indifferent between one additional can of Coke and one additional can of Pepsi. His indifference curves for these two goods are straight, parallel lines with a slope of -1 everywhere along the curve, as in panel a of Figure 4.4. Thus, Bill's marginal rate of substitution is -1 at every point along these indifference curves. The slope of indifference curves of perfect substitutes need not always be -1; it can be any constant rate. For example, Chris knows from reading the labels that Clorox bleach is twice as strong as a generic brand. As a result, Chris is indifferent between one cup of Clorox and two cups of the generic bleach. The slope of his indifference curve is $-\frac{1}{2}$, where the generic bleach is on the horizontal axis.[9]

perfect complements
goods that a consumer is interested in consuming only in fixed proportions

The other extreme case is **perfect complements**: goods that a consumer is interested in consuming only in fixed proportions. Maureen doesn't like pie by itself or ice cream by itself but loves pie à la mode: a slice of pie with a scoop of vanilla ice cream on top. Her indifference curves have right angles in panel b of Figure 4.4. If she has only one piece of pie, she gets the same pleasure from the pie with one scoop of ice cream, Bundle *a*, as with two scoops, Bundle *d*, or three scoops, Bundle *e*. That is, she won't eat the extra scoops because she does not have pieces of pie to go with the ice cream. Therefore, she consumes only bundles like *a*, *b*, and *c* in which pie and ice cream are in equal proportions.

With a bundle like *a*, *b*, or *c*, she will not substitute a piece of pie for an extra scoop of ice cream. For example, if she were at *b*, she would be unwilling to give up an extra slice of pie to get, say, two extra scoops of ice cream, as at point *e*. Indeed, she wouldn't give up the slice of pie for a virtually unlimited amount of extra ice cream because the extra ice cream is worthless to her.

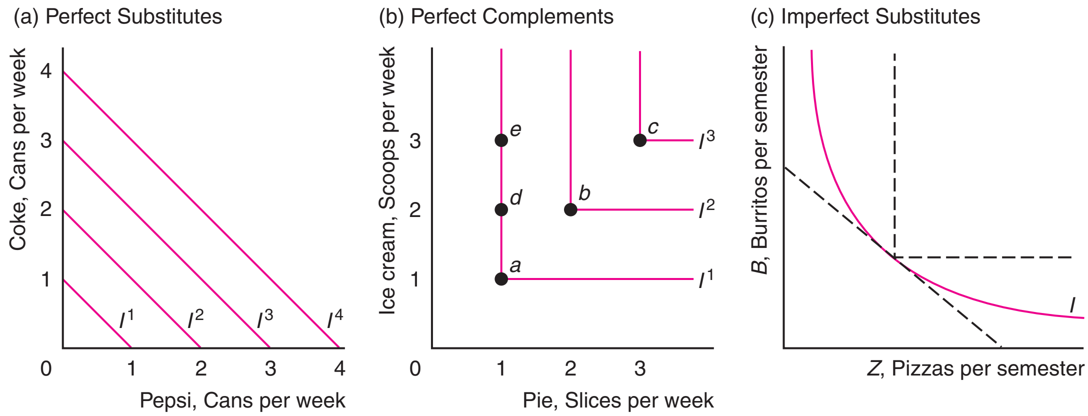
The standard-shaped, convex indifference curve in panel c of Figure 4.4 lies between these two extreme examples. Convex indifference curves show that a consumer views two goods as imperfect substitutes.

---

[9]Sometimes it is difficult to guess which goods are close substitutes. According to *Harper's* Index 1994, flowers, perfume, and fire extinguishers rank 1, 2, and 3 among Mother's Day gifts that Americans consider "very appropriate."

Figure 4.4 Perfect Substitutes, Perfect Complements, Imperfect Substitutes MyLab Economics Video

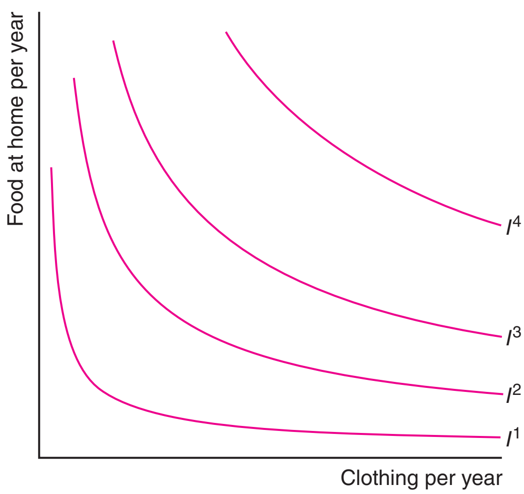
(a) Bill views Coke and Pepsi as perfect substitutes. His indifference curves are straight, parallel lines with a marginal rate of substitution (slope) of −1. Bill is willing to exchange one can of Coke for one can of Pepsi. (b) Maureen likes pie à la mode but does not like pie or ice cream by itself: She views ice cream and pie as perfect complements. She will not substitute between the two; she consumes them only in equal quantities. (c) Lisa views burritos and pizza as imperfect substitutes. Her indifference curve lies between the extreme cases of perfect substitutes and perfect complements.



Application

Indifference Curves Between Food and Clothing

The figure shows estimated indifference curves of the average U.S. consumer between food consumed at home and clothing. The food and clothing measures are weighted averages of various goods. At relatively low quantities of food and clothing, the indifference curves, such as I^1, are nearly right angles: perfect complements. As we move away from the origin, the indifference curves become flatter: closer to perfect substitutes.



One interpretation of these indifference curves is that minimum levels of food and clothing are necessary to support life. The consumer cannot trade one good for the other if it means having less than these critical levels. As the consumer obtains more of both goods, however, the consumer is increasingly willing to trade between the two goods. According to these estimates, food and clothing are perfect complements when the consumer has little of either good and perfect substitutes when the consumer has large quantities of both goods.

4.2 Utility

Underlying our model of consumer behavior is the belief that consumers can compare various bundles of goods and decide which gives them the greatest pleasure. We can summarize a consumer's preferences by assigning a numerical value to each possible bundle to reflect the consumer's relative ranking of these bundles.

utility
a set of numerical values that reflect the relative rankings of various bundles of goods

Following Jeremy Bentham, John Stuart Mill, and other nineteenth-century British economist-philosophers, economists apply the term **utility** to this set of numerical values that reflect the relative rankings of various bundles of goods. The statement that "Bonnie prefers Bundle x to Bundle y" is equivalent to the statement that "consuming Bundle x gives Bonnie more utility than consuming Bundle y." Bonnie prefers x to y if Bundle x gives Bonnie 10 *utils* (*util* is the name given to a unit of utility) and Bundle y gives her 8 utils.

Utility Function

utility function
the relationship between *utility* values and every possible bundle of goods

If we knew the **utility function**—the relationship between utility measures and every possible bundle of goods—we could summarize the information in indifference maps succinctly. Lisa's utility function, $U(Z, B)$, tells us how many utils she gets from Z pizzas and B burritos. Given that her utility function reflects her preferences, if Lisa prefers Bundle 1, (Z_1, B_1), to Bundle 2, (Z_2, B_2), then the utils she gets from the first bundle are greater that from the second bundle: $U(Z_1, B_1) > U(Z_2, B_2)$.

For example, suppose that the utility, U, that Lisa gets from burritos and pizzas is

$$U(Z,B) = \sqrt{ZB}. \tag{4.1}$$

From Equation 4.1, we know that the more she consumes of either good, the greater the utility that Lisa receives. Using this function, we can determine whether she would be happier if she had Bundle x, with 16 pizzas and 9 burritos, or Bundle y, with 13 of each. The utility she gets from x is 12 $(= \sqrt{16 \times 9})$ utils. The utility she gets from y is 13 $(= \sqrt{13 \times 13})$ utils. Therefore, she prefers y to x. The utility function is a concept that economists use to help them think about consumer behavior; utility functions do not exist in any fundamental sense.

Ordinal Preferences

Typically, consumers can easily answer questions about whether they prefer one bundle to another, such as "Do you prefer a bundle with one scoop of ice cream and two pieces of cake to another bundle with two scoops of ice cream and one piece of cake?" But they have difficulty answering questions about how much more they prefer one bundle to another because they don't have a measure to describe how their pleasure from two goods or bundles differs. Therefore, we may know a consumer's rank ordering of bundles, but we are unlikely to know by how much more that consumer prefers one bundle to another.

If we know only consumers' relative rankings of bundles, our measure of pleasure is ordinal rather than cardinal. An *ordinal* measure is a consumer's relative ranking of two goods; it does not measure the degree to which the consumer values one good over the other.

If a professor assigns only letter grades to an exam, we know that a student who receives a grade of A did better than a student who received a B, but we can't say how much better from that ordinal scale. Nor can we tell whether the difference in performance between an A student and a B student is greater or less than the difference between a B student and a C student.

A *cardinal* measure is one by which absolute comparisons between ranks may be made. Money is a cardinal measure. If you have \$100 and your brother has \$50, we know not only that you have more money than your brother but also that you have exactly twice as much money as he does.

Because utility is an ordinal measure, we should not put any weight on the absolute differences between the utility associated with one bundle and another.[10] We care only about the relative utility or ranking of the two bundles.

Utility and Indifference Curves

An indifference curve consists of all those bundles that correspond to a particular level of utility, say $\overline{U}$. If Lisa's utility function is $U(Z, B)$, then the expression for one of her indifference curves is

$$\overline{U} = U(Z, B). \tag{4.2}$$

This expression determines all those bundles of Z and B that give her $\overline{U}$ utils of pleasure. For example, if her utility function is Equation 4.1, $U = \sqrt{ZB}$, then the indifference curve $4 = \overline{U} = \sqrt{ZB}$ includes any (Z, B) bundles such that $ZB = 16$, including the bundles (4, 4), (2, 8), (8, 2), (1, 16), and (16, 1).

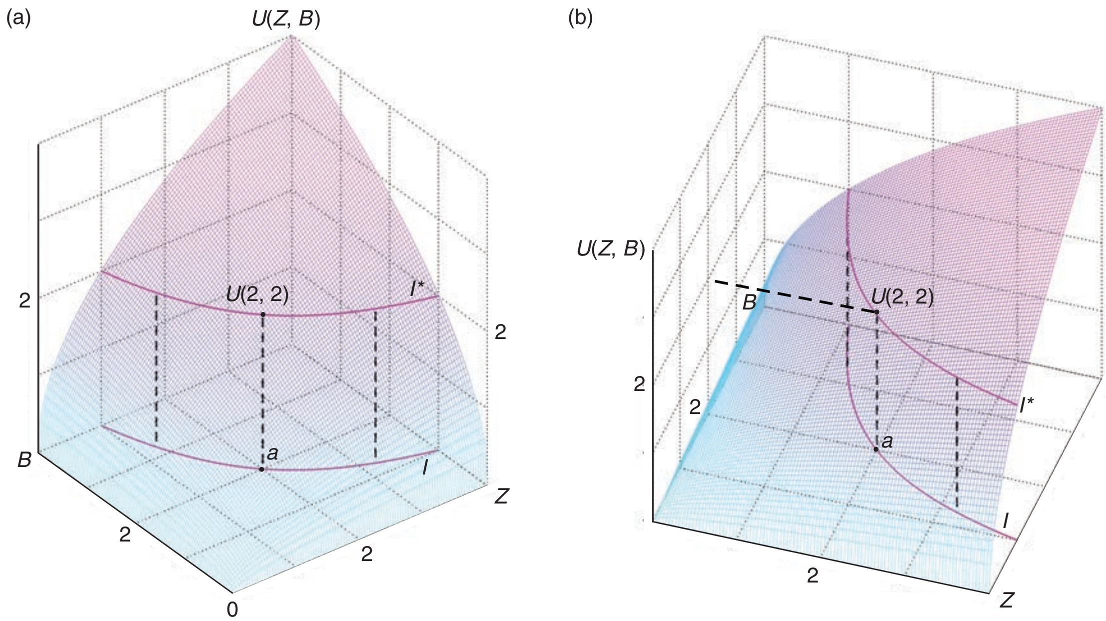
The three-dimensional diagram in Figure 4.5 shows how Lisa's utility varies with the amounts of pizza, Z, and burritos, B, that she consumes. Panel a shows this relationship from a frontal view, while panel b shows the same relationship looking at it from one side. The figure measures Z on one axis on the "floor" of the diagram, B on the other axis on the floor of the diagram, and $U(Z, B)$ on the vertical axis. For example, in the figure, Bundle *a* lies on the floor of the diagram and contains two pizzas and two burritos. Directly above it on the utility surface, or *hill of happiness*, is a point labeled $U(2, 2)$. The vertical height of this point shows how much utility Lisa gets from consuming Bundle *a*. In the figure, $U(Z, B) = \sqrt{ZB}$, so this height is $U(2, 2) = \sqrt{2 \times 2} = 2$. Because she prefers more to less, her utility rises as Z increases, B increases, or both goods increase. That is, Lisa's hill of happiness rises as she consumes more of either or both goods.

What is the relationship between Lisa's utility function and one of her indifference curves—those combinations of Z and B that give Lisa a particular level of utility? Imagine that the hill of happiness is made of clay. If you cut the hill at a particular level of utility, the height corresponding to Bundle *a*, $U(2, 2) = 2$, you get a smaller hill above the cut. The bottom edge of this hill—the edge where you cut—is the curve I^*. Now, suppose that you lower that smaller hill straight down onto the floor and trace the outside edge of this smaller hill. The outer edge of the hill on the two-dimensional floor is indifference curve I. Making other parallel cuts in the hill of happiness, placing the smaller hills on the floor, and tracing their outside edges, you can obtain a map of indifference curves on which each indifference curve reflects a different level of utility.

[10]Let $U(Z, B)$ be the original utility function and $V(Z, B)$ be the new utility function after we have applied a *positive monotonic transformation*: a change that increases the value of the function at every point. These two utility functions give the same ordinal ranking to any bundle of goods. (Economists often express this idea by saying that a *utility function is unique only up to a positive monotonic transformation*.) Suppose that $V(Z, B) = \alpha + \beta U(Z, B)$, where $\beta > 0$. The rank ordering is the same for these utility functions because $V(Z, B) = \alpha + \beta U(Z, B) > V(Z^*, B^*) = \alpha + \beta U(Z^*, B^*)$ if and only if $U(Z, B) > U(Z^*, B^*)$.

Figure 4.5 The Relationship Between the Utility Function and Indifference Curves

Panels a and b show Lisa's utility, $U(Z, B)$, from different angles as a function of the amount of pizza, Z, and burritos, B, that she consumes. Each panel measures Z along one axis on the floor of the diagram, and B along the other axis on the floor. Utility is measured on the vertical axis. As Z, B, or both increase, she has more utility: She is on a higher point on the diagram. If we project onto the floor of the diagram all the points on the curve I^* that are at a given height on the utility surface—a given level of utility—we obtain the indifference curve I.



Marginal Utility

Using Lisa's utility function for burritos and pizza, we can show how her utility changes if she gets to consume more of one of the goods. We now suppose that Lisa has the utility function in panel a of Figure 4.6. The curve shows how Lisa's utility rises as she consumes more pizzas, while we hold her consumption of burritos fixed at 10. Because pizza is a *good*, Lisa's utility rises as she consumes more pizza.

If her consumption of pizzas increases from $Z = 4$ to 5, $\Delta Z = 5 - 4 = 1$, her utility increases from $U = 230$ to 250, $\Delta U = 250 - 230 = 20$. The extra utility ($\Delta U$) that she gets from consuming the last unit of a good ($\Delta Z = 1$) is the **marginal utility** from that good. Thus, marginal utility is the slope of the utility function as we hold the quantity of the other good constant (see Appendix 4A for a calculus derivation):

marginal utility
the extra utility that a consumer gets from consuming the last unit of a good

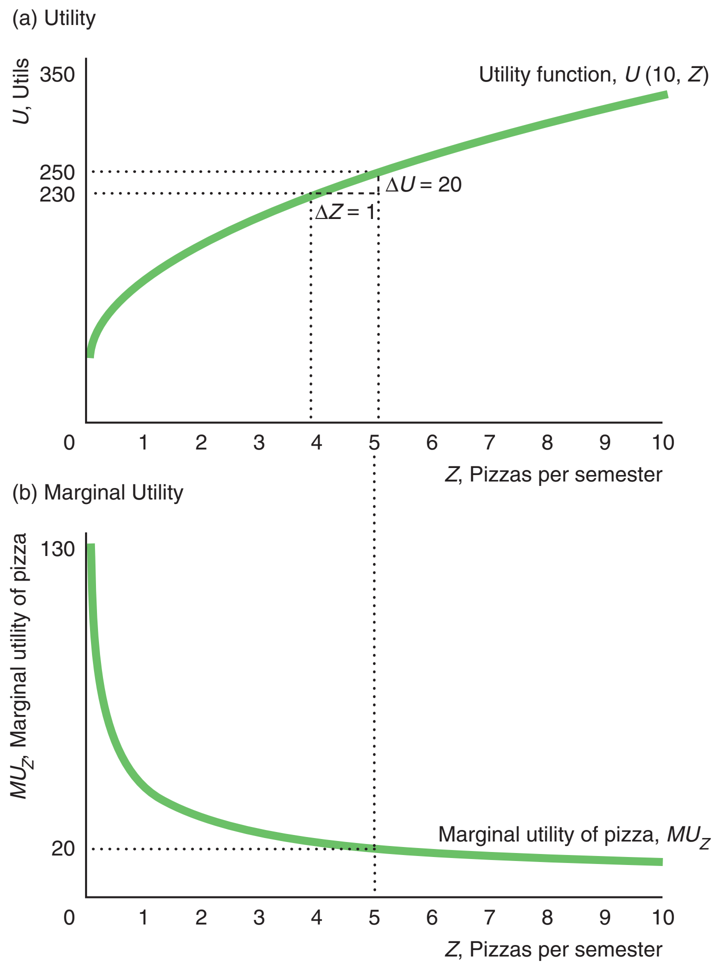
$$MU_Z = \frac{\Delta U}{\Delta Z}.$$

Lisa's marginal utility from increasing her consumption of pizza from 4 to 5 is

$$MU_Z = \frac{\Delta U}{\Delta Z} = \frac{20}{1} = 20.$$

Figure 4.6 Utility and Marginal Utility

As Lisa consumes more pizza, holding her consumption of burritos constant at 10, her total utility, U, increases and her marginal utility of pizza, MU_Z, decreases (though it remains positive).



Panel b in Figure 4.6 shows that Lisa's marginal utility from consuming one more pizza varies with the number of pizzas she consumes, holding her consumption of burritos constant. Her marginal utility of pizza curve falls as her consumption of pizza increases, but the marginal utility remains positive: Each extra pizza gives Lisa pleasure, but it gives her less pleasure than the previous pizza relative to other goods.

Utility and Marginal Rates of Substitution

Earlier we learned that the marginal rate of substitution (*MRS*) is the slope of the indifference curve. The marginal rate of substitution depends on the marginal utilities. If Lisa has 10 burritos and 4 pizzas in a semester and gets one more pizza, her utility rises. That extra utility is the marginal utility from the last pizza, MU_Z.

Similarly, if she received one extra burrito instead, her marginal utility from the last burrito is MU_B.

Suppose that Lisa moves from one bundle on an indifference curve to another by giving up some burritos to gain more pizza. She gains marginal utility from the extra pizza but loses marginal utility from fewer burritos. As Appendix 4A shows, the marginal rate of substitution can be written as

$$MRS = \frac{B}{Z} = -\frac{MU_Z}{MU_B}. \tag{4.3}$$

The *MRS* is the negative of the ratio of the marginal utility of another pizza to the marginal utility of another burrito.

4.3 Budget Constraint

Knowing an individual's preferences is only the first step in analyzing that person's consumption behavior. Consumers maximize their well-being subject to constraints. The most important constraint most of us face in deciding what to consume is our personal budget constraint.

If we cannot save and borrow, our budget is the income we receive in a given period. If we can save and borrow, we can save money early in life to consume later, such as when we retire; or we can borrow money when we are young and repay those sums later in life. Savings is, in effect, a good that consumers can buy. For simplicity, we assume that each consumer has a fixed amount of money to spend now, so we can use the terms *budget* and *income* interchangeably.

For graphical simplicity, we assume that consumers spend their money on only two goods. If Lisa spends all her budget, Y, on pizza and burritos, then

$$p_B B + p_Z Z = Y, \tag{4.4}$$

where $p_B B$ is the amount she spends on burritos and $p_Z Z$ is the amount she spends on pizzas. Equation 4.4 is her budget constraint. It shows that her expenditures on burritos and pizza use up her entire budget.

How many burritos can Lisa buy? Subtracting $p_Z Z$ from both sides of Equation 4.4 and dividing both sides by p_B, we determine the number of burritos she can purchase to be

$$B = \frac{Y}{p_B} - \frac{p_Z}{p_B}Z. \tag{4.5}$$

According to Equation 4.5, she can buy more burritos if her income increases, if the prices of burritos or pizzas fall, or if she buys fewer pizzas.[11] For example, if she has one more dollar of income (Y), she can buy $1/p_B$ more burritos.

If $p_Z = \$1$, $p_B = \$2$, and $Y = \$50$, Equation 4.5 is

$$B = \frac{\$50}{\$2} - \frac{\$1}{\$2}Z = 25 - \tfrac{1}{2}Z. \tag{4.6}$$

[11]Using calculus, we find that $\mathrm{d}B/\mathrm{d}Y = 1/p_B > 0$, $\mathrm{d}B/\mathrm{d}Z = -p_Z/p_B < 0$, $\mathrm{d}B/\mathrm{d}p_Z = -Z/p_B < 0$, and $\mathrm{d}B/\mathrm{d}p_B = -(Y - p_Z Z)/(p_B)^2 = -B/p_B < 0$.

As Equation 4.6 shows, every two pizzas cost Lisa one burrito. How many burritos can she buy if she spends all her money on burritos? By setting $Z = 0$ in Equation 4.5, we find that $B = Y/p_B = \$50/\$2 = 25$. Similarly, if she spends all her money on pizza, $B = 0$ and $Z = Y/p_Z = \$50/\$1 = 50$.

Instead of spending all her money on pizza or all on burritos, she can buy some of each. Table 4.1 shows four possible bundles she could buy. For example, she can buy 20 burritos and 10 pizzas with \$50.

budget line
the bundles of goods that can be bought if the entire budget is spent on those goods at given prices

Equation 4.6 is plotted in Figure 4.7. This line is called a **budget line** or *budget constraint*: the bundles of goods that can be bought if the entire budget is spent on those goods at given prices. This budget line shows the combinations of burritos and pizzas that Lisa can buy if she spends all of her \$50 on these two goods. The four bundles in Table 4.1 are labeled on this line.

opportunity set
all the bundles a consumer can buy, including all the bundles inside the budget constraint and on the budget constraint

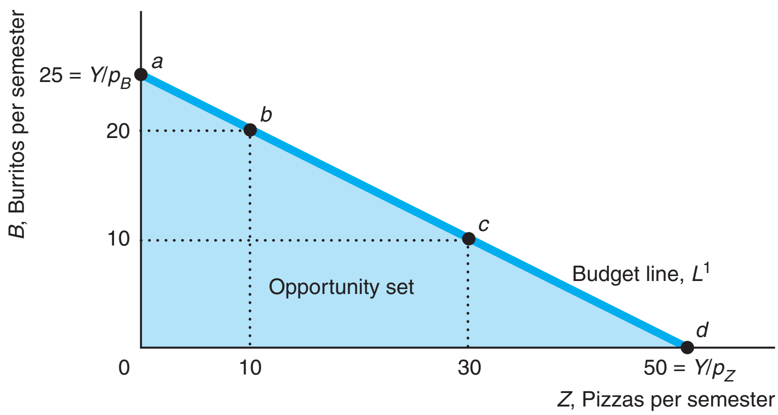
Of course, Lisa could buy any bundle that costs less than \$50. The **opportunity set** is all the bundles a consumer can buy, including all the bundles inside the budget constraint and on the budget constraint (all those bundles of positive Z and B such that $p_B B + p_Z Z \le Y$). Lisa's opportunity set is the shaded area in Figure 4.7. She could buy 10 burritos and 15 pieces of pizza for \$35, which falls inside the budget constraint. Unless she wants to spend the other \$15 on some other good, though, she might as well spend all of it on the food she loves and pick a bundle on the budget constraint rather than inside it.

Table 4.1 Allocations of a \$50 Budget Between Burritos and Pizza

Bundle	Burritos	Pizza
a	25	0
b	20	10
c	10	30
d	0	50

Figure 4.7 Budget Constraint

Lisa's budget line L^1 hits the vertical burritos axis at 25 and the horizontal pizza axis at 50 if $Y = \$50$, $p_Z = \$1$, and $p_B = \$2$. Lisa can buy any bundle in the opportunity set, the shaded area, including points on L^1. The formula for the budget line is $B = Y/p_B - (p_Z/p_B)Z = \$50/\$2 - (\$1/\$2)Z$. If Lisa buys one more unit of Z, she must reduce her consumption of B by $-(p_Z/p_B) = -\frac{1}{2}$ to stay within her budget. Thus the slope, $\Delta B/\Delta Z$, of her budget line, which is also called the marginal rate of transformation (MRT), is $-(p_Z/p_B) = -\frac{1}{2}$.



Slope of the Budget Constraint

The slope of the budget line depends on the relative prices of the two goods. Given that the budget line, Equation 4.5, is $B = Y/p_B - (p_Z/p_B)Z$, every extra unit of Z that Lisa purchases reduces B by $-p_Z/p_B$. That is, the slope of the budget line is $\Delta B/\Delta Z = -p_Z/p_B$.[12] Thus, the slope of the budget line depends on only the relative prices.

Lisa faces prices of $p_Z = \$1$ and $p_B = \$2$, so the slope of her budget line is $-p_Z/p_B = -\$1/\$2 = -\frac{1}{2}$. For example, if we reduce the number of pizzas from 10 at point b in Figure 4.7 to 0 at point a, the number of burritos that Lisa can buy rises from 20 at point b to 25 at point a, so $\Delta B/\Delta Z = (25 - 20)/(0 - 10) = 5/(-10) = -\frac{1}{2}$.[13]

marginal rate of transformation (*MRT*) the trade-off the market imposes on the consumer in terms of the amount of one good the consumer must give up to obtain more of the other good

The slope of the budget line is called the **marginal rate of transformation** (***MRT***): the trade-off the market imposes on the consumer in terms of the amount of one good the consumer must give up to purchase more of the other good:

$$MRT = \frac{B}{Z} = -\frac{p_Z}{p_B}. \tag{4.7}$$

Because Lisa's $MRT = -\frac{1}{2}$, she can "trade" an extra pizza for half a burrito; or, equivalently, she has to give up two pizzas to obtain an extra burrito.

Solved Problem 4.2

Water rationing is common during droughts. In recent years, water quotas have been imposed in areas of California, Texas, Oklahoma, Florida, the U.S. Great Plains, and the U.S. Midwest, as well as in Egypt, Honduras, India, Kenya, New Zealand, Pakistan, and Venezuela. Rationing affects consumers' opportunity sets because they cannot necessarily buy as much as they want at market prices. Suppose that a government rations water by setting a quota on how much a consumer can purchase. If a consumer can afford to buy 12 thousand gallons a week but the government restricts purchases to no more than 10 thousand gallons a week, how does the consumer's opportunity set change?

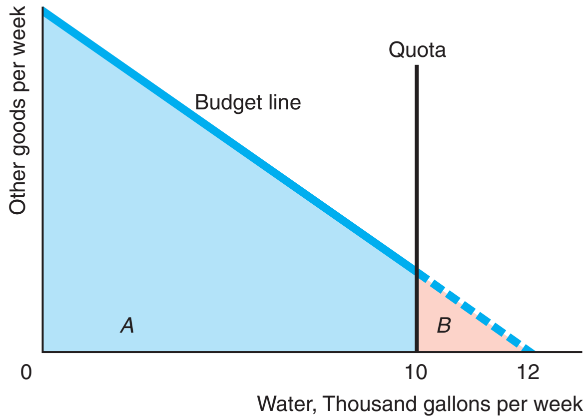
Answer

1. *Draw the original opportunity set using a budget line between water and all other goods*. In the graph, the consumer can afford to buy up to 12 thousand gallons of water per week if not constrained. The axes and the budget line bound the opportunity set, areas A and B.

2. *Add a line to the figure showing the quota, and determine the new opportunity set*. A vertical line at 10 thousand gallons on the water axis indicates the quota.

[12]As the budget line hits the horizontal axis at Y/p_Z and the vertical axis at Y/p_B, we can use the "rise over run" method to determine that the slope of the budget line is $-(Y/p_B)/(Y/p_Z) = -p_Z/p_B$. Alternatively, we can determine the slope by differentiating the budget constraint, Equation 4.5, with respect to Z: $dB/dZ = -p_Z/p_B$.

[13]The budget constraint in Figure 4.7 is a smooth, continuous line, which implies that Lisa can buy fractional numbers of burritos and pizzas. That's plausible because Lisa can buy a burrito at a *rate* of one-half per period, by buying one burrito every other week.



The axes, the budget line, and the quota line bound the new opportunity set, area *A*.

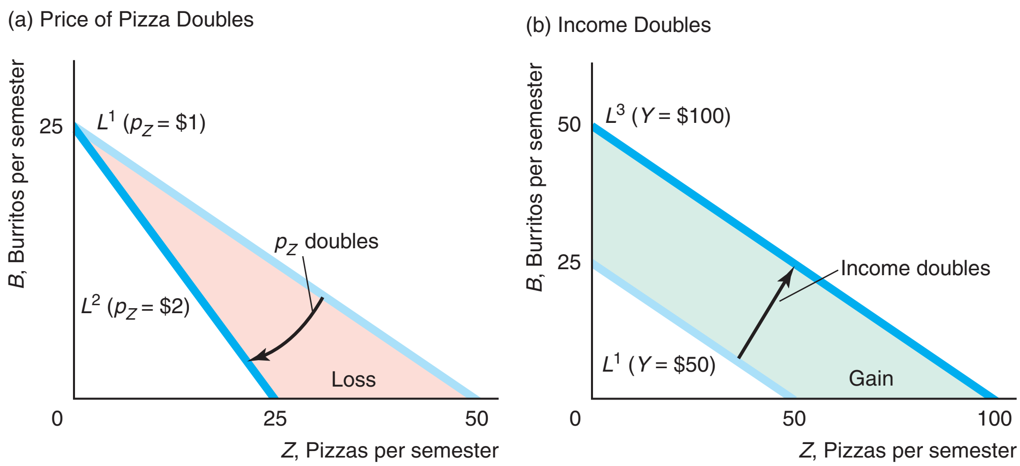
3. *Compare the two opportunity sets*. Because of the rationing, the consumer loses part of the original opportunity set: the triangle *B* to the right of the 10 thousand gallons line. The consumer has fewer opportunities because of rationing.

Effect of a Change in Price on the Opportunity Set

If the price of pizza doubles but the price of burritos is unchanged, the budget constraint swings in toward the origin in panel a of Figure 4.8. If Lisa spends all her money on burritos, she can buy as many burritos as before, so the budget line still

Figure 4.8 Changes in the Budget Constraint

(a) If the price of pizza increases from \$1 to \$2 per slice while the price of a burrito remains \$2, Lisa's budget constraint rotates from L^1 to L^2 around the intercept on the burrito axis. The slope, or *MRT*, of the original budget line, L^1, is $-\frac{1}{2}$, while the *MRT* of the new budget line, L^2, is -1. The shaded area shows the combinations of pizza and burritos that Lisa can no longer afford. (b) If Lisa's income increases by \$50 and prices don't change, her new budget constraint moves from $L^1$to L^3. This shift is parallel: Both budget lines have the same slope (*MRT*) of $-\frac{1}{2}$. The new opportunity set is larger by the shaded "Gain" area.



hits the burrito axis at 25. If she spends all her money on pizza, however, she can now buy only half as many pizzas as before, so the budget line intercepts the pizza axis at 25 instead of at 50.

The new budget constraint is steeper and lies inside the original one. As the price of pizza increases, the slope of the budget line, *MRT*, changes. On the original line, L^1, at the original prices, $MRT = -\frac{1}{2}$, which shows that Lisa could trade half a burrito for one pizza or two pizzas for one burrito. On the new line, L^2, $MRT = p_Z/p_B = -\$2/\$2 = -1$, indicating that she can now trade one burrito for one pizza, due to the increase in the price of pizza.

Unless Lisa only wants to eat burritos, she is unambiguously worse off due to this increase in the price of pizza because she can no longer afford the combinations of pizza and burritos in the shaded "Loss" area.

A decrease in the price of pizza would have the opposite effect: The budget line would rotate outward around the intercept of the line and the burrito axis. As a result, the opportunity set would increase.

Effect of a Change in Income on the Opportunity Set

If the consumer's income increases, the consumer can buy more of all goods. Suppose that Lisa's income increases by \$50 per semester to $Y = \$100$. Her budget constraint shifts outward—away from the origin—and is parallel to the original constraint in panel b of Figure 4.8. Why is the new constraint parallel to the original one? The intercept of the budget line on the burrito axis is Y/p_B, and the intercept on the pizza axis is Y/p_Z. Thus, holding prices constant, the intercepts shift outward in proportion to the change in income. Originally, if she spent all her money on pizza, Lisa could buy $50 = \$50/\$1$ pizzas; now she can buy $100 = \$100/\$1$. Similarly, the burrito axis intercept goes from $25 = \$50/\$2$ to $50 = \$100/\$2$. A change in income affects the position but not the slope of the budget line, because the relative prices of pizza and burritos determine the slope. An equal percentage decrease in the prices of both pizza and burritos has the same effect as an increase in income, as the next Solved Problem shows.

Solved Problem 4.3

Is Lisa better off if her income doubles or if the prices of both the goods she buys fall by half?

Answer

Show that her budget line and her opportunity set are the same with either change. As panel b of Figure 4.8 shows, if her income doubles, her budget line has a parallel shift outward. The new intercepts at $50 = 2Y/p_B = (2 \times 50)/2$ on the burrito axis and $100 = 2Y/p_Z = (2 \times 50)/1$ on the pizza axis are double the original values. If the prices fall by half, her budget line is the same as if her income doubles. The intercept on the burrito axis is $50 = Y/(p_B/2) = 50/(2/2)$, and the intercept on the pizza axis is $100 = Y/(p_Z/2) = 50/(1/2)$.

4.4 Constrained Consumer Choice

My problem lies in reconciling my gross habits with my net income. —Errol Flynn

Were it not for the budget constraint, consumers who prefer more to less would consume unlimited amounts of all goods. Well, they can't have it all! Instead, consumers maximize their well-being subject to their budget constraints. Now, we determine the bundle of goods that maximizes well-being subject to the budget constraint.

The Consumer's Optimal Bundle

Veni, vidi, Visa. (*We came, we saw, we went shopping.*)

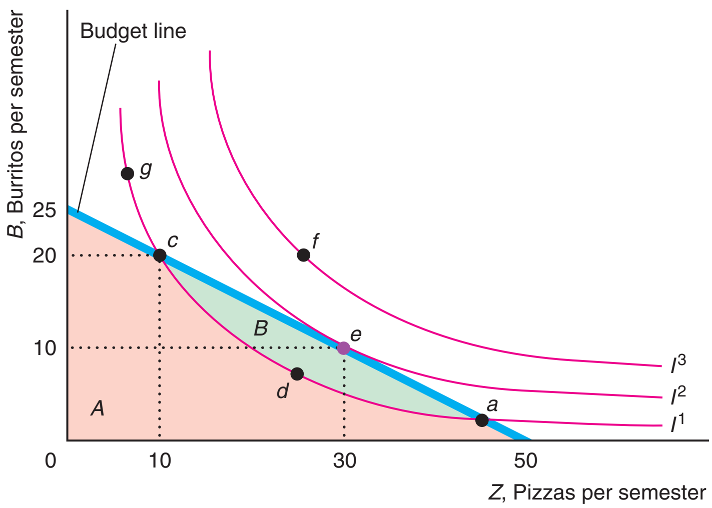
Given information about Lisa's preferences (as summarized by her indifference curves) and how much she can spend (as summarized by her budget constraint), we can determine Lisa's optimal bundle. Her optimal bundle is the bundle out of all the bundles that she can afford that gives her the most pleasure.[14]

We first show that Lisa's optimal bundle must be on the budget line in Figure 4.9. Bundles that lie on indifference curves above the constraint, such as those on I^3, are not in the opportunity set. So even though Lisa prefers *f* on indifference curve I^3 to *e* on I^2, *f* is too expensive and she can't purchase it. Although Lisa could buy a bundle inside the budget constraint, she does not want to do so, because more is better than less: For any bundle inside the constraint (such as *d* on I^1), another bundle on the constraint has more of at least one of the two goods, and hence she prefers the bundle on the constraint. Therefore, the optimal bundle must lie on the budget constraint.

We can also show that bundles that lie on indifference curves that cross the budget constraint (such as I^1, which crosses the constraint at *a* and *c*) are less desirable than certain other bundles on the constraint. Only some of the bundles on indifference curve I^1 lie within the opportunity set: Lisa can purchase Bundles *a* and *c* and all the points on I^1 between them, such as *d*. Because I^1 crosses the budget constraint, the bundles between *a* and *c* on I^1 lie inside the constraint. Some bundles in area *B* are

Figure 4.9 Consumer Maximization, Interior Solution MyLab Economics Video

Lisa's optimal bundle is *e* (10 burritos and 30 pizzas) on indifference curve I^2. Indifference curve $I^2$is tangent to her budget line at *e*. Bundle *e* is the bundle on the highest indifference curve (highest utility) that she can afford. Any bundle that is preferred to *e* (such as points on indifference curve I^3) lies outside of her opportunity set, so she cannot afford them. Bundles inside the opportunity set, such as *d*, are less desirable than *e* because they represent less of one or both goods.



[14]Appendix 4B uses calculus to determine the bundle that maximizes utility subject to the budget constraint. In this section, we use graphical techniques.

preferable to these bundles on I^1 and are affordable. For example, by the more-is-better property, Lisa prefers *e* to *d* because *e* has more of both pizza and burritos than *d*. By transitivity, *e* is preferred to *a*, *c*, and all the other points on I^1—even those, like *g*, that Lisa can't afford. Any indifference curve that crosses the budget constraint, such as I^1, has an area, such as *B*, that contains bundles that are preferred to bundles on the indifference curve. Thus, a consumer's optimal bundle—the *consumer's optimum*—must lie on the budget constraint and be on an indifference curve that does not cross it. If Lisa is consuming this optimal bundle, she has no incentive to change her behavior by substituting one good for another.

Two types of optimal bundles are possible. The first is an *interior solution*, in which the optimal bundle has positive quantities of both goods and lies between the ends of the budget line. The other possibility, called a *corner solution*, occurs when the optimal bundle is at one end of the budget line, where the budget line forms a corner with one of the axes.

Interior Solution In Figure 4.9, Bundle *e* on indifference curve I^2 is the optimal bundle. It is in the interior of the budget line away from the corners. Lisa prefers consuming a balanced diet, *e*, of 10 burritos and 30 pizzas, to eating only one type of food or the other.

For the indifference curve I^2 to touch the budget constraint but not cross it, it must be *tangent* to the budget constraint: The budget constraint and the indifference curve have the same slope at the point *e* where they touch. The slope of the indifference curve, the marginal rate of substitution, measures the rate at which Lisa is *willing* to trade burritos for pizza: $MRS = -MU_Z/MU_B$, Equation 4.3. The slope of the budget line, the marginal rate of transformation, measures the rate at which Lisa *can* trade her money for burritos or pizza in the market: $MRT = -p_Z/p_B$, Equation 4.7. Thus, Lisa maximizes her utility with the bundle where the rate at which she is willing to trade burritos for pizza equals the rate at which she can trade them:

$$MRS = -\frac{MU_Z}{MU_B} = -\frac{p_Z}{p_B} = MRT.$$

Rearranging terms, this condition is equivalent to

$$\frac{MU_Z}{p_Z} = \frac{MU_B}{p_B}. \tag{4.8}$$

Equation 4.8 says that the marginal utility of pizza divided by the price of a pizza (the amount of extra utility from pizza per dollar spent on pizza), MU_Z/p_Z, equals the marginal utility of burritos divided by the price of a burrito, MU_B/p_B. Thus, Lisa is maximizing her utility if the last dollar she spends on pizza gets her as much extra utility as the last dollar she spends on burritos. If the last dollar spent on pizza gave Lisa more extra utility than the last dollar spent on burritos, Lisa could increase her happiness by spending more on pizza and less on burritos. Her cousin Spenser is a different story.

To summarize, Lisa maximizes her utility when the following four equivalent conditions hold:

- The indifference curve between the two goods is tangent to the budget constraint.
- The consumer buys the bundle of goods that is on the highest obtainable indifference curve.
- The consumer's marginal rate of substitution (the slope of the indifference curve) equals the marginal rate of transformation (the slope of the budget line).
- The last dollar spent on Good 1 gives the consumer as much extra utility as the last dollar spent on Good 2.

Application

Substituting Alcohol for Marijuana

Crost and Guerrero (2012) found that young people view alcohol and marijuana as substitutes. They estimated that a 1 percentage point increase in the probability of using alcohol reduces the probability of using marijuana by 0.2 percentage points.

They also found that people are sensitive to changes in the relative prices. When young people turn 21 and can legally drink in the United States, their cost of buying alcohol drops (taking into account their time and the risk of being caught buying it illegally). Consequently, when they turn 21, they drink more alcohol and sharply decrease their consumption of marijuana. This effect is stronger for women, whose consumption of marijuana falls by 17%, than for men, whose consumption drops by 6%. Apparently, one unintended consequence of barring teenagers from legally drinking is that they are more likely to consume marijuana.

Solved Problem 4.4

MyLab Economics Solved Problem

Nate's utility function over jelly, J, and peanut butter, N, is $U = JN$. Nate's marginal utility from jelly is $MU_J = N$, and his marginal utility from peanut butter is $MU_N = J$.[15] The price of a jar of jelly is \$5. The price of a jar of peanut butter is \$10. Nate has a budget of \$100 to allocate to these two items. If Nate maximizes his utility, how much of each good does he consume?

Answer

1. *Derive Nate's budget line by setting his expenditure equal to his budget.* The expenditure on each item is its price times the amount consumed, so Nate's budget, 100, equals the sum of the expenditures on these two goods: $100 = 5J + 10N$.

2. *Use Equation 4.8 to find the relationship between N and J.* Equation 4.8 states that Nate maximizes his utility if he equalizes his marginal utility per dollar across jelly and peanut butter: $MU_J/5 = MU_N/10$. That is, $N/5 = J/10$ or $N = J/2$.

3. *Substitute this utility-maximizing condition into the budget equation to determine J and N.* Substituting this optimality condition into the budget constraint, we learn that $100 = 5J + 10N = 5J + 10(J/2) = 10J$. Solving this expression for J, we find that $J = 10$.

4. *Substitute the solution for J into the budget line to solve for N.* Substituting $J = 10$ into the budget constraint, we learn that $100 = 5J + 10N = 50 + 10N$, or $50 = 10N$, or $N = 5$. Thus, Nate's utility-maximizing bundle is ten jars of jelly, $J = 10$, and five jars of peanut butter, $N = 5$.

Corner Solution Some consumers choose to buy only one of the two goods: a *corner solution*. They so prefer one good to another that they only purchase the preferred good.

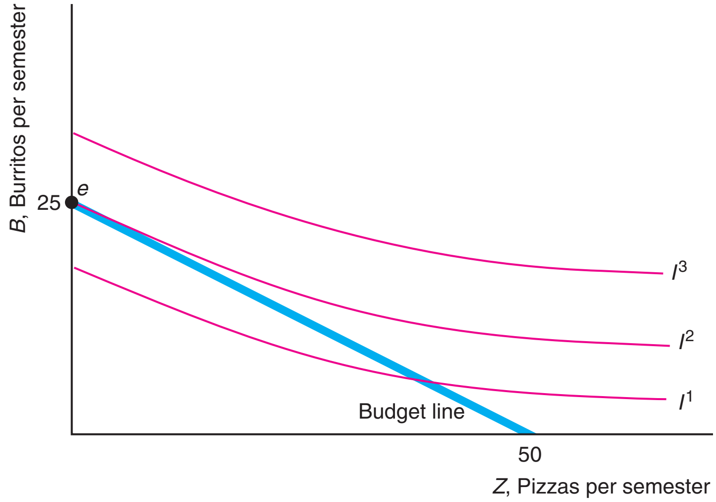
Spenser's indifference curves in Figure 4.10 are flatter than Lisa's in Figure 4.9. His optimal bundle, e, where he buys 25 burritos and no pizza, lies on an indifference curve that touches the budget line only once, at the upper-left corner.

Bundle e is the optimal bundle because the indifference curve does not cross the constraint into the opportunity set. If it did, another bundle would give Spenser more pleasure. Spenser's indifference curve is not tangent to his budget line.

[15] The marginal utility with respect to J, MU_J, is $\partial U/\partial J = \partial(JN)/\partial J = N$. Similarly, $MU_N = \partial(JN)/\partial N = J$.

Figure 4.10 Consumer Maximization, Corner Solution

Spenser's indifference curves are flatter than Lisa's indifference curves in Figure 4.9. That is, he is willing to give up more pizzas for one more burrito than is Lisa. Spenser's optimal bundle occurs at a corner of the opportunity set at Bundle *e*: 25 burritos and 0 pizzas.



Solved Problem 4.5

The concentration of the active ingredient, sodium hypochlorite, in Clorox is twice that of the generic brand. Consequently, Chris views one cup of Clorox to be a perfect substitute for two cups of the generic. If Clorox costs \$3 gallon, the generic costs \$1, and Chris allocates $Y = \$6$ per year for bleach, what bundle does Chris buy? If the price of Clorox falls to \$2, how does Chris' behavior change?

Answer

1. *Draw representative indifference curves.* Chris' indifference curves are straight lines with a slope of $-\frac{1}{2}$. For example, indifference curve I^2 hits the generic axis at 6 and the Clorox axis at 3 in both panels of the figure.

2. *Draw the initial budget line and show the bundle Chris chooses.* The initial budget line is L^1 in panel a of the figure. Chris can purchase 6 (= \$6/\$1) gallons of generic bleach, 2 (= \$6/\$3) gallons of Clorox, or any linear combination of generic and Clorox bleach that costs \$6. The highest indifference curve that

touches L^1 is I^2, which hits L^1 at Bundle e at the generic bleach axis. Because Clorox is only twice as strong as the generic but costs three times as much, Chris buys 6 gallons of the generic and no Clorox. Thus, Chris is at a corner solution.

3. *Draw the new budget line and describe how Chris' behavior changes*. When the price of Clorox falls to \$2 per gallon, the budget line is L^2 in panel b. Now the budget line L^2 and the indifference curve I^2 lie one on top of each other. Chris is indifferent between buying any bundle on I^2 including 6 gallons of the generic, 3 gallons of Clorox, 2 gallons of Clorox and 2 gallons of generic, 1 gallon of Clorox and 4 gallons of generic, and so forth. That is, Chris might buy a bundle at either corner or in the interior.

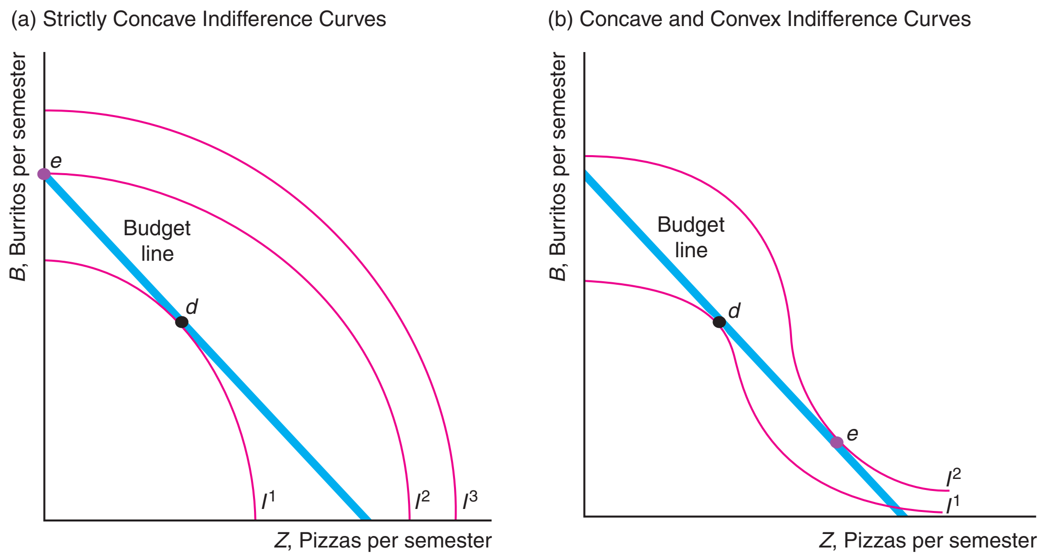
★ Optimal Bundles on Convex Sections of Indifference Curves[16]

Earlier, based on introspection, we argued that most indifference curves are convex to the origin. Now that we know how to determine a consumer's optimal bundle, we can give a more compelling explanation about why we assume that indifference curves are convex. We can show that if indifference curves are smooth, optimal bundles lie either on a convex section of an indifference curve or at a point where the budget constraint hits an axis.

Suppose that indifference curves were strictly concave to the origin, as in panel a of Figure 4.11. Indifference curve I^1 is tangent to the budget line at d, but that bundle

Figure 4.11 Optimal Bundles on Convex Sections of Indifference Curves MyLab Economics Video

(a) Indifference curve I^1 is tangent to the budget line at Bundle d, but Bundle e is superior because it lies on a higher indifference curve, I^2. If indifference curves are strictly concave to the origin, the optimal bundle, e, is always at a corner. (b) If indifference curves have both concave and convex sections, a bundle such as d, which is tangent to the budget line in the concave portion of indifference curve I^1, cannot be an optimal bundle. Here, the consumer prefers Bundle e in the convex portion of a higher indifference curve, I^2.



[16]Starred sections are optional.

is not optimal. Bundle *e* on the corner between the budget constraint and the burrito axis is on a higher indifference curve, I^2, than *d* is. Thus, if a consumer had strictly concave indifference curves, the consumer would buy only one good—here, burritos. Similarly, as we saw in Solved Problem 4.5, consumers with straight-line indifference curves buy only the least expensive good. Because we do not see consumers buying only one good, indifference curves must have convex sections.

If indifference curves have both concave and convex sections, as in panel b of Figure 4.11, the optimal bundle lies in a convex section or at a corner. Bundle *d*, where a concave section of indifference curve I^1 is tangent to the budget line, cannot be an optimal bundle. The optimum, *e*, occurs where the convex portion of a higher indifference curve, I^2, is tangent to the budget line. *If a consumer buys positive quantities of two goods, the indifference curve is convex and tangent to the budget line at the optimal bundle.*

Buying Where More Is Better

Whoever said money can't buy happiness didn't know where to shop.

A key assumption in our analysis of consumer behavior is that more is preferred to less: Consumers are not satiated. We now show that if a person consumes positive quantities of two goods and their prices are positive, more of either good must be preferred to less. Suppose that the opposite were true and that Lisa prefers fewer burritos to more. Because burritos cost her money, she can increase her well-being by reducing the number of burritos she consumes until she consumes no burritos—a scenario that violates our assumption that she consumes positive quantities of both goods.[17] Though it is possible that consumers prefer less to more at some large quantities, we do not observe consumers making purchases where that occurs.

In summary, we do not observe consumer optima at bundles where indifference curves are concave or consumers are satiated. Thus, we can safely assume that indifference curves are convex and that consumers prefer more to less in the ranges of goods that we actually observe.

Food Stamps

I've known what it is to be hungry, but I always went right to a restaurant.
—Ring Lardner

We can use the theory of consumer choice to analyze whether poor people are better off receiving food or a comparable amount of cash. Traditionally, poor U.S. households that met income, asset, and employment eligibility requirements could receive *food stamps*: coupons that recipients could use to purchase food from retail stores.

The U.S. Food Stamp Plan started in 1939. It was renamed the Food Stamp Program in 1964 and the Supplemental Nutrition Assistance Program (SNAP) in 2008. SNAP is one of the nation's largest social welfare programs, with 44 million people (one in seven U.S. residents) receiving food stamps at a cost of \$67 billion in early 2016. The average benefits were \$125 per person per month or \$4.12 per day.

In 2016, the U.S. Department of Agriculture reported that 76% of SNAP benefits go toward households with children, 12% goes to households with disabled people,

[17]Similarly, at her optimal bundle, Lisa cannot be *satiated*—indifferent between consuming more or fewer burritos. Suppose that Lisa earns her income by working and that she does not like working at the margin. Were it not for the goods she can buy with what she earns, she would not work as many hours as she does. Thus, if she were satiated and did not care if she consumed fewer burritos, she would reduce the number of hours that she works, lowering her income until she prefers more to less or she consumes none.

and 10% goes to households with senior citizens. By the time they reach 20 years of age, half of all Americans and 90% of African-American children have received food stamps at least briefly.[18]

Since the food stamp programs started, economists, nutritionists, and policymakers have debated "cashing out" food stamps by providing cash instead of coupons (or the modern equivalent, which is a debit card that recipients can use to buy only food). Legally, recipients may not sell food stamps (though a black market for them exists). Because of technological advances in electronic fund transfers, switching from food stamps to a cash program would lower administrative costs and reduce losses due to fraud and theft.

Would a switch to a comparable cash subsidy instead of food stamps increase the well-being of people who receive food assistance? Would recipients spend less on food and more on other goods?

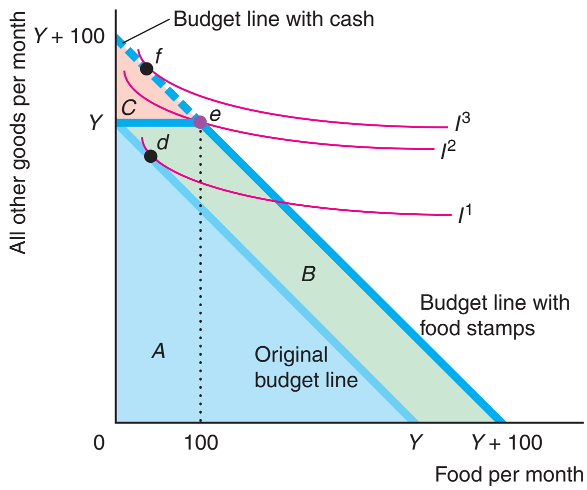
Poor people who receive cash have more choices than those who receive a comparable amount of food stamps. With cash, recipients could buy either food or other goods—not just food as with food stamps. As a result, a cash grant increases a recipient's opportunity set by more than do food stamps of the same value. Following this reasoning, many people overgeneralize about how many people would benefit from cash.

Common Confusion: Most poor people would be better off with cash than with food stamps.

People are better off with cash only if the amount of food stamps they receive exceeds what they would spend on food if they had received cash instead, which is not true for most food stamp recipients.

In Figure 4.12, one unit of food and one unit of all other goods each has a price of $1. Felicity has a monthly income of Y, so her budget line hits both axes at Y. Her opportunity set is area A.

Figure 4.12 Food Stamps Versus Cash MyLab Economics Video



The lighter line shows Felicity's original budget line for an income of Y per month. The darker solid line shows the budget constraint with $100 worth of food stamps. The budget constraint with a grant of $100 in cash is a line between $Y + 100$ on both axes. The opportunity set increases by area B with food stamps but by areas $B + C$ with cash. Given Felicity's indifference curves, she consumes Bundle d (with less than 100 units of food) with no subsidy, e (Y units of all other goods and 100 units of food) with food stamps, and f (more than Y units of all other goods and less than 100 units of food) with a cash subsidy. Felicity's utility is greater with a cash subsidy than with food stamps.

[18]According to Professor Mark Rank (Jason DeParle and Robert Gebeloff, "The Safety Net: Food Stamp Use Soars, and Stigma Fades," *New York Times*, November 29, 2009).

If Felicity receives a subsidy of \$100 in cash per month, her new monthly income is $Y + \$100$. The budget constraint with cash hits both axes at $Y + 100$ and is parallel to the original budget constraint. The opportunity set increases by $B + C$ to $A + B + C$.

If Felicity receives \$100 worth of food stamps, the food stamp budget constraint has a kink. Because recipients can spend food stamps only on food, the budget constraint shifts 100 units to the right for any quantity of other goods up to Y units. For example, if Felicity buys only food, she can now consume $Y + 100$ units of food. If she buys only other goods with the original Y income, she now can consume Y units of other goods plus 100 units of food. Because she cannot use food stamps to buy other goods, Felicity cannot buy $Y + 100$ units of other goods, as she can under a cash-transfer program. The food stamps opportunity set is area $A + B$, which is larger than the pre-subsidy opportunity set by B. The opportunity set with food stamps is smaller than with the cash-transfer program by C.

Felicity benefits as much from cash or an equivalent amount of food stamps if she would spend at least \$100 on food if given cash. In other words, she is indifferent between cash and food stamps if her indifference curve is tangent to the downward-sloping section of the food stamp budget constraint.

Conversely, if she would spend less than \$100 on food if given cash, she prefers receiving cash to food stamps. Given that she has the indifference curves in Figure 4.12, Felicity prefers cash to food stamps. She chooses Bundle e (Y units of all other goods and 100 units of food) if given food stamps but Bundle f (more than Y units of all other goods and less than 100 units of food) if given cash. She is on a higher indifference curve, I^2 rather than I^1, if given cash rather than food stamps.

Application

Benefiting from Food Stamps

Your food stamps will be stopped effective March 1992 because we received notice that you passed away. May God bless you. You may reapply if there is a change in your circumstances. —Department of Social Services, Greeneville, South Carolina

Consumer theory predicts that if the government gave food stamp recipients an equivalent amount of cash instead, their utility would remain the same or rise and *some* recipients would consume less food and more of other goods.

Whitmore (2002) estimated that between 20% and 30% of food stamp recipients would be better off if they were given cash instead of an equivalent value in food stamps. They would spend less on food than their food stamp benefit amount if they received cash instead of stamps, and therefore would be better off with cash. Of those who would trade their food stamps for cash, the average food stamp recipient values the stamps at 80% of their face value (although the average price on the underground economy is only 65%). Thus, Whitmore estimated that across all such recipients, the government wasted \$500 million by giving food stamps rather than cash.

Hoynes and Schanzenbach (2009) found that food stamps result in a decrease in out-of-pocket expenditures on food and an increase in overall food expenditures. For those households that would prefer cash to food stamps—those that spend relatively little of their income on food—food stamps cause them to increase their food consumption by about 22%, compared to 15% for other recipients, and 18% overall. Bruich (2014) estimated that each extra \$1 of SNAP leads to 37¢ more grocery store spending. However, MacEwan, Smith, and Alston (2016) concluded that participating in the food stamp program does not cause weight gain.

4.5 Behavioral Economics

He who has choice has trouble. —Dutch proverb

behavioral economics by adding insights from psychology and empirical research on human cognition and emotional biases to the rational economic model, economists try to better predict economic decision making

So far, we have assumed that consumers are rational, maximizing individuals. A new field of study, **behavioral economics**, adds insights from psychology and empirical research on human cognition and emotional biases to the rational economic model to better predict economic decision making.[19] We discuss three applications of behavioral economics in this section: tests of transitivity, the endowment effect, and salience and bounded rationality. Later in the book, we examine whether a consumer is influenced by the purchasing behavior of others (Chapter 11), why many people lack self-control (Chapter 16), and the psychology of decision making under uncertainty (Chapter 17).

Tests of Transitivity

In our presentation of the basic consumer choice model at the beginning of this chapter, we assumed that consumers make transitive choices. But, do they actually make transitive choices?

A number of studies of animals and humans show that preferences usually are transitive. Monteiro et al. (2013) report that even starlings (birds) have transitive preferences.

Weinstein (1968) used an experiment to determine how frequently people give intransitive responses. Subjects could choose between ten goods, offered in pairs, in every possible combination. To ensure that monetary value would not affect their calculations, they were told that all of the goods had a value of $3. (None of the subjects knew the purpose of the experiment.) Weinstein found that 93.5% of the responses of adults—people over 18 years old—were transitive. However, only 79.2% of children aged 9–12 gave transitive responses.

Psychologists have also tested for transitivity using preferences for colors, photos of faces, and so forth. Bradbury and Ross (1990) found that, given a choice of three colors, nearly half of 4–5 year olds are intransitive, compared to 15% for 11–13-year-olds, and 5% for adults. Bradbury and Ross showed that novelty (a preference for a new color) is responsible for most intransitive responses, and that this effect is especially strong in children.

Based on these results, one might conclude that it is appropriate to assume that adults exhibit transitivity for most economic decisions. On the other hand, one might modify the theory for children or consumers who first see novel goods.

Economists normally argue that government should allow rational people to make their own consumption choices so that they may maximize their well-being. However, some might conclude that children's lack of transitivity or rationality provides a justification for political and economic restrictions and protections placed on young people.

Endowment Effect

endowment effect people place a higher value on a good if they own it than they do if they are considering buying it

Experiments show that people have a tendency to stick with the bundle of goods that they currently possess. One important reason for this tendency is the **endowment effect**, which occurs when people place a higher value on a good if they own it than they do if they are considering buying it.

[19] The introductory chapter of Camerer et al. (2004) and DellaVigna (2009) are excellent surveys of the major papers in this field and heavily influenced the following discussion.

We normally assume that an individual can buy or sell goods at the market price. Rather than relying on income to buy some mix of two goods, an individual who has an *endowment* of several units of one good could sell some and use that money to buy units of another good.

We assume that a consumer's endowment does not affect the indifference curve map. In a classic buying and selling experiment, Kahneman et al. (1990) challenged this assumption. In an undergraduate law and economics class at Cornell University, the experimenters divided 44 students randomly into two groups. The experiments gave members of one group coffee mugs that were available at the student store for $6. They told the students endowed with a mug that they could sell their mugs. The experimenters asked the students what was the lowest price that they would accept for the mug. The experimenters also asked the students in the other group who did not receive a mug how much they would pay to buy the mug. Given the standard assumptions of our model and that the subjects were chosen randomly, we would expect no difference between the selling and buying prices. However, the median selling price was $5.75 and the median buying price was $2.25, so sellers wanted more than twice what buyers would pay.

Researchers have conducted many variations of this experiment and usually have found an endowment effect. However, Plott and Zeiler (2005) argued that if you take adequate care to train the subjects in the procedures and make sure they understand them, that subjects no longer exhibit an endowment effect. List (2003) examined the behavior of sports memorabilia collectors and found that amateurs who do not trade frequently exhibited an endowment effect, unlike professionals and amateurs who traded a lot. Thus, experience may minimize or eliminate the endowment effect, and people who buy goods for resale may be less likely to form an attachment to these goods.

Others accept the results and have considered how to modify the standard model to reflect the endowment effect (Knetsch, 1992). One implication of these experimental results is that people will only trade away from their endowments if prices change substantially. We can capture this resistance to trade with a kink in the indifference curve at the endowment bundle. (We showed indifference curves with a 90° kink in panel b of Figure 4.4.) Such indifference curves could have an angle greater than 90°, and an indifference curve could be curved at points other than at the kink. If the indifference curve has a kink, the consumer does not shift to a new bundle in response to a small price change, but may shift if the price change is large.

Application

Opt In Versus Opt Out

One practical implication of the endowment effect is that consumers' behavior may differ depending on how a choice is posed. Many firms offer their workers a choice of enrolling in the firm's voluntary tax-deferred retirement (pension) plan, called a 401(k) plan. The firm can pose the choice in two ways: It can automatically sign up employees for the program and let them opt out if they want, or it can tell employees that to participate in the program they must sign up (opt in) to participate.

These two approaches might seem identical, but the behaviors they lead to are not. Madrian and Shea (2001, 2002) found that well over twice as many workers participate if they are automatically enrolled (but may opt out) than if they must opt in: 86% versus 37%. In short, inertia matters.

Because of this type of evidence, Congress changed the federal law in 2006 and 2007 to make it easier for employers to enroll their employees in their 401(k) plans automatically. According to Aon Hewitt, the share of large firms that automatically enroll new hires in 401(k) plans was 67% in 2012, up from 58% in 2007. According to the U.S. Bureau of Labor Statistics, the enrollment rate in defined-contribution retirement plans in firms with over 500 employees rose from 45% in 2008 to 72% in 2015, due to the increased use of automatic enrollment.

Salience and Bounded Rationality

Up to now in this chapter, we have assumed that consumers know their own income or endowment, the relevant prices, and their own tastes, and hence they make informed decisions. Those assumptions are not always valid. Behavioral economists and psychologists have demonstrated that people are more likely to consider information if it is presented in a way that grabs their attention or if it takes relatively little thought or calculation to understand. Economists use the term *salience*, in the sense of *striking* or *obvious*, to describe this idea. This issue is particularly important in predicting the effects of sales taxes.

Common Confusion People necessarily buy less when the sales tax rises.

tax salience
awareness of a tax

bounded rationality
people have a limited capacity to anticipate, solve complex problems, or enumerate all options

The belief that people will buy less with a higher sales tax seems reasonable given the Law of Demand: the higher the price, the less people demand. However, for this reasoning to be true, people must have **tax salience**: awareness of a tax. If a store posts only the pre-tax price of a good, some consumers may ignore the tax when making decisions (see the next Application, "Unaware of Taxes").

An alternative explanation for ignoring taxes is **bounded rationality**: people have a limited capacity to anticipate, solve complex problems, or enumerate all options. To avoid having to perform hundreds of calculations when making purchasing decisions at a grocery store, many people choose not to calculate the tax-inclusive price. However, when that post-tax price information is easily available to them, consumers make use of it.

One way to modify the standard model is to assume that people incur a cost when making calculations—such as the time taken or the mental strain—and that deciding whether to incur this cost is part of their rational decision-making process. People incur this calculation cost only if they think the gain from a better choice of goods exceeds the calculation cost. More people pay attention to a tax when the tax rate is high or when their demand for the good is elastic (they are sensitive to price). Similarly, some people are more likely to pay attention to taxes when making large, one-time purchases—such as for a computer or car—rather than small, repeated purchases—such as for a bar of soap.

Tax salience and bounded rationality have important implications for tax policy. In Chapter 3, where we assumed that consumers pay attention to prices and taxes, we showed that the tax incidence on consumers is the same regardless of whether the tax is collected from consumers or sellers. However, if consumers are inattentive to taxes, they're more likely to bear the tax burden if they're taxed. If a tax on consumers rises and consumers don't notice, their demand for the good becomes relatively inelastic, causing consumers to bear more of the tax incidence (see Equation 3.12). In contrast, if the government taxes sellers and the sellers want to pass at least some of the tax on to consumers, they raise their price, which consumers observe.

Application

Unaware of Taxes

If a grocery store's posted price includes the sales tax, consumers observe a change in the price as the tax rises. On the other hand, if a store posts the pre-tax price and collects the tax at the cash register, consumers are less likely to note that the post-tax price has increased when the tax rate increases.

Chetty et al. (2009) compared consumers' response to a rise in an ad valorem sales tax on beer, which is included in the posted price of beer, to an increase in a general ad valorem sales tax on beer and other goods, which is collected at the cash register but not included in the posted price. An increase in either tax has the same effect on the final price, so an increase in either tax should have the same effect on purchases if consumers pay attention to both taxes.[20]

Chetty et al. found that a 10% increase in the posted price, which includes the excise tax, reduces beer consumption by 9%, while a 10% increase in the price due to a rise in a non-posted sales tax reduces consumption by only 2%. They also conducted an experiment where they posted tax-inclusive prices for 750 products in a grocery store and found that demand for these products fell by about 8% relative to control products in that store and comparable products at nearby stores.

Some consumers may ignore the tax because they forgot about it—salience—while others may not know that a tax applies at all. In grocery stores in some states, taxes may apply to alcohol and toiletries but not to food. Consequently, someone buying toothpaste may not realize that it is taxed. Zheng et al. (2013) used the finding by Chetty et al. that 20% of shoppers mistakenly think that toothpaste is untaxed to calculate that the information effect explains 31% of the sales drop in the Chetty et al. study when taxes were included in the posted price.

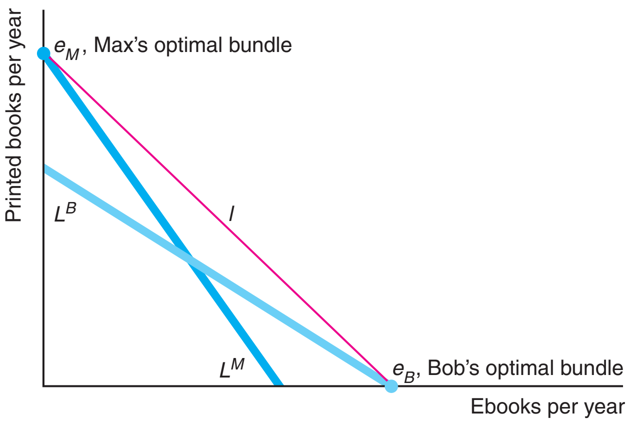
Goldin and Homonoff (2013) found that all customers respond to changes in cigarettes' posted price, but that only low-income consumers respond to taxes levied at the register. Thus, a tax collected at the register is less effective in discouraging smoking than if everyone paid attention to the tax. However, such a tax is progressive, falling more heavily on rich people who are more likely to ignore it.

Challenge Solution

Why Americans Buy More Ebooks Than Do Germans

MyLab Economics Solved Problem

Why do Germans largely ignore ebooks, while many Americans are quickly switching to this technology? While it's possible that this difference is due to different tastes in the two countries, some evidence indicates that attitudes toward ebooks are similar in the two countries. For example, according to surveys, 59% of Americans and 56% of Germans report that they have no interest in ebooks. Price differences provide a better explanation.



Suppose that Max, a German, and Bob, a Yank, are avid readers with identical incomes and tastes. Each is indifferent between reading a novel in a traditional book and using an e-reader so that their indifference curves have a slope of −1, as the red line in the figure illustrates. We can use an indifference curve–budget

[20]The final price consumers pay is $p^* = p(1 + V)(1 + v)$, where p is the pre-tax price, v is the general ad valorem sales tax rate, and V is the tax rate on beer.

line analysis to explain why Max buys printed books while Bob chooses electronic ones.

In both countries, the pre-tax price of ebooks is lower than that of printed books. In the United States, the after-tax price of ebooks remains lower, so Bob's budget line L^B is flatter than his indifference curve. However, because the German tax system sets a lower tax rate for printed books than for ebooks, the after-tax price of ebooks is higher in Germany, so Max's budget line L^M is steeper than his indifference curve. Thus, as the figure shows, Bob maximizes his utility by spending his entire book budget on ebooks. He chooses the Bundle e_B, where his indifference curve I hits his budget line L^B on the ebook axis. In contrast, Max spends his entire book budget on printed books, at point e_M.

If Bob and Max viewed the two types of books as imperfect substitutes and had the usual convex indifference curves, they would each buy a mix of ebooks and printed books. However, because of the relatively lower price of ebooks in the United States, Bob would buy relatively more ebooks.

Summary

Consumers maximize their utility (well-being) subject to constraints based on their income and the prices of goods.

1. **Preferences.** To predict consumers' responses to changes in constraints, economists use a theory about individuals' preferences. One way of summarizing consumers' preferences is with a map of indifference curves. An indifference curve consists of all bundles of goods that give the consumer a particular level of utility. Based on observations of consumers' behavior, economists assume that consumers' preferences have three properties: completeness, transitivity, and more is better. Given these three assumptions, indifference curves have the following properties:
   - Consumers get more pleasure from bundles on indifference farther from the origin.
   - An indifference curve goes through any given bundle.
   - Indifference curves cannot cross.
   - Indifference curves slope downward.
   - Indifference curves are thin.

2. **Utility.** Economists call the set of numerical values that reflect the relative rankings of bundles of goods *utility*. Utility is an ordinal measure: By comparing the utility a consumer gets from each of two bundles, we know that the consumer prefers the bundle with the higher utility, but we can't tell by how much the consumer prefers that bundle. The marginal utility from a good is the extra utility a person gets from consuming one more unit of that good, holding the consumption of all other goods constant. The rate at which a consumer is willing to substitute Good 1 for Good 2, the marginal rate of substitution, *MRS*, depends on the relative amounts of marginal utility the consumer gets from each of the two goods.

3. **Budget Constraint.** The amount of goods consumers can buy at given prices is limited by their income. As a result, the greater their income and the lower the prices of goods, the better off they are. The rate at which they can exchange Good 1 for Good 2 in the market, the marginal rate of transformation, *MRT*, depends on the relative prices of the two goods.

4. **Constrained Consumer Choice.** Each person picks an affordable bundle of goods to consume so as to maximize his or her pleasure. If an individual consumes both Good 1 and Good 2 (an interior solution), utility is maximized when the following four equivalent conditions hold:
   - The indifference curve between the two goods is tangent to the budget constraint.
   - The consumer buys the bundle of goods that is on the highest obtainable indifference curve.
   - The consumer's marginal rate of substitution (the slope of the indifference curve) equals the marginal rate of transformation (the slope of the budget line).
   - The last dollar spent on Good 1 gives the consumer as much extra utility as the last dollar spent on Good 2.

   However, consumers may not buy some of all possible goods (corner solutions). The last dollar spent on a good that is actually purchased gives more extra utility than would a dollar's worth of a good the consumer chose not to buy.

5. **Behavioral Economics.** Using insights from psychology and empirical research on human cognition and emotional biases, economists are starting to modify the rational economic model to better predict economic decision making. While adults tend to make transitive choices, children are less likely to do so, especially when novelty is involved. Consequently, some would

argue that children's ability to make economic choices should be limited. If consumers have an endowment effect, such that they place a higher value on a good if they own it than they do if they are considering buying it, they are less sensitive to price changes and hence less likely to trade than would be predicted by the standard economic model. Consumers sometimes ignore sales taxes when making purchasing decisions.

Questions

Select questions are available on MyLab Economics; ** = answer appears at the back of this book;* **A** *= algebra problem;* **C** *= calculus problem.*

1. Preferences

1.1 Give as many reasons as you can why we believe that economists assume that the more-is-better property holds and describe how these explanations relate to the results in the Application "You Can't Have Too Much Money."

1.2 Can an indifference curve be downward sloping in one section, but then bend backward so that it forms a "hook" at the end of the indifference curve? (*Hint*: See Solved Problem 4.1.)

1.3 Give as many reasons as you can why we believe that indifference curves are convex and explain.

1.4 Don is altruistic. Show the possible shape of his indifference curves between charity and all other goods.

*1.5 Arthur spends his income on bread and chocolate. He views chocolate as a good but is neutral about bread, in that he doesn't care if he consumes it or not. Draw his indifference curve map.

1.6 Which of the following pairs of goods are complements and which are substitutes? Are the goods that are complements likely to be perfect complements for some or all consumers?

a. Cars and tires

b. Apple cider and hot chocolate

c. Printers and ink cartridges

d. Soybeans and chickpeas

2. Utility

2.1 Divit enjoys watching team sports and considers tickets to kabaddi and badminton games to be perfect substitutes. Show his preference map. What is his utility function?

*2.2 Whenever Aliza buys a flashlight, she also always buys exactly two batteries along with it. Any more than two batteries will serve no purpose, because she will not be able to use them in the flashlight. Any more flashlights will not be useful either, because there will not be enough batteries for her to make use of them. Show her preference map. What is her utility function?

2.3 Does the utility function $V(Z, B) = \alpha + [U(Z, B)]^2$ give the same ordering of bundles as does $U(Z, B)$? **A**

2.4 Fiona requires a minimum level of consumption, a *threshold*, to derive additional utility: $U(X, Z)$ is 0 if $X + Z \leq 5$ and is $X + Z$ otherwise. Draw Fiona's indifference curves. Which of our usual assumptions about preferences are violated by this example?

2.5 If Joe views two candy bars and one piece of cake as perfect substitutes, what is his marginal rate of substitution between candy bars and cake?

2.6 Daniela enjoys eating cabbage rolls and often has them with her meals. She also likes polenta, and while making it, she uses the recipe that her grandmother and mother used. Daniela views cabbage rolls and polenta as imperfect substitutes, so her indifference curves for the two goods are convex. If she consumes 6 slices of polenta and 5 cabbage rolls per week, and her marginal rate of substitution of polenta for cabbage rolls is $MRS = \Delta P/\Delta C = -2$, will she prefer a bundle with 4 slices of polenta and 6 cabbage rolls? Why? **A**

*2.7 José Maria's utility function is $U(B, Z) = AB^{\alpha}Z^{\beta}$. What is his marginal utility of B? What is his marginal utility of Z? What is his marginal rate of substitution between B and Z? **C**

3. Budget Constraint

3.1 Suppose Gregg consumes chocolate candy bars and oranges. He is given four chocolate bars and three oranges. He can buy or sell a candy bar for \$2 each. Similarly, he can buy or sell an orange for \$1. If he has no other source of income, draw his budget constraint and write the equation. What is the most he can spend, Y, on these goods?

3.2 What happens to the budget line if the government applies a specific tax of \$1 per gallon on gasoline but does not tax other goods? What happens to the budget line if the tax applies only to purchases of gasoline in excess of 10 gallons per week?

3.3 What is the effect of a quota of 13 thousand gallons of water per month on the opportunity set of the consumer in Solved Problem 4.2?

*3.4 What is the effect of a 50% income tax on Dale's budget line and opportunity set? (*Hint*: See Solved Problem 4.3.)

3.5 Change Solved Problem 4.3 so that Lisa's budget and the price of pizza double, but the price of burritos remains constant. Show how her budget constraint and opportunity set changes. Is Lisa necessarily better off than before these changes? (*Hint*: What happens to the intercepts of the budget line?)

4. Constrained Consumer Choice

4.1 What happens to a consumer's optimum if all prices and income double? (*Hint*: What happens to the intercepts of the budget line?)

4.2 Suppose that, to support domestic employment, the government imposes a 33.3% tariff on imports of clothing. Fareeha spends ₹8,000 on dresses each month. Before the tariff, the price of a dress produced domestically is ₹400 and the price of a foreign-made dress is ₹300. Use an indifference curve–budget line analysis to show how imposing this tariff affects the quantity of dresses she buys compared to what she would have bought in the absence of the tariff. Will she buy a relatively higher number of domestically produced dresses after the tariff? Why or why not?

4.3 Suppose that consumers in Bangladesh pay twice as much per unit for mangoes as they pay for bananas, whereas consumers in Indonesia pay half as much for mangoes as they pay for bananas. Assuming that consumers maximize their utility, which country's consumers have a higher marginal rate of substitution of mangoes for bananas? Explain your answer.

4.4 On their weekly visits to a local restaurant, Bram and Mila usually buy one slice of *appeltaart* (Dutch apple pie) to share and two cups of coffee each. However, the restaurant offers a special discount for the month of April to celebrate Koningsdag (King's Day): buy one slice of *appeltaart* and get the next at half-price. Show the original and new budget constraint. What can you say about the bundle Bran and Mila will choose when faced with the new constraint?

*4.5 Andy purchases only two goods, apples (a) and kumquats (k). He has an income of \$40 and can buy apples at \$2 per pound and kumquats at \$4 per pound. His utility function is $U(a, k) = 3a + 5k$. That is, his (constant) marginal utility for apples is 3 and his marginal utility for kumquats is 5. What bundle of apples and kumquats should he purchase to maximize his utility? Why? **A**

*4.6 David's utility function is $U = B + 2Z$, so $MU_B = 1$ and $MU_Z = 2$. Describe the location of his optimal bundle (if possible) in terms of the relative prices of B and Z. **A**

4.7 Linda loves buying shoes and going out to dance. Her utility function for pairs of shoes, S, and the number of times she goes dancing per month, T, is $U(S, T) = 2ST$, so $MU_S = 2T$ and $MU_T = 2S$. It costs Linda \$50 to buy a new pair of shoes or to spend an evening out dancing. Assume that she has \$500 to spend on shoes and dancing.

a. What is the equation for her budget line? Draw it (with T on the vertical axis), and label the slope and intercepts.

b. What is Linda's marginal rate of substitution? Explain.

c. Solve mathematically for her optimal bundle. Show how to determine this bundle in a diagram using indifference curves and a budget line. (*Hint*: See Solved Problem 4.4.) **A**

4.8 Vasco's utility function is $U = 10X^2Z$. The price of X is $p_X = \$10$, the price of Z is $p_z = \$5$, and his income is $Y = \$150$. What is his optimal consumption bundle? (*Hint*: See Appendix 4B.) Show this bundle in a graph. (*Hint*: See Solved Problem 4.4.) **C**

*4.9 Diogo has a utility function $U(B, Z) = AB^{\alpha}Z^{\beta}$, where A, α, and β are constants, B is burritos, and Z is pizzas. If the price of burritos, p_B, is \$2 and the price of pizzas, p_Z, is \$1, and Y is \$100, what is Diogo's optimal bundle? (*Hint*: See Solved Problem 4.4.) **C**

4.10 A dry spell in Kenya reduced the water level at the Ndakaini Dam in Kenya by 75% of its capacity in April 2017, resulting in water rationing in Nairobi and calls for the public to use water sparingly. A typical resident of Nairobi normally uses about 1.5 cubic meters of water per week. Suppose that, due to the dry spell, the government orders every resident to restrict their water use to 0.75 cubic meters per week. Use an indifference curve–budget line diagram, with water on the horizontal axis and all other goods on the vertical axis, to show how imposing such a restriction would affect the optimal bundle for a typical resident. If Mbaruku typically uses 0.75 cubic meters of water per week, would his optimal bundle be affected by the water restriction?

4.11 Salvo and Huse (2013) found that roughly one-fifth of owners of flexible-fuel cars (which can run on a mix of ethanol and gasoline) choose gasoline when the price of gas is 20% above that of ethanol (in energy-adjusted terms) and, similarly, one-fifth choose ethanol when ethanol is 20% more expensive than gasoline. What can you say about these people's tastes?

4.12 The household saving rate in Poland is low relative to other OECD countries, and Poland's population is aging. In 2016, the Polish government announced significant reforms to its pension system to help ensure that Poles will have adequate income in their retirement as numbers grow. It is also hoped that the reforms, which will go into effect in 2018, will encourage higher household saving by providing incentives to participate voluntarily in pension plans. Economic research shows that higher-income households tend to save more and that people respond to incentives. Use an indifference curve-budget line diagram, with the good "saving" on the horizontal axis and all other goods on the vertical axis, to show how the amount of saving in pension plans can increase as household income increases. Assume that no saving occurs when income is at or below the average household income level.

4.13 Maureen only drinks a cup of coffee with one teaspoon of sugar. In a figure, show the bundle of coffee and sugar that gives her the most pleasure. Are her indifference curve and the budget line tangent at that bundle?

4.14 Studies put the cost of raising a child up to the age of 18 at about €9,000 per year, so when a child is born, parents may have to make some sacrifices. Before having their first child, Ivan and Emma would often go out for dinner at fancy restaurants. However, dining out now involves extra costs associated with it, for example, planning and preparing for an evening out, and locating, engaging, and paying for a suitable caregiver. As the cost of restaurant meals has gone up, they now dine at home most evenings. Use an indifference curve-budget line diagram with "Dining out" on the horizontal axis and "Dining in" on the vertical axis to show how having a child has affected Ivan and Emma's optimal choice, assuming a constant budget.

4.15 A poor person who has an income of $1,000 receives $100 worth of food stamps. Draw the budget constraint if the food stamp recipient can sell these coupons on the black market for less than their face value.

4.16 Parents with a child in subsidized childcare in the province of Québec, Canada, pay a basic amount and, depending on family income, may pay an additional amount. As of January 1, 2017, families with a net income of $50,920 or less pay only the basic amount of $7.75 per day per child. Families with a net income of $161,380 or more pay a daily rate of $21.20 per child, which is composed of the basic amount plus the maximum additional amount of $13.45. How does the budget constraint for parents with four children who pay the basic amount for subsidized childcare in Québec compare to the budget constraint they would face in the absence of the subsidy? If the subsidy were provided only for their first two children, how would their budget constraint change?

4.17 Under the Healthy Start scheme, low-income women in the United Kingdom who are pregnant or have a child under 4 years of age can get free vouchers every week to spend on milk, fruit and vegetables, and infant formula milk. If the vouchers were instead sold at a subsidized rate (for example, 20% of the face value of a voucher), how would the mother's opportunity set change?

4.18 Is a poor person more likely to benefit from $100 a month worth of food stamps (which can be spent only on food) or $100 a month worth of clothing stamps (which can be spent only on clothing)? Why?

*4.19 Is a wealthy person more likely than a poor person to prefer to receive a government payment of $100 in cash to $100 worth of food stamps? Why or why not?

4.20 Under the Healthy Start scheme, low-income women in the United Kingdom who are pregnant or have a child under 4 years of age can get free vouchers every week to spend on milk, fruit and vegetables, and infant formula milk. They may also get housing subsidies under Housing Benefit to help pay their rent. Suppose Lily's income is £250 per week, which she spends on food and housing. Draw her budget line when the price of food and housing is £1 per unit each. If she receives £5 per week in food vouchers and the housing benefit is £45 per week, how do her budget line and opportunity set change?

4.21 The local swimming pool charges nonmembers $10 per visit. If you join the pool, you can swim for $5 per visit but you must pay an annual fee of *F*. Use an indifference curve diagram to find the value of *F* such that you are indifferent between joining and not joining. Suppose the annual fee for the pool (*F*) was exactly that amount. Would you go to the pool more or fewer times than if you did not join? For simplicity, assume that the price of all other goods is $1.

4.22 In Solved Problem 4.5, if the generic bleach increases its strength to equal that of Clorox without changing its price, what bundles will Chris buy if the price of Clorox is $3 or $2?

5. Behavioral Economics

5.1 Illustrate the logic of the endowment effect using a kinked indifference curve. Let the angle be greater than 90°. Suppose that the prices change, so the slope of the budget line through the endowment changes. Use a diagram to explain why an individual whose endowment point is at the kink will only trade from the endowment point if the price change is substantial.

*5.2 Why would a consumer's demand for a supermarket product change when the product price is quoted inclusive of taxes rather than before tax? Is the same effect as likely for people buying a car? (*Hint*: See the Application "Unaware of Taxes.")

6. Challenge

6.1 Suppose the Challenge Solution were changed so that Max and Bob have identical tastes, with the usual-shaped indifference curves. Use a figure to discuss how the different slopes of their budget lines affect the bundles of printed books and ebooks that each chooses. Can you make any unambiguous statements about how their bundles differ? Can you make an unambiguous statement if you know that Bob's budget line goes through Max's optimal bundle?

*6.2 State governments in India levy a value-added tax on the sale of goods. The rate of this sales tax varies by type of good and by state. In Maharashtra, the general rate of sales tax is 12.5%, and in the bordering state of Karnataka, it is 14.5%. This difference in tax rates can lead to a good being cheaper in one state than in another. Why might a government that shares a border with another state not wish to set a rate of sales tax that is significantly different from the rate applicable in the other state? While explaining your answer, consider how a large sales tax rate differential might affect the behavior of consumers who live near the border and can shop easily on the other side of it. What could the higher-taxed state do to mitigate any concerns that businesses in the border region might raise?

6.3 Einav et al. (2012) found that people who live in high sales tax locations are much more likely than other consumers to purchase goods over the Internet because Internet purchases are generally exempt from the sales tax if the firm is located in another state. They found that a 1% increase in a state's sales tax increases online purchases by that state's residents by just under 2%. Is the explanation for this result similar to that in the Challenge Solution? Why or why not?

6.4 Under a new Australian law, large foreign businesses are required to register with the Australian Taxation Office to collect the goods and services tax (GST, a 10% value-added tax) on all the goods they sell. Prior to July 2017, online sales of imported products worth up to AU$1,000 did not attract the GST. The Treasurer of Australia stated that the change would remove a disadvantage the low-value threshold causes for Australian businesses. Retailers feel the change will help level the playing field with online retailers like Amazon, Asos, and eBay. Use an indifference curve–budget line diagram to show what will happen when all sellers, both online and over-the-counter, pay the same amount of tax on the sale of the same good. (*Hint*: Start by drawing a typical consumer's indifference curve for buying an identical good, online and over-the-counter, at the same price without tax. Then add the tax to over-the-counter sales and finally to online sales.)

5 Applying Consumer Theory

The IRS = Theirs.

Challenge

Per-Hour Versus Lump-Sum Childcare Subsidies

Government childcare subsidies are common throughout the world but vary widely across countries. According to an Organization for Economic Cooperation and Development report in 2016, childcare spending as a percentage of gross domestic product was 0.1% in Germany, Mexico, and the United States; 0.3% in Japan; 0.4% in the United Kingdom; 0.6% in France; 0.8% in Finland; and 1.1% in Sweden.

The increased employment of mothers outside the home has led to a steep rise in childcare over the past half century. In the United States today, nearly seven out of ten mothers work outside the home—more than twice the rate in 1970. Eight out of ten employed mothers with children under age six are likely to have some form of nonparental childcare arrangement. Six out of ten children under the age of six are in childcare, as are 45% of children under age one. In contrast, in Korea, where parents receive 100% childcare subsidies for children up to five years old, 90% of children up to two years old are in childcare, as are 49% of three- to four-year-olds (Lee, 2016).

Childcare is a major burden for the poor, and the expense may prevent poor mothers from working. Paying for childcare for children under the age of five absorbed 25% of the earnings for families with annual incomes under $14,400, but only 6% for families with incomes of $54,000 or more. Government childcare subsidies increase the probability that a single mother will work at a standard job by 7% (Tekin, 2007). As one would expect, childcare subsidies have larger impacts on welfare recipients than on wealthier mothers.

In large part to help poor families obtain childcare so that the parents can work, the U.S. Child Care and Development Fund provided $9.4 billion to states in fiscal year 2016. Childcare programs vary substantially across states in their generosity and in the form of the subsidy.[1] Most states provide an ad valorem or a specific subsidy (see Chapter 3) to lower the hourly rate that a poor family pays for childcare.

Rather than subsidizing the price of childcare, some governments provide an unrestricted lump-sum payment that families can spend on childcare or on all other goods, such as food and housing. Canada provides such lump-sum payments.

For a given government expenditure, does a per-hour subsidy or a lump-sum subsidy provide greater benefit to recipients? Which option increases the demand for childcare services by more? Which one inflicts less cost on other consumers of childcare?

We can answer these questions posed in this chapter's Challenge using consumer theory. We can also use consumer theory to derive demand curves, to analyze the effects of providing cost-of-living adjustments to deal with inflation, and to derive labor supply curves.

[1]Some states provide a lump-sum amount of money while others provide a percentage of the cost of care.

We start by using consumer theory to show how to determine the shape of a demand curve for a good by varying the price of a good, holding other prices and income constant. Firms use information about the shape of demand curves when setting prices. Governments apply this information in predicting the impact of policies such as taxes and price controls.

We then use consumer theory to show how an increase in income causes the demand curve to shift. Firms use information about the relationship between income and demand to predict which less-developed countries will substantially increase their demand for the firms' products.

Next, we show that an increase in the price of a good has two effects on demand. First, consumers would buy less of the now relatively more expensive good even if they were compensated with cash for the price increase. Second, consumers' incomes can't buy as much as before because of the higher price, so consumers buy less of at least some goods.

We use this analysis of these two demand effects of a price increase to show why the government's measure of inflation, the Consumer Price Index (CPI), overestimates the amount of inflation. Because of this bias in the CPI, some people gain and some lose from contracts that adjust payment based on the government's inflation index. If you signed a long-term lease for an apartment in which your rent payments increase over time in proportion to the change in the CPI, you lose and your landlord gains from this bias.

Finally, we show how we can use the consumer theory of demand to determine an individual's labor *supply* curve. Knowing the shape of workers' labor supply curves is important in analyzing the effect of income tax rates on work and on tax collections. Many politicians, including Presidents John F. Kennedy, Ronald Reagan, and George W. Bush, have argued that if the income tax rates were cut, workers would work so many more hours that tax revenues would increase. If so, everyone would benefit from a tax cut. If not, the deficit could grow to record levels. Economists use empirical studies based on consumer theory to predict the effect of the tax rate cut on tax collections, as we discuss at the end of this chapter.

In this chapter, we examine five main topics

1. **Deriving Demand Curves.** We use consumer theory to derive demand curves, showing how a change in price causes a movement along a demand curve.
2. **How Changes in Income Shift Demand Curves.** We use consumer theory to determine how a demand curve shifts because of a change in income.
3. **Effects of a Price Change.** A change in price has two effects on demand, one having to do with a change in relative prices and the other concerning a change in the consumer's opportunities.
4. **Cost-of-Living Adjustments.** Using an analysis of the two effects of price changes, we show that the CPI overestimates the rate of inflation.
5. **Deriving Labor Supply Curves.** Using consumer theory to derive the demand curve for leisure, we can derive workers' labor supply curves and use them to determine how a reduction in the income tax rate affects labor supply and tax revenues.

5.1 Deriving Demand Curves

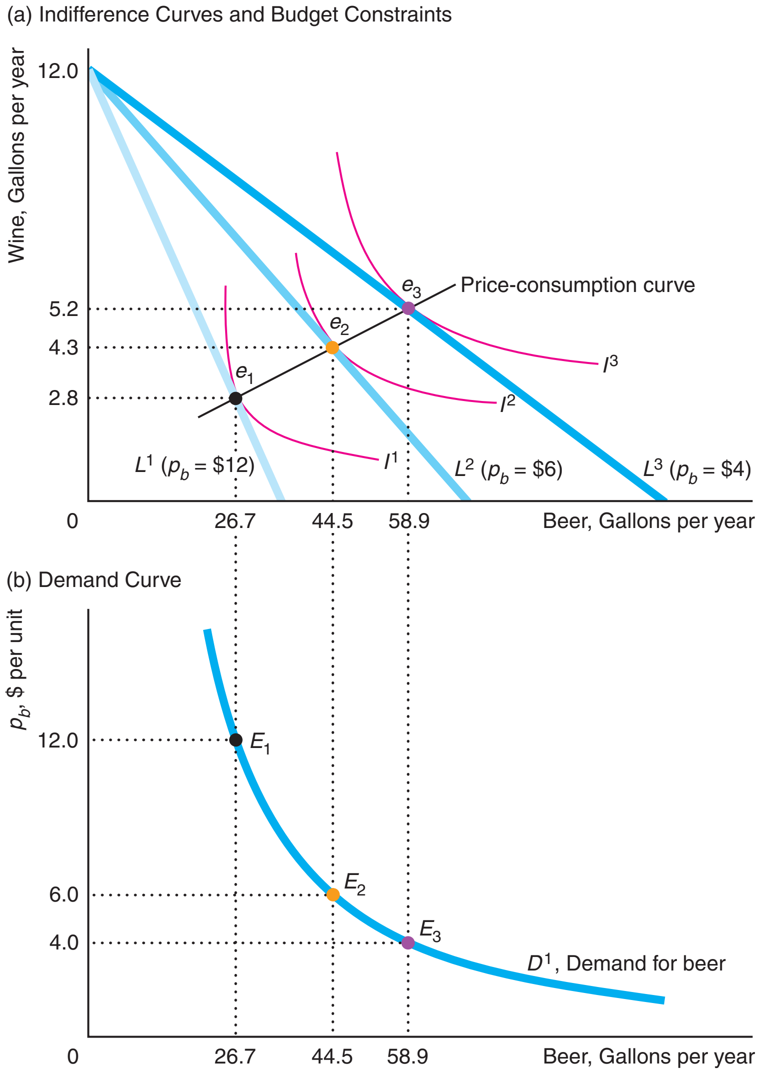
We use consumer theory to show by how much the quantity demanded of a good falls as its price rises. An individual chooses an optimal bundle of goods by picking the point on the highest indifference curve that touches the budget line (Chapter 4). When a price changes, the budget constraint the consumer faces shifts, so the consumer chooses a new optimal bundle. By varying one price and holding other prices and income constant, we determine how the quantity demanded changes as the price changes, which is the information we need to draw the demand curve. After deriving an individual's demand curve, we show the relationship between consumer tastes and the shape of the demand curve, which is summarized by the elasticity of demand (Chapter 3).

Indifference Curves and a Rotating Budget Line

We derive a demand curve using the information about tastes from indifference curves (see Appendix 4B for a mathematical approach). To illustrate how to construct a demand curve, we estimated a set of indifference curves between wine and beer, using data for U.S. consumers. Panel a of Figure 5.1 shows three estimated indifference curves for a typical U.S. consumer, whom we call Mimi. These indifference curves are

Figure 5.1 Deriving an Individual's Demand Curve MyLab Economics Video

If the price of beer falls, holding the price of wine, the budget, and tastes constant, the typical American consumer buys more beer, according to our estimates. (a) At the actual budget line, L^1, where the price of beer is \$12 per unit and the price of wine is \$35 per unit, the average consumer's indifference curve, I^1, is tangent at Bundle e_1, 26.7 gallons of beer per year and 2.8 gallons of wine per year. If the price of beer falls to \$6 per unit, the new budget constraint is L^2, and the average consumer buys 44.5 gallons of beer per year and 4.3 gallons of wine per year. (b) By varying the price of beer, we trace out the individual's demand curve, D^1. The beer price-quantity combinations E_1, E_2, and E_3 on the demand curve for beer in panel b correspond to optimal Bundles e_1, e_2, and e_3 in panel a.



convex to the origin: Mimi views beer and wine as imperfect substitutes (Chapter 4). We can construct Mimi's demand curve for beer by holding her budget, her tastes, and the price of wine constant at their initial levels and varying the price of beer.

The vertical axis in panel a measures the number of gallons of wine Mimi consumes each year, and the horizontal axis measures how many gallons of beer she drinks per year. Mimi spends $Y = \$419$ per year on beer and wine. The price of beer, p_b, is \$12 per unit, and the price of wine, p_w, is \$35 per unit.[2] The slope of her budget line, L^1, is $-p_b/p_w = -12/35 \approx \frac{1}{3}$. At those prices, Mimi consumes Bundle e_1, 26.7 gallons of beer per year and 2.8 gallons of wine per year, a combination that the tangency of indifference curve I^1 and budget line L^1 determine.[3]

If the price of beer falls in half to \$6 per unit while the price of wine and her budget remain constant, Mimi's budget line rotates outward to L^2. If she were to spend all her money on wine, she could buy the same 12 ($\approx$ 419/35) gallons of wine per year as before, so the intercept on the vertical axis of L^2 is the same as for L^1. However, if Mimi were to spend all her money on beer, she could buy twice as much as before (70 [$\approx$ 419/6] instead of 35 gallons of beer), so L^2 hits the horizontal axis twice as far from the origin as L^1. As a result, L^2 has a flatter slope than L^1, $-6/35 \approx -\frac{1}{6}$. The slope is flatter because the price of beer has fallen relative to the price of wine.

Because beer is now relatively less expensive, Mimi drinks relatively more beer. She chooses Bundle e_2, 44.5 gallons of beer per year and 4.3 gallons of wine per year, where her indifference curve I^2 is tangent to L^2. If the price of beer falls again, say, to \$4 per unit, Mimi consumes Bundle e_3, 58.9 gallons of beer per year and 5.2 gallons of wine per year.[4] The lower the price of beer, the happier Mimi is because she can consume more on the same budget: She is on a higher indifference curve (or perhaps just higher).

Price-Consumption Curve

Panel a also shows the *price-consumption curve*, which is the line through the optimal bundles, such as e_1, e_2, and e_3, that Mimi would consume at each price of beer, when the price of wine and Mimi's budget are held constant. Because the price-consumption curve is upward sloping, we know that Mimi's consumption of both beer and wine increases as the price of beer falls.

With different tastes—different shaped indifference curves—the price-consumption curve could be flat or downward sloping. If it were flat, then as the price of beer fell, the consumer would continue to purchase the same amount of wine and consume more beer. If the price-consumption curve were downward sloping, the individual would consume more beer and less wine as the price of beer fell.

[2]To ensure that the prices are whole numbers, we state the prices with respect to an unusual unit of measure (not gallons).

[3]These figures are the U.S. average annual per capita consumption of wine and beer. These numbers are startlingly high given that they reflect an average that includes teetotalers and (apparently heavy) drinkers. According to the World Health Organization, as of 2015, the consumption of liters of pure alcohol per capita by people 15 years and older was 9.0 in the United States, compared to 2.4 in Turkey, 7.0 in Norway, 7.5 in Japan, 8.7 in Sweden, 10.3 in Canada, 10.6 in Germany and Spain, 10.9 in Ireland, 11.2 in New Zealand, 12.0 in the United Kingdom, 12.6 in Australia, and 14.1 in the Czech Republic (**http://apps.who.int/gho/data/node.sdg.3-5-viz?lang=en**).

[4]These quantity numbers are probably higher than they would be in reality because we are assuming that Mimi continues to spend the same total amount of money on beer and wine as the price of beer drops.

Application

Smoking Versus Eating and Phoning

Tobacco use, one of the biggest public health threats the world has ever faced, killed 100 million people in the twentieth century. In 2016, the U.S. Centers for Disease Control and Prevention (CDC) reported that cigarette smoking and secondhand smoke are responsible for one in every five deaths each year in the United States—480,000 deaths every year. Half of all smokers die of tobacco-related causes. Worldwide, tobacco kills 6 million people a year.

Of the more than one billion smokers in the world, over 80% live in low- and middle-income countries. One way to get people—particularly poor people—to quit smoking is to raise the price of tobacco relative to the prices of other goods (thereby changing the slope of the budget constraints that individuals face). In poorer countries, smokers are giving up cigarettes to buy cell phones. As cell phones have recently become affordable in many poorer countries, the price ratio of cell phones to tobacco has fallen substantially.

According to Labonne and Chase (2011), in 2003, before cell phones were common, 42% of households in the Philippine villages they studied used tobacco, and 2% of total village income was spent on tobacco. After the price of cell phones fell, ownership of the phones quadrupled from 2003 to 2006. As consumers spent more on mobile phones, tobacco use fell by one-third in households in which at least one member had smoked (so that consumption fell by one-fifth for the entire population). That is, if we put cell phones on the horizontal axis and tobacco on the vertical axis and reduce the price of cell phones, the price-consumption curve is downward sloping (unlike that in panel a of Figure 5.1).

At least 163 countries tax cigarettes to raise tax revenue and to discourage smoking. Lower-income and younger populations are more likely than others to quit smoking when the price increases. Although the estimated U.S. price elasticity is about 0.25 (Pesko et al., 2016), Nikaj and Chaloupka (2013) estimated the demand elasticity for cigarettes on youth smoking across 38 countries to be -1.5. Moreover, in relatively poor countries, the elasticity was -2.2.

But what happens to those who continue to smoke heavily? To pay for their now more expensive habit, they have to reduce their expenditures on other goods, such as housing and food. Busch et al. (2004) found that a 10% increase in the price of cigarettes caused poor, smoking families to cut back on cigarettes by 9%, alcohol and transportation by 11%, food by 17%, and health care by 12%. These smoking families allocated 36% of their expenditures to housing compared to 40% for nonsmokers. Thus, to continue to smoke, these people cut back on many basic goods. If we put tobacco on the horizontal axis and all other goods on the vertical axis, the price-consumption curve is upward sloping, so that as the price of tobacco rises, the consumer buys less of both tobacco and all other goods.

The Demand Curve Corresponds to the Price-Consumption Curve

We can use the same information in the price-consumption curve to draw Mimi's demand curve for beer, D^1, in panel b of Figure 5.1. Corresponding to each possible price of beer on the vertical axis of panel b, we record on the horizontal axis the quantity of beer demanded by Mimi from the price-consumption curve.

Points E_1, E_2, and E_3 on the demand curve in panel b correspond to Bundles e_1, e_2, and e_3 on the price-consumption curve in panel a. Both e_1 and E_1 show that when the price of beer is \$12, Mimi demands 26.7 gallons of beer per year. When the price falls to \$6 per unit, Mimi increases her consumption to 44.5 gallons of beer, point E_2. The demand curve, D^1, is downward sloping as predicted by the Law of Demand.

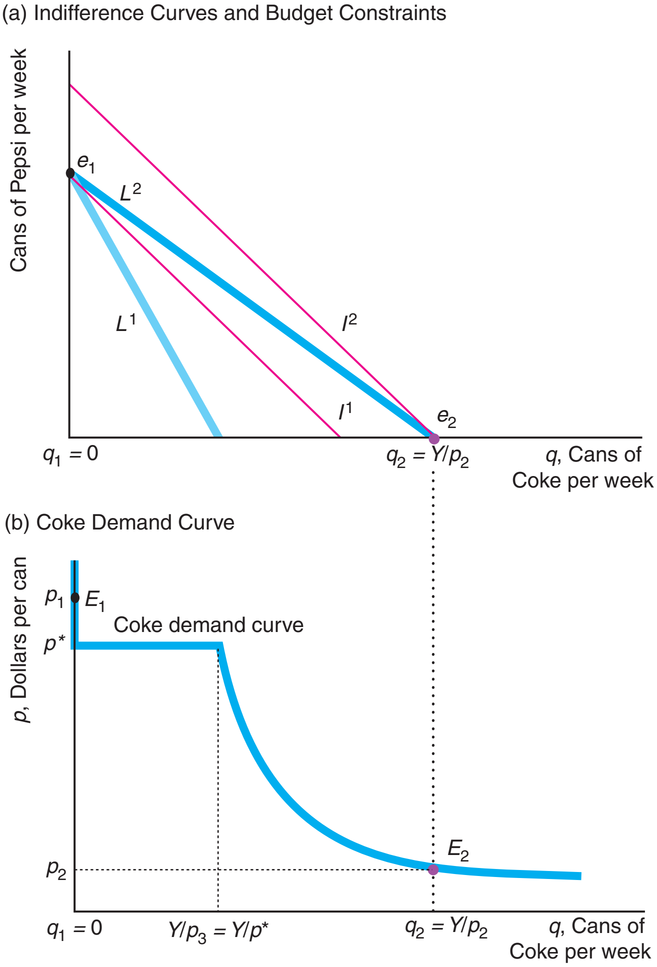
Solved Problem 5.1

MyLab Economics Solved Problem

Mahdu views Coke, q, and Pepsi as perfect substitutes: He is indifferent as to which one he drinks. The price of a 12-ounce can of Coke is p, the price of a 12-ounce can of Pepsi is p^*, and his weekly cola budget is Y. Derive Mahdu's demand curve for Coke using the method illustrated in Figure 5.1. (*Hint*: See Solved Problem 4.4.)

Answer

1. *Use indifference curves to derive Mahdu's equilibrium choice.* Panel a of the figure shows that his indifference curves I^1 and I^2 have a slope of -1 because Mahdu is indifferent as to which good he buys (see Chapter 4). We keep the price of Pepsi, p^*, fixed and vary the price of Coke, p. Initially, the budget line L^1 is steeper than the indifference curves because the price of Coke is greater than that of Pepsi, $p_1 > p^*$. Mahdu maximizes his utility by choosing Bundle e_1, where he purchases only Pepsi (a corner solution, see Chapter 4). If the price of Coke is $p_2 < p^*$, the budget line L^2 is flatter than the indifference curves. Mahdu maximizes his utility at e_2, where he spends his cola budget on Coke, buying as many cans of Coke as he can afford, $q_2 = Y/p_2$, and he consumes no Pepsi. If the price of Coke is $p_3 = p^*$, his budget line would have the same slope as his indifference curves, and one indifference curve would lie on top of the budget line. Consequently, he would be indifferent between buying any quantity of q between 0 and $Y/p_3 = Y/p^*$ (and his total purchases of Coke and Pepsi would add to $Y/p_3 = Y/p^*$).



2. *Use the information in panel a to draw his Coke demand curve.* Panel b shows Mahdu's demand curve for Coke, q, for a given price of Pepsi, p^*, and Y. When the price of Coke is above p^*, his demand curve lies on the vertical axis, where he demands zero units of Coke, such as point E_1 in panel b, which corresponds to e_1 in panel a. If the prices are equal, he buys any amount of Coke up to a maximum of $Y/p_3 = Y/p^*$. If the price of Coke is $p_2 < p^*$, he buys Y/p_2 units at point E_2, which corresponds to e_2 in panel a. When the price of Coke is less than that of Pepsi, the Coke demand curve asymptotically approaches the horizontal axis as the price of Coke approaches zero.

5.2 How Changes in Income Shift Demand Curves

To trace out the demand curve, we have looked at how an increase in the good's price—holding income, tastes, and other prices constant—causes a downward *movement along the demand curve*. Now we examine how an increase in income, when all prices are held constant, causes a *shift of the demand curve*.

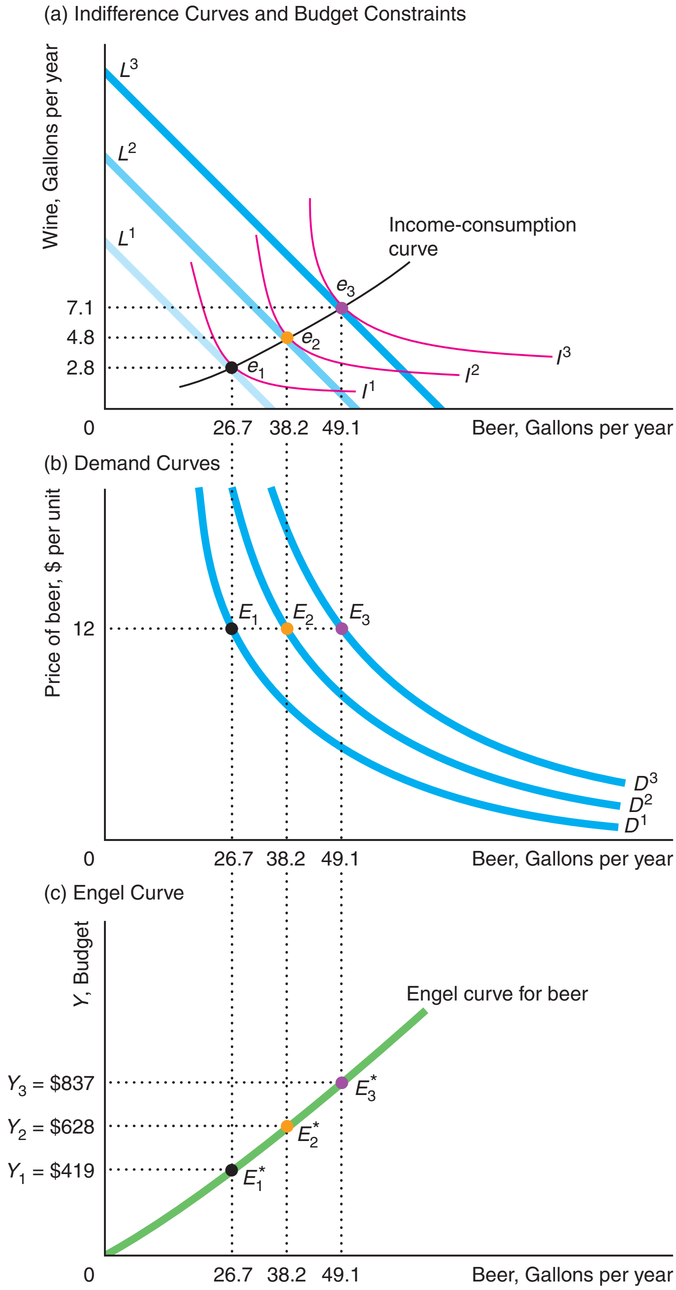
Businesses routinely use information on the relationship between income and the quantity demanded. For example, in deciding where to market its products, Whirlpool wants to know which countries are likely to spend a relatively large percentage of any extra income on refrigerators and washing machines.

Effects of a Rise in Income

We illustrate the relationship between the quantity demanded and income by examining how Mimi's behavior changes when her income rises while the prices of beer and wine remain constant. Figure 5.2 shows three ways of looking at the

Figure 5.2 Effect of a Budget Increase on an Individual's Demand Curve

As Mimi's annual budget for wine and beer, *Y*, increases from \$419 to \$628 and then to \$837, holding prices constant, Mimi buys more of both products, so her income-consumption curve slopes upward (a). The outward shift of her demand curve for beer occurs because Mimi buys more beer as her income increases (b) so her Engel curve for beer slopes upward (c).



relationship between income and the quantity demanded. All three diagrams have the same horizontal axis: the quantity of beer consumed per year. In the consumer theory diagram, panel a, the vertical axis is the quantity of wine consumed per year. In the demand curve diagram, panel b, the vertical axis is the price of beer per unit. Finally, in panel c, which shows the relationship between income and quantity directly, the vertical axis is Mimi's budget, Y, which is referred to as an Engel curve.

A rise in Mimi's income causes the budget constraint to shift outward in panel a, which increases Mimi's opportunities. Her budget constraint L^1 at her original income, $Y = \$419$, is tangent to her indifference curve I^1 at e_1.

As before, Mimi's demand curve for beer is D^1 in panel b. Point E_1 on D^1, which corresponds to point e_1 in panel a, shows how much beer—26.7 gallons per year—Mimi consumes when the price of beer is \$12 per unit (and the price of wine is \$35 per unit).

Now suppose that Mimi's beer and wine budget, Y, increases by roughly 50% to \$628 per year. Her new budget line, L^2 in panel a, is farther from the origin but parallel to her original budget constraint, L^1, because the prices of beer and wine are unchanged. Given this larger budget, Mimi chooses Bundle e_2. The increase in her income causes her demand curve to shift to D^2 in panel b. Holding Y at \$628, we can derive D^2 by varying the price of beer, in the same way as we derived D^1 in Figure 5.1. When the price of beer is \$12 per unit, Mimi buys 38.2 gallons of beer per year, E_2 on D^2. Similarly, if Mimi's income increases to \$837 per year, her demand curve shifts to D^3.

The *income-consumption curve* through Bundles e_1, e_2, and e_3 in panel a shows how Mimi's consumption of beer and wine increases as her income rises. As Mimi's income goes up, her consumption of both wine and beer increases.

Engel curve
the relationship between the quantity demanded of a single good and income, holding prices constant

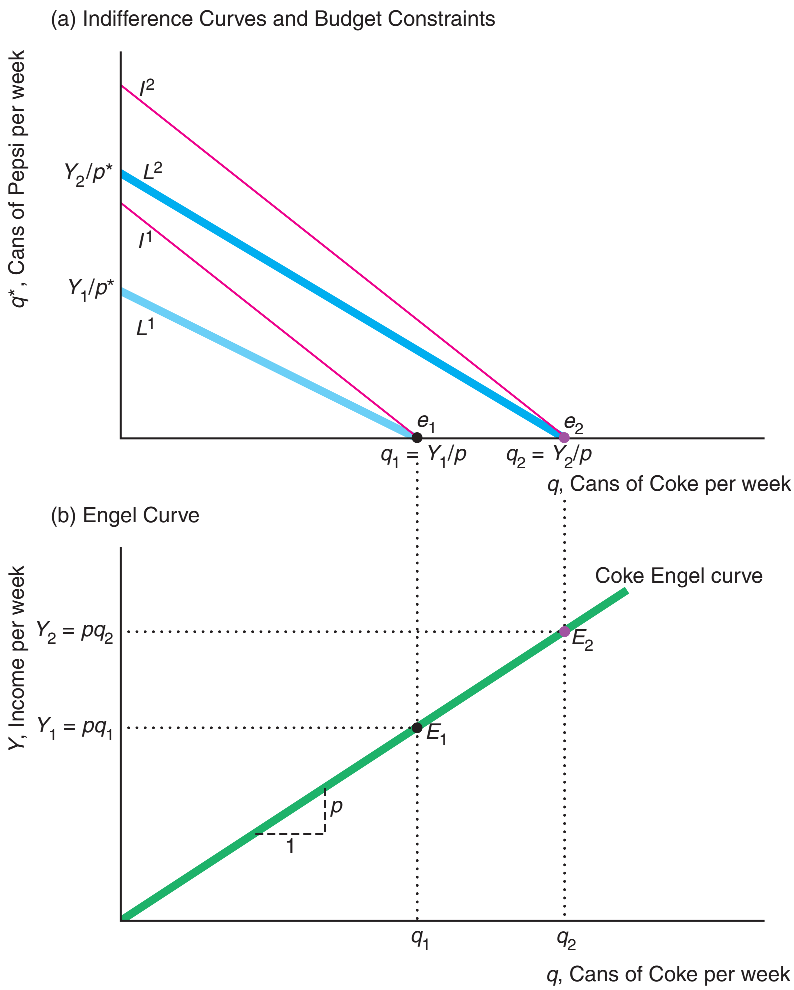
We can show the relationship between the quantity demanded and income directly rather than by shifting demand curves to illustrate the effect. In panel c, we plot an **Engel curve**, which shows the relationship between the quantity demanded of a single good and income, holding prices constant. Income is on the vertical axis, and the quantity of beer demanded is on the horizontal axis. On Mimi's Engel curve for beer, points E_1^*, E_2^*, and E_3^* correspond to points E_1, E_2, and E_3 in panel b and to e_1, e_2, and e_3 in panel a.

Solved Problem 5.2

Mahdu views Coke and Pepsi as perfect substitutes. The price of a 12-ounce can of Coke, p, is less than the price of a 12-ounce can of Pepsi, p^*. What does Mahdu's Engel curve for Coke look like? How much does his weekly cola budget have to rise for Mahdu to buy one more can of Coke per week?

Answer

1. *Use indifference curves to derive Mahdu's optimal choice*. Because Mahdu views the two brands as perfect substitutes, his indifference curves, such as I^1 and I^2 in panel a of the graphs, are straight lines with a slope of -1. When his income is Y_1, his budget line hits the Pepsi axis at Y_1/p^* and the Coke axis at Y_1/p. Mahdu maximizes his utility by consuming Y_1/p cans of the less expensive Coke and no Pepsi (a corner solution). As his income rises, say, to Y_2, his budget line shifts outward and is parallel to the original one, with the same slope of $-p/p^*$. Thus, at each income level, his budget lines are flatter than his indifference curves, so his equilibria lie along the Coke axis.



2. *Use the first figure to derive his Engel curve.* Because his entire budget, Y, goes to buying Coke, Mahdu buys $q = Y/p$ cans of Coke. This expression, which shows the relationship between his income and the quantity of Coke he buys, is Mahdu's Engel curve for Coke. The points E_1 and E_2 on the Engel curve in panel b correspond to e_1 and e_2 in panel a. We can rewrite this expression for his Engel curve as $Y = pq$. As a result, his Engel curve in panel b is a straight line with a slope of p. As q increases by one can ("run"), Y increases by p ("rise"). Because he spends his entire cola budget on Coke, Mahdu's income needs to rise by only p for him to buy one more can of Coke per week.

Consumer Theory and Income Elasticities

Income elasticities tell us how much the quantity demanded changes as income increases. We can use income elasticities to summarize the shape of the Engel curve, the shape of the income-consumption curve, or the movement of the demand curves when income increases. For example, firms use income elasticities to predict the impact of income taxes on consumption. We first discuss the definition of income elasticities and then examine their connection to the income-consumption curve.

Income Elasticities We defined the income elasticity of demand in Chapter 3 as

$$\xi = \frac{\text{percentage change in quantity demanded}}{\text{percentage change in income}} = \frac{\Delta Q/Q}{\Delta Y/Y},$$

where ξ is the Greek letter xi. Mimi's income elasticity of beer, ξ_b, is 0.88, and that of wine, ξ_w, is 1.38 (based on our estimates for the average U.S. consumer). When her income goes up by 1%, she consumes 0.88% more beer and 1.38% more wine. Thus, according to these estimates, as income falls, consumption of beer and wine by the average American falls—contrary to frequent (unsubstantiated) claims in the media that people drink more as their incomes fall during recessions.

normal good
a commodity of which as much or more is demanded as income rises

inferior good
a commodity of which less is demanded as income rises

Most goods, like beer and wine, have positive income elasticities. A good is called a **normal good** if as much or more of it is demanded as income rises. Thus, a good is a normal good if its income elasticity is greater than or equal to zero: $\xi \geq 0$.

Some goods, however, have negative income elasticities: $\xi < 0$. A good is called an **inferior good** if less of it is demanded as income rises. We are not making a value judgment when we say a good is *inferior*. An inferior good need not be defective or of low quality. Some of the better-known examples of inferior goods are foods such as potatoes and cassava that very poor people typically eat in large quantities. Some economists—apparently seriously—claim that human meat is an inferior good: Only when the price of other foods is very high and people are starving will they turn to cannibalism. Bezmen and Depken (2006) estimate that pirated goods are inferior: a 1% increase in per-capita income leads to a 0.25% reduction in piracy.

Another strange example concerns treating children as a consumption good. Even though people can't buy children in a market, people can decide how many children to have. Guinnane (2011) surveyed the literature and reported that most studies find that the income elasticity for the number of children in a family is negative but close to zero. Thus, the number of children demanded is not very sensitive to income.

Income-Consumption Curves and Income Elasticities The slope of the income-consumption curve for two goods tells us the sign of the income elasticities: whether the income elasticities for those goods are positive or negative. We know that Mimi's income elasticities of beer and wine are positive because the income-consumption curve in panel a of Figure 5.2 is upward sloping. As income rises, the budget line shifts outward and hits the upward-sloping income-consumption line at higher levels of both goods. Thus, as her income rises, Mimi demands more beer and wine, so her income elasticities for beer and wine are positive. Because the income elasticity for beer is positive, the demand curve for beer shifts to the right in panel b of Figure 5.2 as income increases.

To illustrate the relationship between the slope of the income-consumption curve and the sign of income elasticities, we examine Peter's choices of food and housing. Peter purchases Bundle *e* in Figure 5.3 when his budget constraint is L^1. When his income increases, so that his budget constraint is L^2, he selects a bundle on L^2. Which bundle he buys depends on his tastes—his indifference curves.

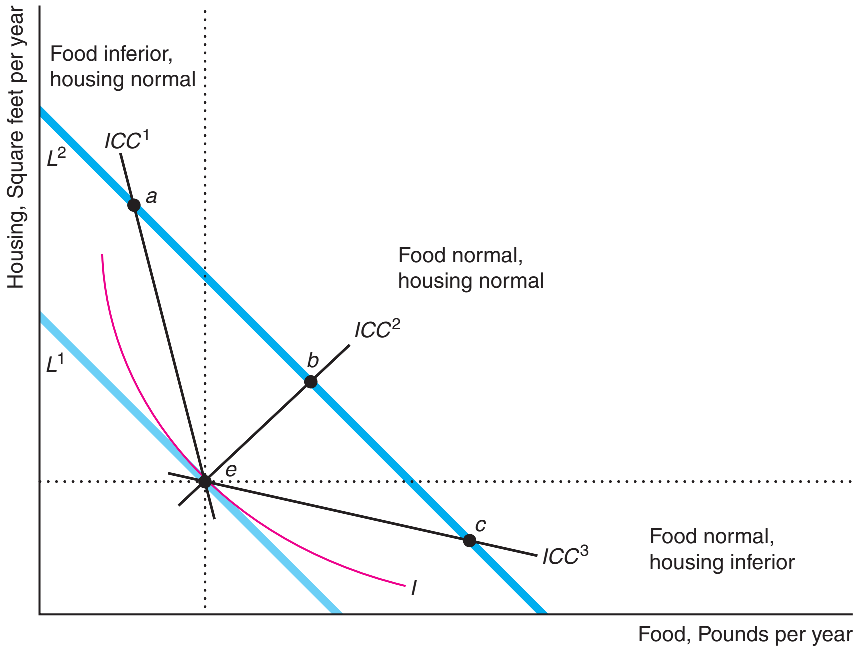
The horizontal and vertical dotted lines through *e* divide the new budget line, L^2, into three sections. In which of these three sections the new optimal bundle is located determines Peter's income elasticities of food and clothing.

Suppose that Peter's indifference curve is tangent to L^2 at a point in the upper-left section of L^2 (to the left of the vertical dotted line that goes through *e*), such as Bundle *a*. If Peter's income-consumption curve is ICC^1, which goes from *e* through *a*, he buys more housing and less food as his income rises. (We draw the possible *ICC* curves as straight lines for simplicity. In general, they may curve.) Here, housing is a normal good and food an inferior good.

If instead the new optimal bundle is located in the middle section of L^2 (above the horizontal dotted line and to the right of the vertical dotted line), such as at *b*, his

Figure 5.3 Income-Consumption Curves and Income Elasticities

At the initial income, the budget constraint is L^1 and the optimal bundle is e. After income rises, the new constraint is L^2. With an upward-sloping income-consumption curve such as ICC^2, both goods are normal. With an income-consumption curve such as ICC^1 that goes through the upper-left section of L^2 (to the left of the vertical dotted line through e), housing is normal and food is inferior. With an income-consumption curve such as ICC^3 that cuts L^2 in the lower-right section (below the horizontal dotted line through e), food is normal and housing is inferior.



income-consumption curve ICC^2 through e and b is upward sloping. He buys more of both goods as his income rises, so both food and housing are normal goods.

Now suppose that his new optimal bundle is in the bottom-right segment of L^2 (below the horizontal dotted line). If his new optimal bundle is c, his income-consumption curve ICC^3 slopes downward from e through c. As his income rises, Peter consumes more food and less housing, so food is a normal good and housing is an inferior good.

Some Goods Must Be Normal It is impossible for all goods to be inferior. We illustrate this point using Figure 5.3. At his original income, Peter faced budget constraint L^1 and bought the combination of food and housing e. When his income goes up, his budget constraint shifts outward to L^2. Depending on his tastes (the shape of his indifference curves), he may buy more housing and less food, such as Bundle a; more of both, such as b; or more food and less housing, such as c. Therefore, either both goods are normal or one good is normal and the other is inferior.

If both goods were inferior, Peter would buy less of both goods as his income rises—which makes no sense. Were he to buy less of both, he would be buying a bundle that lies inside his original budget constraint L^1. Even at his original income, he could have purchased that bundle but bought e instead. According to the more-is-better assumption in Chapter 4, for any given bundle inside the constraint, a bundle on the budget constraint gives Peter more utility.

Even if an individual does not buy more of the usual goods and services, that person may put the extra money into savings. Empirical studies find that savings is a normal good.

Application

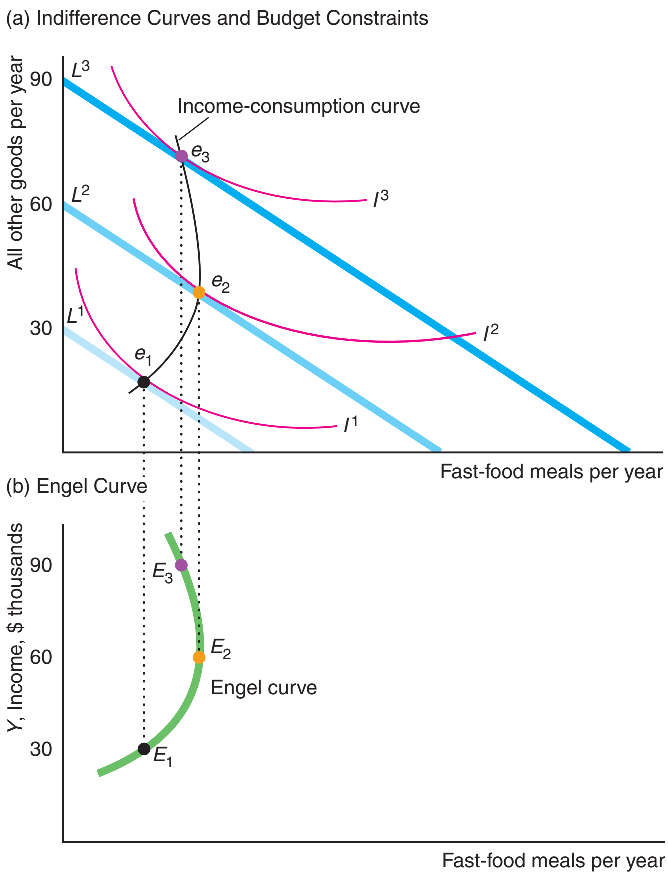
Fast-Food Engel Curve

Is a meal at a fast-food restaurant a normal or inferior good? This question is important because, as incomes have risen over the last quarter century, Americans have spent a larger share of their income on fast food, which many nutritionists blame for increased obesity rates. However, a number of studies find that obesity falls with income, which suggests that a fast-food meal may be an inferior good, at least at high incomes.

Kim and Leigh (2011) estimated the demand for fast-food restaurant visits as a function of prices, income, and various socioeconomic variables such as age, family size, and whether the family received food stamps (which lowers the price of supermarket food relative to restaurant food). They found that fast-food restaurant visits increase with income up to \$60,000, and then decrease as income rises more.[5]

The figure derives the Engel curve for Gail, a typical consumer, based on Kim and Leigh's estimates. Panel a shows that Gail spends her money on fast-food meals (horizontal axis, where Y is measured in thousands) and all other goods (vertical axis). As Gail's income increases from \$30,000 to \$60,000, her budget line shifts outward, from L^1 to L^2. As a result, she eats more restaurant meals: Her new optimal bundle, e_2, lies to the right of e_1. Thus, a fast-food meal is a normal good in this range.

As her income increases further to \$90,000, her budget line shifts outward to L^3, and she reduces her consumption of hamburgers: Bundle e_3 lies to the left of e_2. Thus, at higher incomes, Gail views a fast-food meal as an inferior good.



[5] In contrast, they found that full-service restaurant visits increase with income up to \$95,000.

Panel b shows her corresponding Engel curve for fast food. As her income rises from \$30,000 to \$60,000, she moves up and to the right from E_1 (which corresponds to e_1 in panel a) to E_2. Her Engel curve is upward sloping in this range, indicating that she buys more fast-food meals as her income rises. As her income rises further, her Engel curve is backward bending.

5.3 Effects of a Price Change

substitution effect
the change in the quantity of a good that a consumer demands when that good's price rises, holding other prices and the consumer's utility constant

income effect
the change in quantity of a good that a consumer demands because of a change in income, holding prices constant

Holding tastes, other prices, and income constant, an increase in the price of a good has two effects on an individual's demand. One is the **substitution effect**: the change in the quantity of a good that a consumer demands when that good's price rises, holding other prices and the consumer's utility constant. Holding utility constant, as the price of the good increases, consumers *substitute* other now relatively cheaper goods for that one.

The other effect is the **income effect**: the change in quantity of a good that a consumer demands because of a change in income, holding prices constant. An increase in price reduces a consumer's buying power, effectively reducing the consumer's *income* or opportunity set and causing the consumer to buy less of at least some goods. A doubling of the price of all the goods the consumer buys is equivalent to reducing income to half its original level. Even an increase in the price of only one good reduces a consumer's ability to buy the same amount of all goods as previously. For example, if the price of food increases in China, the effective purchasing power of a Chinese consumer falls substantially because urban consumers spent 36% of their income on food and rural residents spent 40% (*Statistical Yearbook of China*, 2012).

When a price goes up, the total change in the quantity purchased is the sum of the substitution and income effects.[6] When estimating the effects of a price change on the quantity an individual demands, economists decompose the combined income and substitution effects into the two separate components. By doing so, they gain extra information that they can use to answer questions about whether inflation measures are accurate, whether an increase in tax rates will raise tax revenue, and what the effects are of government policies that compensate some consumers. For example, Presidents Jimmy Carter, Bill Clinton, and Barack Obama, when calling for increases in various energy taxes, proposed providing an income compensation for poor consumers to offset the burden of these higher taxes. We can use knowledge of the substitution and income effects resulting from a price change of energy to evaluate the effect of such policies.

Income and Substitution Effects with a Normal Good

To illustrate the substitution and income effects, we examine the choice between music tracks (songs) and live music. In 2008, a typical British young person (ages 14 to 24), whom we call Laura, bought 24 music tracks, T, per quarter and consumed 18 units of live music, M, per quarter.[7] We estimate Laura's utility function and use it to draw Laura's indifference curves in Figure 5.4.[8]

---

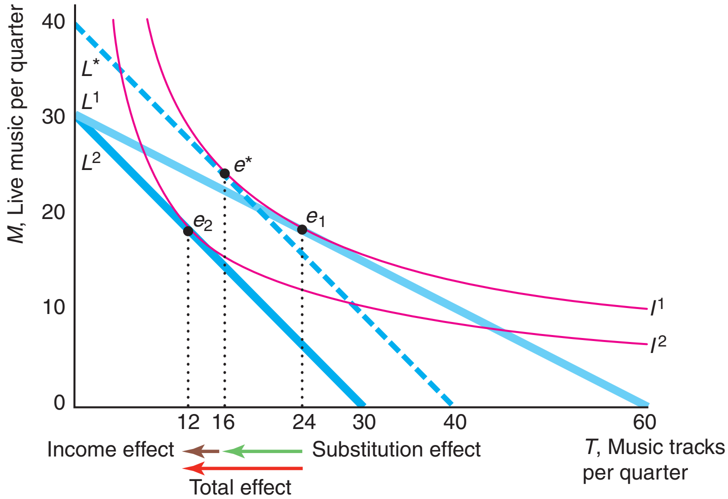
[6]See Appendix 5A for the mathematical relationship, called the *Slutsky equation*. See also the discussion of the Slutsky equation at MyLab Economics, Chapter 5, "Measuring the Substitution and Income Effects."

[7]A unit of live music is the amount that £1 purchases (that is, it does not correspond to a full concert or a performance in a pub). Data on total expenditures are from *The Student Experience Report, 2007*, **http://www.unite-students.com**, while budget allocations between live and recorded music are from the 2008 survey of the *Music Experience and Behaviour in Young People* produced by British Music Rights and the University of Hertfordshire.

[8]Laura's estimated utility function is $U = T^{0.4}M^{0.6}$, which is a type of Cobb-Douglas utility function (Appendix 4A).

Figure 5.4 Substitution and Income Effects with Normal Goods MyLab Economics Video

A doubling of the price of music tracks from £0.5 to £1 causes Laura's budget line to rotate from L^1 to L^2. The imaginary budget line L^* has the same slope as L^2 and is tangent to indifference curve I^1. The shift of the optimal bundle from e_1 to e_2 is the *total effect* of the price change. The total effect can be decomposed into the *substitution effect*—the movement from e_1 to e^*—and the *income effect*—the movement from e^* to e_2.



Because Laura's entertainment budget for the quarter is $Y = £30$, the price of a music track from Amazon.com or its major competitors is £0.5, and the price for a unit of live music is £1 (where we pick the unit appropriately; see footnote 8), her original budget constraint is L^1 in Figure 5.4. She can afford to buy 60 music tracks and no live music, 30 units of live music and no music tracks, or any combination between these extremes.

Given her estimated utility function, Laura's demand functions are $T = 0.4Y/p_T$ music tracks and $M = 0.6Y/p_M$ units of live music. At the original prices and with an entertainment budget of $Y = £30$ per quarter, Laura chooses Bundle e_1, $T = 0.4 \times £30/£0.5 = 24$ music tracks and $M = 0.6 \times £30/£1 = 18$ units of live music per quarter, where her indifference curve I^1 is tangent to her budget constraint L^1.

Now suppose that the price of a music track doubles to £1, causing Laura's budget constraint to rotate inward from L^1 to L^2 in Figure 5.4. Because music tracks are not twice as expensive, the new budget constraint, L^2, is twice as steep as L^1. That is, the slope of L^2 is $-p_T/p_M = -1/1 = -1$, while the slope of L^1, $-p_T/p_M = -0.5/1 = -0.5$. Laura's opportunity set is smaller, so she can choose between fewer bundles of music tracks and live music units than she could at the lower music track price. The area between the two budget constraints reflects the decrease in her opportunity set resulting from the price increase of music tracks. At this higher price for music tracks, Laura's new optimal bundle is e_2 (where she buys $T = 0.4 \times 30/1 = 12$ music tracks), located where her indifference curve I^2 is tangent to L^2.

The movement from e_1 to e_2 is the total change in her consumption due to the higher price of music tracks. In particular, the *total effect* on Laura's consumption of music tracks due to the price increase is that she now buys 12 (= 24 − 12) fewer music tracks per quarter. In the figure, the red arrow pointing to the left and labeled "Total effect" shows this decrease.

We can break the total effect into its two components: substitution effect and income effect. As the price of music tracks increases, Laura's opportunity set shrinks even though her income is unchanged. If, as a thought experiment, we compensate her for this loss by giving her extra income, we can determine her substitution effect. The *substitution effect* is the change in the quantity demanded from a *compensated change in the price* of music tracks, which occurs when we increase Laura's income by enough to offset the rise in the price of music tracks so that her utility stays constant.

To determine the substitution effect, we draw an imaginary budget constraint, L^*, that is parallel to L^2 and tangent to Laura's original indifference curve, I^1. This imaginary budget constraint, L^*, has the same slope, -1, as L^2, because both curves are based on the new, higher price of music tracks. For L^* to be tangent to I^1, we need to increase Laura's budget from £30 to £40 to offset the reduction in utility from the higher price of music tracks. If Laura's budget constraint were L^*, she would choose Bundle e^*, where she buys $T = 0.4 \times 40/1 = 16$ tracks.

Thus, if the price of tracks rises relative to that of live music and we hold Laura's utility constant by raising her income, Laura's optimal bundle shifts from e_1 to e^*, which is the substitution effect. She buys 8 (= 24 − 16) fewer tracks per quarter, as the green arrow pointing to the left labeled "Substitution effect" shows.

Laura also faces an income effect because the higher price of tracks shrinks her opportunity set, so she must buy a bundle on a lower indifference curve. As a thought experiment, we can ask how much we would have to lower Laura's income, while holding prices constant, for her to choose a bundle on this new, lower indifference curve. The *income effect* is the change in the quantity of a good a consumer demands because of a change in income, holding prices constant. The parallel shift of the budget constraint from L^* to L^2 captures this effective decrease in income. The movement from e^* to e_2 is the income effect, as the brown arrow pointing to the left labeled "Income effect" shows. As her budget decreases from £40 to £30, Laura consumes 4 (= 16 − 12) fewer tracks per year.

The *total effect* from the price change is the *sum of the substitution and income effects*, as the arrows show. The total effect on Laura's demand for music tracks per year of an increase in the price of music tracks is

$$\begin{array}{cccccc} \text{Total effect} & = & \text{substitution effect} & + & \text{income effect} \\ -12 & = & -8 & + & (-4). \end{array}$$

Because indifference curves are convex to the origin, *the substitution effect is unambiguous*: When a good's price rises, a consumer substitutes a less expensive good for this more expensive one, holding utility constant. Because we are holding utility constant, the substitution effect causes a *movement along an indifference curve*.

The income effect causes a shift to another indifference curve due to a change in the consumer's opportunity set. The direction of the income effect depends on the income elasticity. Because a music track is a normal good for Laura, her income effect is negative. Thus, both Laura's substitution effect and her income effect go in the same direction, so the total effect of the price rise must be negative.

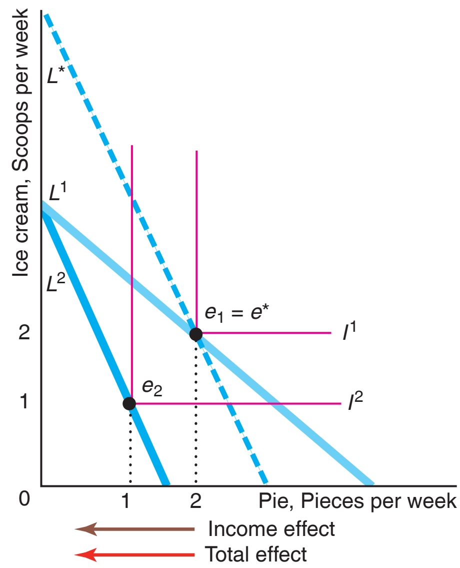
Solved Problem 5.3

MyLab Economics Solved Problem

Maureen loves apple pie à la mode (a slice of pie with a scoop of vanilla ice cream on top) but she doesn't like apple pie by itself or vanilla ice cream by itself (panel b of Figure 4.4). That is, she views apple pie and vanilla ice cream as perfect complements. At the initial prices, she consumed two pieces of pie per week. After the price of pie rises, she chooses to consume only one piece of pie. In a graph similar to Figure 5.4, show the substitution, income, and total effects of the higher price of pie on Maureen's demand for pie.

Answer

1. *Show that the price increase causes the budget line to rotate at the intersection on the ice cream axis and that Maureen's optimal bundle shifts from two units of pie and ice cream to one unit.* In the figure, her initial budget line is L^1 and her optimal bundle is e_1, where her indifference curve I^1 touches L^1. When the price of pie increases, her new budget line is L^2 and her new optimal bundle is e_2.



2. *Draw a line, L^*, that is parallel to L^2 and that touches Maureen's original indifference curve, I^1, and show the relationship between the new tangency point, e^*, and her original one, e^1.* The indifference curve I^1 touches L^* at e^*, which is the same point as e_1.

3. *Discuss the substitution, income, and total effects.* No substitution effect occurs because Maureen is unwilling to substitute between pie and ice cream. The brown arrow shows the income effect of the price increase is a decrease in demand from two pieces of pie per week to one. The red arrow shows that the total effect is identical to the income effect.

Solved Problem 5.4

Next to its plant, a manufacturer of dinner plates has an outlet store that sells plates of both first quality (perfect plates) and second quality (plates with slight blemishes). The outlet store sells a relatively large share of seconds. At its regular stores elsewhere, the firm sells many more first-quality plates than second-quality plates. Why? (Assume that consumers' tastes with respect to plates are the same everywhere and the cost of shipping each plate from the factory to the firm's other stores is s.)

Answer

1. *Determine how the relative prices of plates differ between the two types of stores.* The slope of the budget line at the factory outlet store is $-p_1/p_2$, where p_1 is the price of first-quality plates and p_2 is the price of the seconds. It costs the same to ship, s, a first-quality plate as a second because they weigh the same and have to be handled in the same way. At all other stores, the firm adds the cost of shipping to the price it charges at its factory outlet store, so the price of a first-quality plate at those stores is $p_1 + s$ and the price of a second is $p_2 + s$. As a result, the slope of the budget line consumers face at the other retail stores is $-(p_1 + s)/(p_2 + s)$. The seconds are relatively less expensive at the factory outlet than at other stores. For example, if $p_1 = \$2$, $p_2 = \$1$, and $s = \$1$ per plate, the slope of the budget line is -2 at the outlet store and $-3/2$ elsewhere. Thus, the first-quality plate costs twice as much as a second at the outlet store but only 1.5 times as much at other stores.

2. *Use the relative price difference to explain why consumers buy relatively more seconds at the factory outlet*. Holding a consumer's income and tastes fixed, if the price of seconds rises relative to that of firsts (as we go from the factory outlet to other retail shops), most consumers will buy relatively more firsts. The substitution effect is unambiguous: If they received compensation so that their utilities remained constant, consumers would unambiguously substitute firsts for seconds. It is possible that the income effect could go in the other direction. However, as most consumers spend relatively little of their total budget on plates, the income effect is presumably small relative to the substitution effect. Thus, we expect consumers to buy relatively fewer seconds at the retail stores than at the factory outlet.

Income and Substitution Effects with an Inferior Good

Giffen good
a commodity for which a decrease in its price causes the quantity demanded to fall

If a good is inferior, the income effect and the substitution effect move in opposite directions. For most inferior goods, the income effect is smaller than the substitution effect. As a result, the total effect moves in the same direction as the substitution effect, but the total effect is smaller. However, if the income effect more than offsets the substitution effect, we have a **Giffen good**, for which a decrease in its price causes the quantity demanded to fall.[9]

Jensen and Miller (2008) showed that rice is a Giffen good in Hunan, China. Ximing, a typical inhabitant of Hunan, views rice as a Giffen good. A fall in the rice price saves him money that he spends on other goods. Indeed, he decides to increase his spending on other goods even further by buying less rice. Thus, his demand curve for this Giffen good has an *upward* slope.

However, the Law of Demand (Chapter 2) says that demand curves slope downward. You're no doubt wondering how I'm going to worm my way out of this apparent contradiction. The answer is that I said that we believe in the Law of Demand because we do not observe upward sloping demand curves, not because our theory says that it must be true. It is theoretically possible for a demand curve to slope upward, but other than rice consumption in Hunan, China, economists have found few real-world examples of Giffen goods.[10]

Solved Problem 5.5

MyLab Economics Solved Problem

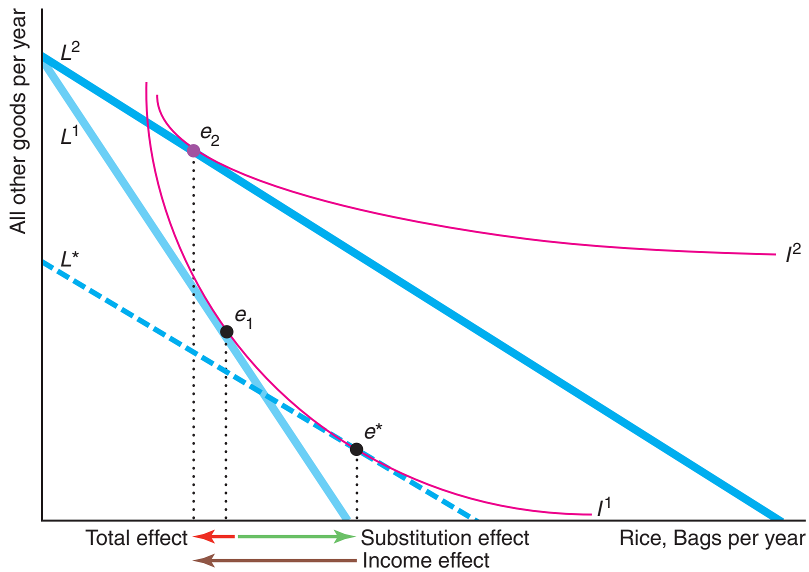
Ximing spends his money on rice, a Giffen good, and all other goods. Show that when the price of rice falls, Ximing buys less rice. Decompose this total effect of a price change on his rice consumption into a substitution effect and an income effect.

Answer

1. *Determine Ximing's original optimal bundle, e_1, using the tangency between his original budget line and one of his indifference curves*. In the figure, his original budget line, L^1, is tangent to his indifference curve I^1 at e_1.

[9] Robert Giffen, a nineteenth-century British economist, argued that poor people in Ireland increased their consumption of potatoes when the price rose because of a blight. However, recent studies of the Irish potato famine dispute this observation.

[10] However, Battalio, Kagel, and Kogut (1991) conducted an experiment that showed that quinine water is a Giffen good for lab rats.



2. *Show how the optimal bundle changes from a drop in the price of rice*. As the price of rice drops, his new budget line, L^2, becomes flatter, rotating around the original budget line's intercept on the vertical axis. The tangency between L^2 and indifference curve I^2 occurs at e_2, where Ximing consumes less rice than before because rice is a Giffen good.

3. *Draw a new, hypothetical budget line, L^*, that is based on the new price that keeps Ximing on the original indifference curve*. Ximing's opportunity set grows when the rice price falls. To keep him on his original indifference curve, his income would have to fall by enough so that his new budget line, L^2, shifts down to L^*, which is tangent to his original indifference curve I^1 at e^*.

4. *Identify the substitution and income effects*. The substitution effect is the change in rice consumption shown by the movement from e_1 to e^*: Ximing buys more rice when the price of rice drops, but he remains on his original indifference curve. The movement from e^* to e_2 determines the income effect: Ximing buys less rice as his income increases, holding prices constant. The total effect, the movement from e_1 to e_2, is the sum of the substitution effect that increases rice purchases and the income effect that reduces them. Ximing buys less rice because the income effect is larger than the substitution effect.

★ Compensating Variation and Equivalent Variation

In Figure 5.4 and in the figure in Solved Problem 5.5, we determine the substitution and income effects using an imaginary budget constraint, L^*, that is parallel to the new budget constraint, L^2, and tangent to the original indifference curve, I^1. We are calculating the **compensating variation** (CV), which is the amount of money one would have to *give* to a consumer to offset the harm from a price increase or *take* from a consumer to offset the benefit from a price decrease. That is, in Figure 5.4, the compensating variation keeps the consumer on the original indifference curve, I^1. Economists call this measure of the harm from a price increase the compensating variation because we give money to the consumer to compensate the consumer.

compensating variation (*CV*)
the amount of money one would have to *give* to a consumer to offset the harm from a price increase or *take* from a consumer to offset the benefit from a price decrease

equivalent variation (*EV*) the amount of money one would have to *take* from a consumer to harm the consumer by as much as the price increase or *give* to a consumer to benefit the consumer by as much as a price increase

Alternatively, we could draw L^* so that it was parallel to L^1 and tangent to I^2 to determine the substitution and income effects. By so doing, we calculate the **equivalent variation** (EV), which is the amount of money one would have to *take* from a consumer to harm the consumer by as much as the price increase or *give* to a consumer to benefit the consumer by as much as a price increase. That is, the equivalent variation moves the consumer to the new indifference curve, I^2. Economists call this measure of the harm from a price increase the equivalent variation because it causes the same or equivalent harm.

Application

What's Your Smart Phone Worth to You?

What value do you place on surfing the Internet using your smart phone, tablet, or wearable? In 2015, the Boston Consulting Group (BCG) surveyed consumers in 13 wealthy countries about their compensated variation, CV, and equivalent variation, EV (although the survey did not use those terms).

The surveyors asked consumers how much money they would want in exchange for giving up mobile Internet access: their compensating variation. Across the 13 countries, the average CV was \$4,000, seven times what consumers pay for the device and access. For all the consumers in these countries, the total CV is about \$3.5 trillion.

The survey also asked an equivalent variation question: What would you give up for a year to maintain your mobile Internet connection? That is, giving up what good or activity would hurt you as much as losing the Internet? To keep mobile Internet access, 74% of consumers would give up newspapers, 70% chocolate, 67% alcohol, 62% books, 49% television, 44% exercise, 27% sex, and 19% showering.

Many economists question whether answers to survey questions are reliable. For this reason, economists typically calculate consumer surplus, compensating variation, and equivalent variation by using estimated demand curves based on observed purchasing behavior.

5.4 Cost-of-Living Adjustments

By knowing both the substitution and income effects, we can answer questions that we couldn't if we knew only the total effect. For example, if firms have an estimate of the income effect, they can predict the impact of a negative income tax (a gift of money from the government) on the consumption of their products. Similarly, if we know the size of both effects, we can determine how accurately the government measures inflation.

Many long-term business contracts and government programs include *cost-of-living adjustments* (*COLAs*), which raise prices or incomes in proportion to an index of inflation. Rental contracts, alimony payments, salaries, pensions, and Social Security payments are frequently adjusted in this manner over time. We will use consumer theory to show that the most commonly used cost-of-living measure overestimates how the true cost of living changes over time. Because of this overestimate, you overpay your landlord if the rent on your apartment rises with this measure.

Inflation Indexes

The prices of most goods rise over time. We refer to the percentage increase in the overall price level as *inflation*.

Real Versus Nominal Prices Economists refer to the actual price of a good as the *nominal price*. The price adjusted for inflation is the *real price*.

Because the overall level of prices rises over time, nominal prices usually increase more rapidly than do real prices. For example, the nominal price of a McDonald's hamburger rose from 15¢ in 1955 to $1.10 in 2016, which is more than a sevenfold increase. However, the real price of a burger fell because the prices of other goods rose more rapidly than that of a burger.

How do we adjust for inflation to calculate the real price? Governments measure the cost of a standard bundle of goods for use in comparing prices over time using the Consumer Price Index (CPI). Each month, the government reports how much it costs to buy the bundle of goods that an average consumer purchased in a *base* year (with the base year changing every few years).

By comparing the cost of buying this bundle over time, we can determine how much the overall price level has increased. In the United States, the CPI was 26.8 in 1955 and 240.2 in May 2016.[11] The cost of buying the bundle of goods increased 896% ($\approx$ 240.2/26.8) from 1955 to 2016.

We can use the CPI to calculate the real price of a hamburger over time. In terms of 2016 dollars, the real price of a hamburger in 1955 was

$$\frac{\text{CPI for 2016}}{\text{CPI for 1955}} \times \text{price of a burger} = \frac{240.2}{26.8} \times 15¢ \approx 1.34.$$

That is, if you could have purchased the hamburger in 1955 with 2016 dollars—which are worth less than 1955 dollars—the hamburger would have cost $1.34. But, the actual price of a hamburger in 2016 was only $1.10. Thus, the price of a hamburger fell by nearly one-fifth. If we compared the real prices in both years using 1955 dollars, we would reach the same conclusion that the price of hamburgers fell by about one-fifth.

Calculating Inflation Indexes The government collects data on the quantities and prices of 364 individual goods and services, such as housing, dental services, watch and jewelry repairs, college tuition fees, taxi fares, women's hairpieces and wigs, hearing aids, slipcovers and decorative pillows, bananas, pork sausage, and funeral expenses. These prices rise at different rates. If the government merely reported all these price increases separately, most of us would find this information overwhelming. It is much more convenient to use a single summary statistic, the CPI, which tells us how prices rose *on average*.

We can use an example with only two goods, clothing and food, to show how the CPI is calculated. In the first year, consumers buy C_1 units of clothing and F_1 units of food at prices p_C^1 and p_F^1. We use this bundle of goods, C_1 and F_1, as our base bundle for comparison. In the second year, consumers buy C_2 and F_2 units at prices p_C^2 and p_F^2.

The government knows from its survey of prices each year that the price of clothing in the second year is p_C^2/p_C^1 times as large as the price the previous year and the price of food is p_F^2/p_F^1 times as large. If the price of clothing was $1 in the first year and $2 in the second year, the price of clothing in the second year is $\frac{2}{1} = 2$ times, or 100%, larger than in the first year.

One way we can average the price increases of each good is to weight them equally. But do we really want to do that? Do we want to give as much weight to the price increase for skateboards as to the price increase for automobiles? An alternative approach is to give a larger weight to the price change of a good as we spend more

[11]The CPI number 240.2 is not an actual dollar amount. Rather, it is the price of the bundle in the relevant period divided by the price during the 1982–1984 base period times 100. Thus, the CPI for 1982–1984 is 100.

of our income on that good—its budget share. The CPI takes this approach to weighting, using budget shares.[12]

The CPI for the first year is the amount of income it takes to buy the market basket actually purchased that year:

$$Y_1 = p_C^1 C_1 + p_F^1 F_1. \tag{5.1}$$

The cost of buying the first year's bundle in the second year is

$$Y_2 = p_C^2 C_1 + p_F^2 F_1. \tag{5.2}$$

To calculate the rate of inflation, we determine how much more income it would take to buy the first year's bundle in the second year, which is the ratio of Equation 5.1 to Equation 5.2:

$$\frac{Y_2}{Y_1} = \frac{p_C^2 C_1 + p_F^2 F_1}{p_C^1 C_1 + p_F^1 F_1}.$$

For example, from May 2015 to May 2016, the U.S. CPI rose from $Y_1 = 237.8$ to $Y_2 = 240.2$, so $Y_2/Y_1 \approx 1.01$. Thus, it cost 1% more in 2016 than in 2015 to buy the same bundle of goods.

The ratio Y_2/Y_1 reflects how much prices rise on average. By multiplying and dividing the first term in the numerator by p_C^1 and multiplying and dividing the second term by p_F^1, we find that this index is equivalent to

$$\frac{Y_2}{Y_1} = \frac{\left(\frac{p_C^2}{p_C^1}\right)p_C^1 C_1 + \left(\frac{p_F^2}{p_F^1}\right)p_F^1 F_1}{Y_1} = \left(\frac{p_C^2}{p_C^1}\right)\theta_C + \left(\frac{p_F^2}{p_F^1}\right)\theta_F,$$

where $\theta_C = p_C^1 C_1/Y_1$ and $\theta_F = p_F^1 F_1/Y_1$ are the budget shares of clothing and food in the first or base year. The CPI is a *weighted average* of the price increase for each good, p_C^2/p_C^1 and p_F^2/p_F^1, where the weights are each good's budget share in the base year, θ_C and θ_F.

Effects of Inflation Adjustments

A CPI adjustment of prices in a long-term contract overcompensates for inflation. We use an example involving an employment contract to illustrate the difference between using the CPI to adjust a long-term contract and using a true cost-of-living adjustment, which holds utility constant.

CPI Adjustment Klaas signed a long-term contract when he was hired. According to the COLA clause in his contract, his employer increases his salary each year by the same percentage as that by which the CPI increases. If the CPI this year is 5% higher than the CPI last year, Klaas' salary rises automatically by 5% over last year's.

Klaas spends all his money on clothing and food. His budget constraint in the first year is $Y_1 = p_C^1 C + p_F^1 F$, which we rewrite as

$$C = \frac{Y_1}{p_C^1} - \frac{p_F^1}{p_C^1}F.$$

[12]We've simplified this discussion of the CPI in a number of ways. The government makes sophisticated adjustments to the CPI that we ignore here, including repeated updating of the base year (chaining).

The intercept of the budget constraint, L^1, on the vertical (clothing) axis in Figure 5.5 is Y_1/p_C^1, and the slope of the constraint is $-p_F^1/p_C^1$. The tangency of his indifference curve I^1 and the budget constraint L^1 determine his optimal consumption bundle in the first year, e_1, where he purchases C_1 and F_1.

In the second year, his salary rises with the CPI to Y_2, so his budget constraint, L^2, in that year is

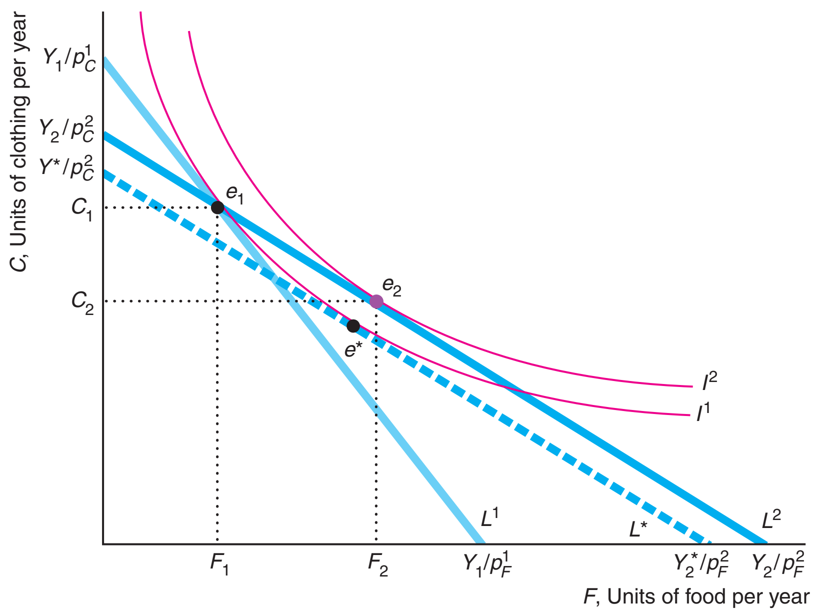
$$C = \frac{Y_2}{p_C^2} - \frac{p_F^2}{p_C^2}F.$$

The new constraint, L^2, has a flatter slope, $-p_F^2/p_C^2$, than L^1 because we assume that the price of clothing rose more than the price of food. The new constraint goes through the original optimal bundle, e_1, because, by increasing his salary using the CPI, the firm ensures that Klaas can buy the same bundle of goods in the second year that he chose in the first year.

He *can* buy the same bundle, but *does* he? The answer is no. His optimal bundle in the second year is e_2, where indifference curve I^2 is tangent to his new budget constraint, L^2. The movement from e_1 to e_2 is the *total effect* from the changes in the real prices of clothing and food. *This adjustment to his income does not keep him on his original indifference curve*, I^1.

Figure 5.5 The Consumer Price Index

In the first year, when Klaas has an income of Y_1, his optimal bundle is e_1, where indifference curve I^1 is tangent to his budget constraint, L^1. In the second year, the price of clothing rises more than the price of food. Because his salary increases in proportion to the CPI, his second-year budget constraint, L^2, goes through e_1, so he can buy the same bundle as in the first year. His new optimal bundle, however, is e_2, where I^2 is tangent to L^2. The CPI adjustment overcompensates him for the increase in prices: Klaas is better off in the second year because his utility is greater on I^2 than on I^1. With a smaller true cost-of-living adjustment, Klaas' budget constraint, L^*, is tangent to I^1 at e^*.



Indeed, Klaas is better off in the second year than in the first. The CPI adjustment *overcompensates* for the change in inflation in the sense that his utility increases.

Klaas is better off because the prices of clothing and food did not increase by the same amount. Suppose that the price of clothing and food had both increased by *exactly* the same amount. After a CPI adjustment, Klaas' budget constraint in the second year, L^2, would be the same as in the first year, L^1, so he would choose exactly the same bundle, e_1, in the second year as in the first year.

Because the price of food rose by less than the price of clothing, L^2 is not the same as L^1. Food became cheaper relative to clothing, so by consuming more food and less clothing Klaas has higher utility in the second year.

Had clothing become relatively less expensive, Klaas would have raised his utility in the second year by consuming relatively more clothing. Thus, it doesn't matter which good becomes relatively less expensive over time—it's only necessary for one of them to become a relative bargain for Klaas to benefit from the CPI compensation.

Application

Paying Employees to Relocate

International firms are increasingly relocating workers throughout their home countries and internationally. Companies that frequently relocate workers include investment banks such as Goldman Sachs, multinational corporations such as Google, and consulting firms such as Bain. A 2016 survey of 163 global companies found that 61% of the companies told employees that such assignments were important to their careers and that about 10% of international assignees were Millennials.

As you might expect, some workers do not want to relocate. In a survey by Runzheimer International, 79% of relocation managers responded that they confronted resistance from employees who were asked to relocate to high-cost areas.

One common approach to entice employees to relocate is for the firm to assess the goods and services consumed by employees in the original location and then offer to pay those employees enough to allow them to consume essentially the same items in the new location. According to a survey by Mercer, 79% of international firms reported that they provided their workers with enough income abroad to maintain their home lifestyle based on CPI-type indexes.

If employees are indifferent about where they live as long as they can buy the goods they want, then these firms are overcompensating them. We can show this result using Figure 5.5. In the figure, L^1 is a worker's budget constraint in the home country, and e_1 is the worker's optimal bundle. The firm's compensation allows the worker to buy the same bundle, e_1, on the new budget line, L^2, in the new country. The worker is better off at e_2 than at e_1. A smaller compensation that would shift the budget line from L^1 to L^* would be sufficient to induce the worker to move.

★ **True Cost-of-Living Adjustment** We now know that a CPI adjustment overcompensates for inflation. What we want is a *true cost-of-living index*: an inflation index that holds utility constant over time.

How big an increase in Klaas' salary would leave him exactly as well off in the second year as in the first? We can answer this question by applying the same technique we use to identify the substitution and income effects. We draw an imaginary budget line, L^* in Figure 5.5, that is tangent to I^1, so that Klaas' utility remains constant but has the same slope as L^2. The income, Y^*, corresponding to that imaginary budget constraint, is the amount that leaves Klaas' utility constant. Had Klaas received

Y^* in the second year instead of Y_2, he would have chosen Bundle e^* instead of e_2. Because e^* is on the same indifference curve, I^1, as e_1, Klaas' utility would be the same in both years.

The COLA example in Table 5.1 illustrates how the CPI overcompensates Klaas.[13] Suppose that p_C^1 is \$1, p_C^2 is \$2, p_F^1 is \$4, and p_F^2 is \$5. In the first year, Klaas spends his income, Y_1, of \$400 on $C_1 = 200$ units of clothing and $F_1 = 50$ units of food and has a utility of 2,000, which is the level of utility on I^1. If his income did not increase in the second year, he would substitute toward the relatively inexpensive food, cutting his consumption of clothing in half but reducing his consumption of food by only a fifth. His utility would fall to 1,265.

If his second-year income increases in proportion to the CPI, he can buy the same bundle, e_1, in the second year as in the first. His second-year income is $Y_2 = \$650$ $(= p_C^2 C_1 + p_F^2 F_1 = [\$2 \times 200] + [\$5 \times 50])$. Klaas is better off if his budget increases to Y_2. He substitutes toward the relatively inexpensive food, buying less clothing than in the first year but more food, e_2. His utility rises from 2,000 to approximately 2,055 (the level of utility on I^2).

How much would his income have to rise to leave him only as well off as he was in the first year? If his second-year income is $Y^* \approx \$632.50$, by appropriate substitution toward food, e^*, he can achieve the same level of utility, 2,000, as in the first year.

We can use the income that just compensates Klaas, Y^*, to construct a true cost-of-living index. In this example, the true cost-of-living index rose 58.1% ($\approx [632.50 - 400]/400$), while the CPI rose 62.5% ($= [650 - 400]/400$).

Size of the CPI Substitution Bias We have just demonstrated that the CPI has an *upward bias*: an individual's utility rises if we increase that person's income by the same percentage as the CPI rises. If we make the CPI adjustment, we are implicitly assuming—incorrectly—that consumers do not substitute toward relatively inexpensive goods when prices change but keep buying the same bundle of goods over time. We call this overcompensation a *substitution bias*.

The CPI calculates the increase in prices as Y_2/Y_1. We can rewrite this expression as

$$\frac{Y_2}{Y_1} = \frac{Y^*}{Y_1}\frac{Y_2}{Y^*}.$$

The first term to the right of the equal sign, Y^*/Y_1, is the increase in the true cost of living. The second term, Y_2/Y^*, reflects the substitution bias in the CPI. This term is greater than one because $Y_2 > Y^*$. In the example in Table 5.1, $Y_2/Y^* = 650/632.50 \approx 1.028$, so the CPI overestimates the increase in the cost of

Table 5.1 Cost-of-Living Adjustments

	p_C	p_F	Income, Y	Clothing	Food	Utility, U
First year	\$1	\$4	$Y_1 = \$400$	200	50	2,000
Second year	\$2	\$5				
No adjustment			$Y_1 = \$400$	100	40	≈ 1,265
CPI adjustment			$Y_2 = \$650$	162.5	65	≈ 2,055
True COLA			$Y^* \approx \$632.50$	≈ 158.1	≈ 63.2	2,000

[13] In Table 5.1 and Figure 5.5, we assume that Klaas has a utility function $U = 20\sqrt{CF}$.

living by about 2.8%. A number of studies estimate that the U.S. substitution bias was at least 0.5%. However, in recent years, the government has modified its CPI formula to reduce the substitution bias.

No substitution bias occurs if all prices increase at the same rate so that relative prices remain constant. The faster some prices rise relative to others, the more pronounced is the upward bias from consumers substituting the now less expensive goods.

5.5 Deriving Labor Supply Curves

The human race faces a cruel choice: work or daytime television.

Throughout this chapter, we've used consumer theory to examine consumers' *demand* behavior. Perhaps surprisingly, we can also use the consumer theory model to derive the *supply curve* of labor. We will to do that by deriving a demand curve for time spent *not* working. We then use that demand curve to determine the supply curve of hours spent working.

Labor-Leisure Choice

People choose between working to earn money to buy goods and services and consuming *leisure*: time spent not working. In addition to sleeping, eating, and playing, leisure includes time spent cooking meals and fixing things around the house. The number of hours worked per day, H, equals 24 minus the hours of leisure or nonwork, N, in a day:

$$H = 24 - N.$$

If I get less than 8 hours sleep, I stay awake for more than 16 hours.

Using consumer theory, we can determine the demand curve for leisure once we know the price of leisure. What does it cost you to watch TV or go to school or do anything for an hour other than work? It costs you the wage, w, you could have earned from an hour's work: The price of leisure is forgone earnings. The higher your wage, the more an hour of leisure costs you. For this reason, taking an afternoon off costs a lawyer who earns \$250 an hour much more than it costs a minimum wage worker.

We use an example to show how the number of hours of leisure and work depends on the wage, unearned income (such as inheritances and gifts from parents), and tastes. Jackie spends her total income, Y, on various goods. For simplicity, we assume that the price of these goods is \$1 per unit, so she buys Y goods. Her utility, U, depends on how many goods and how much leisure she consumes:

$$U = U(Y, N).$$

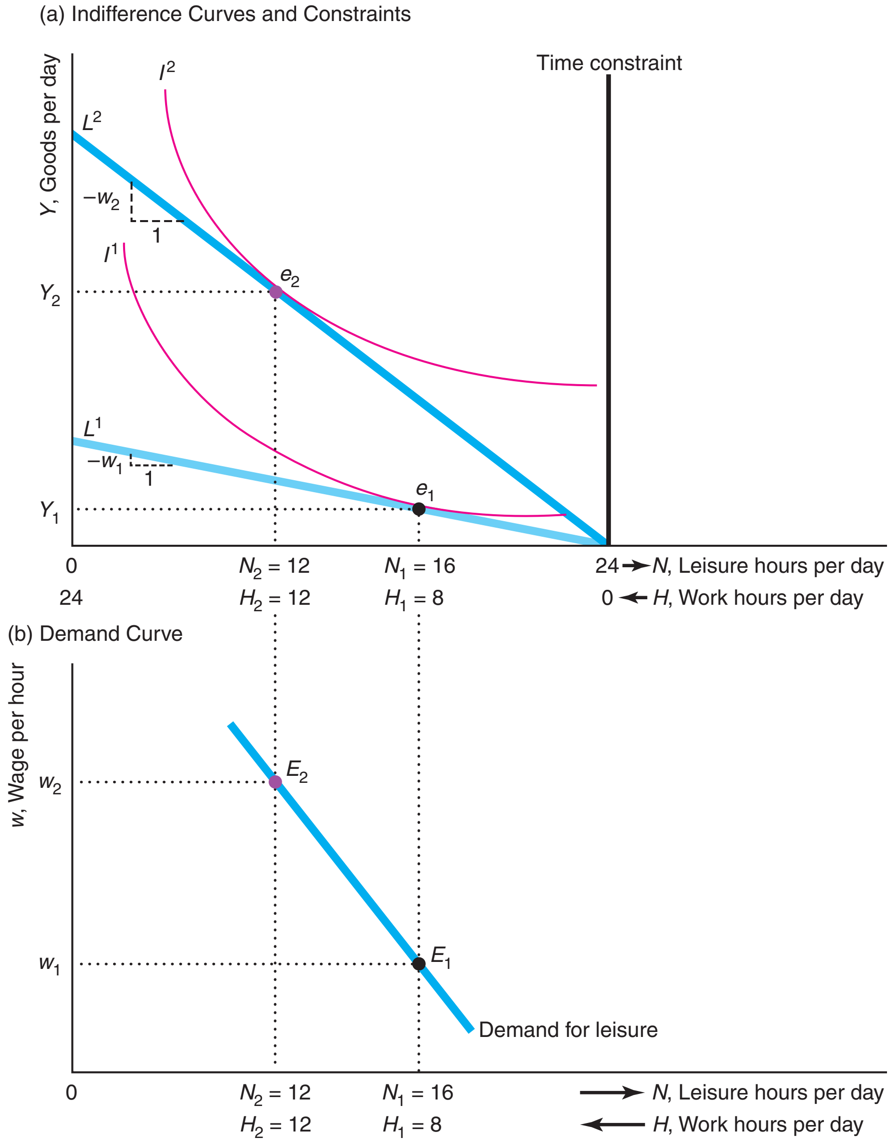
Initially, we assume that Jackie can choose to work as many or as few hours as she wants for an hourly wage of w. Jackie's earned income equals her wage times the number of hours she works, wH. Her total income, Y, is her earned income plus her unearned income, Y^*:

$$Y = wH + Y^*.$$

Panel a of Figure 5.6 shows Jackie's choice between leisure and goods. The vertical axis shows how many goods, Y, Jackie buys. The horizontal axis shows both

Figure 5.6 Demand for Leisure

(a) Jackie chooses between leisure, N, and other goods, Y, subject to a time constraint (vertical line at 24 hours) and a budget constraint, L^1, which is $Y = w_1H = w_1(24 - N)$, with a slope of $-w_1$. The tangency of her indifference curve, I^1, with her budget constraint, L^1, determines her optimal bundle, e_1, where she has $N_1 = 16$ hours of leisure and works $H_1 = 24 - N_1 = 8$ hours. If her wage rises from w_1 to w_2, Jackie shifts from optimal bundle e_1 to e_2. (b) Bundles e_1 and e_2 correspond to E_1 and E_2 on her leisure demand curve.



hours of leisure, N, measured from left to right, and hours of work, H, measured from right to left. Jackie maximizes her utility given the *two* constraints she faces. First, she faces a time constraint, which is a vertical line at 24 hours of leisure. The number of hours in a day is fixed at 24, so all the money in the world won't buy her more hours in a day. Second, Jackie faces a budget constraint. Because Jackie has no unearned income, her initial budget constraint, L^1, is $Y = w_1H = w_1(24 - N)$.

The slope of her budget constraint is $-w_1$, because each extra hour of leisure she consumes costs her w_1 goods.

Jackie picks her optimal hours of leisure, $N_1 = 16$, so that she is on the highest indifference curve, I^1, that touches her budget constraint. She works $H_1 = 24 - N_1 = 8$ hours per day and earns an income of $Y_1 = w_1 H_1 = 8w_1$.

We derive Jackie's demand curve for leisure using the same method that we used to derive Mimi's demand curve for beer. We raise the price of leisure—the wage—in panel a of Figure 5.6 to trace out Jackie's demand curve for leisure in panel b. As the wage increases from w_1 to w_2, leisure becomes more expensive, and Jackie demands less of it.

By subtracting her demand for leisure at each wage (on her demand curve for leisure in panel a of Figure 5.7) from 24, we construct her labor supply curve—the hours she is willing to work as a function of the wage—in panel b.[14] Thus, her supply curve for hours worked is the mirror image of the demand curve for leisure: For every extra hour of leisure that Jackie consumes, she works one hour less.

Income and Substitution Effects

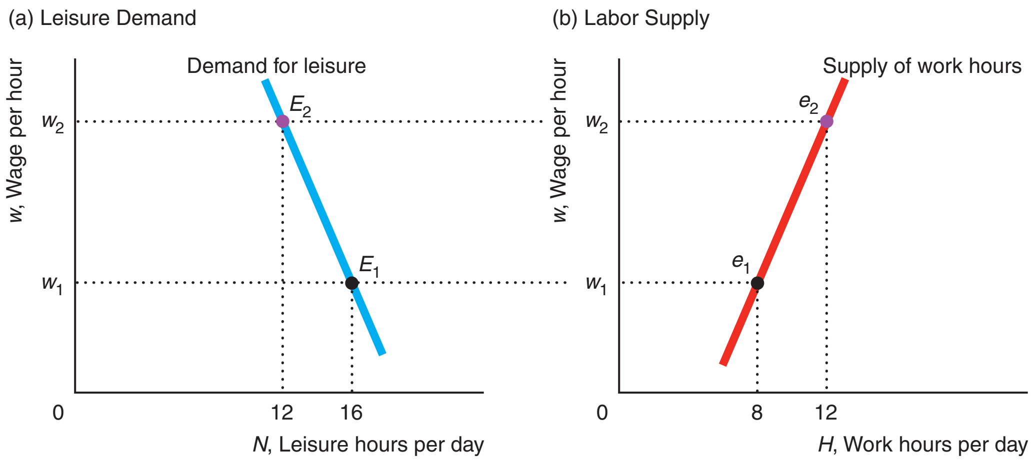
An increase in the wage causes both income and substitution effects, which alter an individual's demand for leisure and supply of hours worked. The *total effect* of an increase in Jackie's wage from w_1 to w_2 is the movement from e_1 to e_2 in Figure 5.8. Jackie works $H_2 - H_1$ fewer hours and consumes $N_2 - N_1$ more hours of leisure.

By drawing an imaginary budget constraint, L^*, that is tangent to her original indifference curve with the slope of the new wage, we can divide the total effect into substitution and income effects. The *substitution effect*, the movement from e_1 to e^*, must be negative: A compensated wage increase causes Jackie to consume fewer hours of leisure, N^*, and work more hours, H^*.

As the wage rises, if Jackie works the same number of hours as before, she has a higher income. The *income effect* is the movement from e^* to e_2. Because leisure

Figure 5.7 Supply Curve of Labor

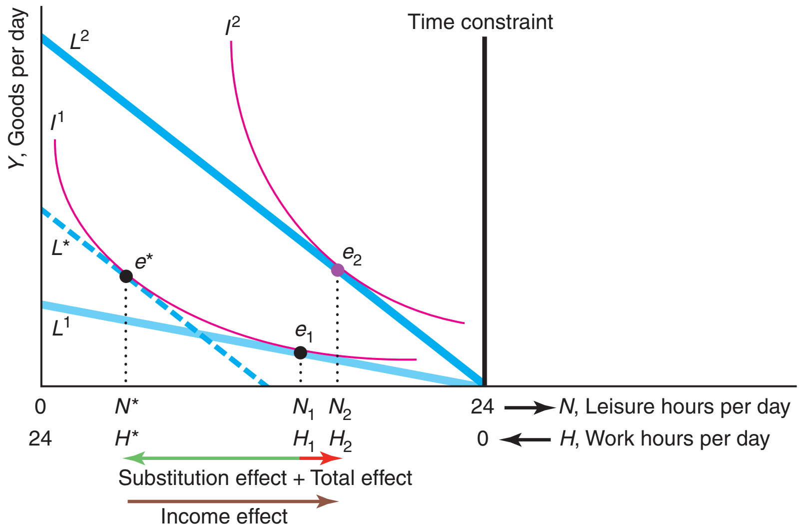
(a) Jackie's demand for leisure is downward sloping. (b) At any given wage, the number of hours that Jackie works, *H*, and the number of hours of leisure, *N*, that she consumes add to 24. Thus, her supply curve for hours worked, which equals 24 hours minus the number of hours of leisure she demands, is upward sloping.



[14]Appendix 5B shows how to derive the labor supply curve using calculus.

Figure 5.8 Income and Substitution Effects of a Wage Change

A wage change causes both a substitution and an income effect. The movement from e_1 to e^* is the substitution effect, the movement from e^* to e_2 is the income effect, and the movement from e_1 to e_2 is the total effect.



is a normal good for Jackie, as her income rises, she consumes more leisure. When leisure is a normal good, the substitution and income effects work in opposite directions, so whether leisure demand increases or not depends on which effect is larger. Jackie's income effect dominates the substitution effect, so the total effect for leisure is positive: $N_2 > N_1$. Jackie works fewer hours as the wage rises, so her labor supply curve is backward bending.

If leisure is an inferior good, both the substitution effect and the income effect work in the same direction, and hours of leisure definitely fall. As a result, if leisure is an inferior good, a wage increase unambiguously causes the hours worked to rise.

Solved Problem 5.6

Enrico receives a no-strings-attached scholarship that pays him an extra Y^* per day. How does this scholarship affect the number of hours he wants to work? Does his utility increase?

Answer

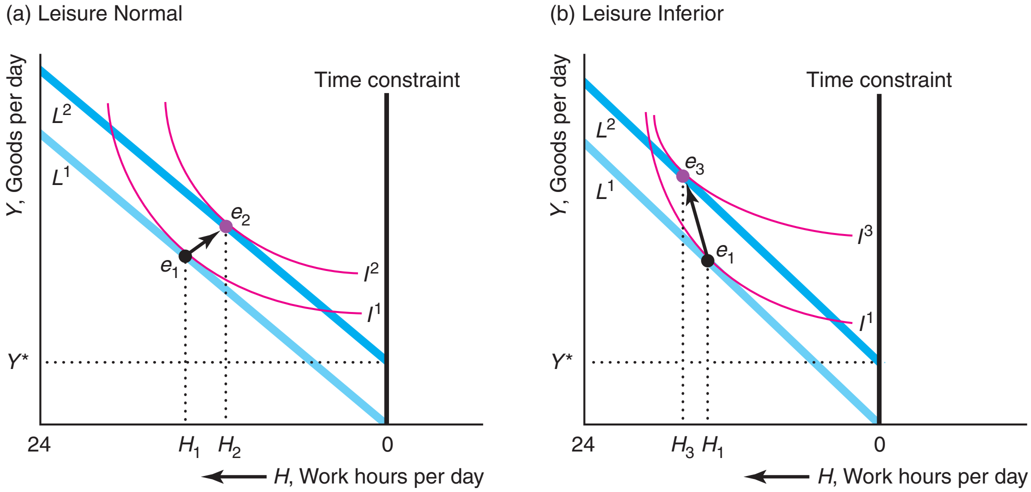
1. *Show his consumer optimum without unearned income.* When Enrico had no unearned income, his budget constraint, L^1 in the graphs, hit the hours-leisure axis at 0 hours and had a slope of $-w$.

2. *Show how the unearned income affects his budget constraint.* The extra income causes a parallel upward shift of his budget constraint by Y^*. His new budget constraint, L^2, has the same slope as before, $-w$, because his wage does not change. The extra income cannot buy Enrico more time, of course, so L^2 cannot extend to the right of the time constraint. It hits the vertical time constraint at Y^*: If he works no hours, he has Y^* income.

3. *Show that the relative position of the new to the original optimum depends on his tastes.* The change in the number of hours he works depends on Enrico's tastes. Panels a and b show two possible sets of indifference curves. In both diagrams, when facing budget constraint L^1, Enrico chooses to work H_1 hours. In panel a,

leisure is a normal good, so as his income rises, Enrico consumes more leisure than originally: He moves from Bundle e_1 to Bundle e_2. In panel b, he views leisure as an inferior good and consumes fewer hours of leisure than originally: He moves from e_1 to e_3. (Another possibility is that the number of hours he works is unaffected by the extra unearned income.)

4. *Discuss how his utility changes*. Regardless of his tastes, Enrico has more income in the new optimum and is on a higher indifference curve after receiving the scholarship. In short, he believes that more money is better than less.



Shape of the Labor Supply Curve

Whether the labor supply curve slopes upward, bends backward, or has sections with both properties depends on the income elasticity of leisure. Suppose that a worker views leisure as an inferior good at low wages and as a normal good at high wages. As the wage increases, the worker's demand for leisure first falls and then rises. Consequently, as the wage rises, the hours supplied to the market first rise and then fall. (Alternatively, the labor supply curve may slope upward and then backward even if leisure is normal at all wages: At low wages, the substitution effect—work more hours—dominates the income effect—work fewer hours—while the opposite occurs at higher wages.)

The budget line rotates upward from L^1 to L^2 as the wage rises in panel a of Figure 5.9. Because leisure is an inferior good at low incomes, in the new optimal bundle, e_2, this worker consumes less leisure and more goods than at the original bundle, e_1.

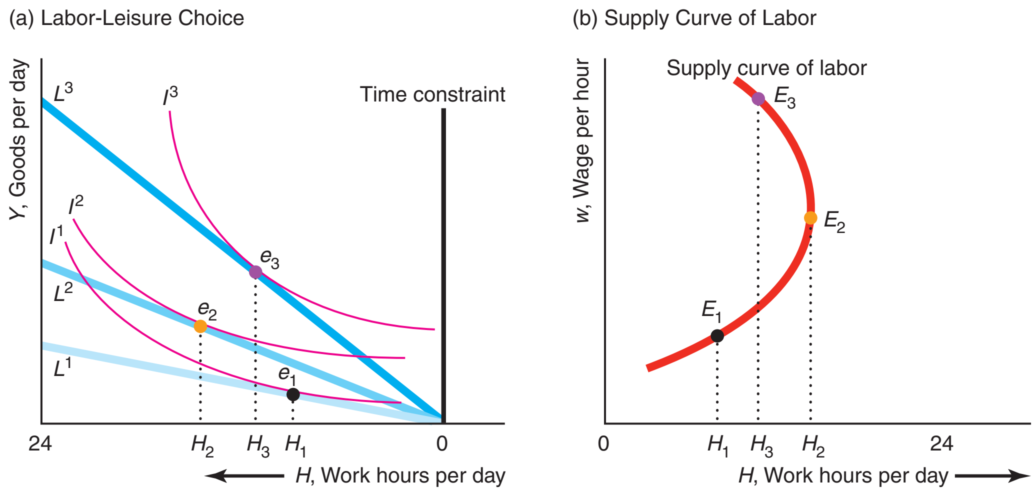
At higher incomes, however, leisure is a normal good. At an even higher wage, the new optimum is e_3 on budget line L^3, where the quantity of leisure demanded is higher and the number of hours worked is lower. Thus, the corresponding supply curve for labor slopes upward at low wages and bends backward at higher wages in panel b.

Do labor supply curves slope upward or backward? Economic theory alone cannot answer this question: Both forward-sloping and backward-bending supply curves are *theoretically* possible. Empirical research is necessary to resolve this question.

Most studies (see the survey in Keane, 2011) find that the labor supply curves for British and American men are relatively vertical because the income and substitution

Figure 5.9 Deriving a Labor Supply Curve That Slopes Upward and Then Bends Backward

At low incomes, an increase in the wage causes the worker to work more: the movement from e_1 to e_2 in panel a or from E_1 to E_2 in panel b. At higher incomes, an increase in the wage causes the worker to work fewer hours: the movement from e_2 to e_3 or from E_2 to E_3.



effects are offsetting or both are small. Economists have found similar results in other countries such as Japan (Kuroda and Yamamoto, 2008) and the Netherlands (Evers et al., 2008). Keane's average across all studies of males' pure substitution wage elasticity is about 0.31 (although most of the estimates are below 0.15). In contrast, most studies in Keane's survey found that females' long-run wage elasticity estimates range from 1.25 to 5.6.

Application

Working After Winning the Lottery

Would you stop working if you won a lottery or inherited a large sum? Economists want to know how unearned income affects the amount of labor people are willing to supply, because this question plays a crucial role in many government debates on taxes and welfare. For example, some legislators oppose negative income tax and welfare programs because they claim that giving money to poor people will stop them from working. Is that assertion true?

We could clearly answer this question if we could observe the behavior of a large group of people, only some of whom were randomly selected to receive varying but large amounts of unearned income each year for decades. Luckily for us, governments conduct such experiments by running lotteries.

Imbens et al. (2001) compared the behavior of winners of major prizes to that of other people who played the Massachusetts Megabucks lottery. Major prizes ranged from \$22,000 to \$9.7 million, with an average prize of \$1.1 million, which the government paid in yearly installments over two decades.

The average winner received \$55,200 in prize money per year and chose to work slightly fewer hours so that his or her labor earnings fell by \$1,877 per year. That is, winners increased their consumption and savings but did not substantially decrease how much they worked.

For every dollar of unearned income, winners reduced their work effort and hence their labor earnings by 11¢ an hour on average. Men and women, big and very big prize winners, and people of all education levels behaved the same way. However, the behavior of winners differed by age and by income groups. People ages 55 to 65 reduced their labor efforts by about a third more than younger people did, presumably because they decided to retire early. Most striking, people with no earnings in the year before winning the lottery tended to increase their labor earnings after winning.

Cesarini et al. (2015) examined the Swedish lotteries and found that winning a lottery prize reduces labor earnings by about 10% of the prize. According to a study of the Dutch lottery (Picchio et al., 2015), winners reduce their hours of work but are unlikely to withdraw from the labor market.

Income Tax Rates and Labor Supply

Why do we care about the shape of labor supply curves? One reason is that we can tell from the shape of the labor supply curve whether an increase in the income tax rate—a percent of earnings—will cause a substantial reduction in the hours of work.[15] Taxes on earnings are an unattractive way of collecting money for the government if supply curves are upward sloping, because the taxes cause people to work fewer hours, reducing the amount of goods society produces and raising less tax revenue than if the supply curve were vertical or backward bending. On the other hand, if supply curves are backward bending, a small increase in the tax rate may increase tax revenue *and* boosts total production (but reduces leisure). Consequently, many people overgeneralize about the effects of a tax cut.

Common Confusion: A cut in the income tax rate must raise government tax revenue.

This belief is not correct in general.[16] A tax cut may cause tax revenue to rise or fall depending on the initial tax rate.

Presidents John F. Kennedy, Ronald Reagan, and George W. Bush and other politicians have advocated cutting the marginal income tax rate (the percentage of the last dollar earned that the government takes in taxes) to stimulate people to work longer and produce more, which they believed would cause tax receipts to increase.[17] Because U.S. tax rates have changed substantially over time, we have a natural experiment with which to test this hypothesis.

The Kennedy-Johnson tax cuts lowered the top federal personal marginal tax rate from 91% to 70% and other rates fell too. The Reagan tax cuts lowered the top marginal tax rate to 50% in 1982, 38.5% in 1987, and 28% in 1988. The top rate rose to 31% in 1991 under President George H. W. Bush, a Republican. It rose to

---

[15] Although taxes are ancient, the income tax is a relatively recent invention. William Pitt the Younger introduced the British income tax (10% on annual incomes above £60) in 1798 to finance the war with Napoleon. The U.S. Congress followed suit in 1861, using the income taxes (3% on annual incomes over $800) to pay for the Civil War.

[16] A 2012 University of Chicago Booth School of Business poll of 40 distinguished economists (including Republicans, Independents, and Democrats) did not find a single economist who believed that a U.S. tax cut would cause tax revenue to rise.

[17] However, none of these politicians have been unwilling to emulate Lady Godiva's tax-fighting technique. Allegedly, her husband, Leofric, the Earl of Mercia, agreed to her request to eliminate taxes if she rode naked through the Coventry marketplace.

39.6% in 1993 under President Bill Clinton, a Democrat. Under President George W. Bush, a Republican, the top tax rate fell temporarily to 38.6% in 2001, 37.6% in 2004, and 35% in 2006. The George W. Bush tax cuts were to have expired under the administration of President Barack Obama, a Democrat. However, the cuts to the lower rates were maintained, while the top rate reduction expired so that rate returned to 39.6% in 2013.[18]

If the tax does not affect the pre-tax wage, the effect of imposing a marginal tax rate of $v = 28\% = 0.28$ is to reduce the effective wage from w to $(1 - v)w = 0.72w$.[19] The tax reduces the after-tax wage by 28%, so a worker's budget constraint rotates downward, similar to rotating the budget constraint downward from L^3 to L^2 or from L^2 to L^1, in Figure 5.9.

As that figure indicates, if the budget constraint rotates downward, the hours of work may increase or decrease, depending on whether leisure is a normal or an inferior good. The worker in panel b has a labor supply curve that at first slopes upward and then bends backward. If the worker's wage is very high, the worker is in the backward-bending section of the labor supply curve.

If so, the relationship between the marginal tax rate, v, and tax revenue, vwH, is bell-shaped, as in Figure 5.10. This figure is the estimated U.S. tax revenue curve (Trabandt and Uhlig, 2011, 2013). At the marginal rate for the typical person, $v = 28\%$, the government collects 100% of the amount of tax revenue that it is currently collecting. At a zero tax rate, a small increase in the tax rate *must* increase the tax revenue because the government collects no revenue if the tax rate is zero. However, if the tax rate rises a little more, the tax revenue collected must rise even higher, for two reasons: First, the government collects a larger percentage of every dollar earned because the tax rate is higher. Second, employees work more hours as the tax rate rises because workers are in the backward-bending sections of their labor supply curves.

As the marginal rate increases, tax revenue rises until the marginal rate reaches $v^* = 63\%$, where the U.S. tax revenue would be 130% of its current level.[20] If the marginal tax rate increases more, workers are in the upward-sloping sections of their labor supply curves, so an increase in the tax rate reduces the number of hours worked. When the tax rate rises high enough, the reduction in hours worked more than offsets the gain from the higher rate, so the tax revenue falls.

It makes little sense for a government to operate at very high marginal tax rates in the downward-sloping portion of this bell-shaped curve where the government could get more output *and* more tax revenue by cutting the marginal tax rate.

---

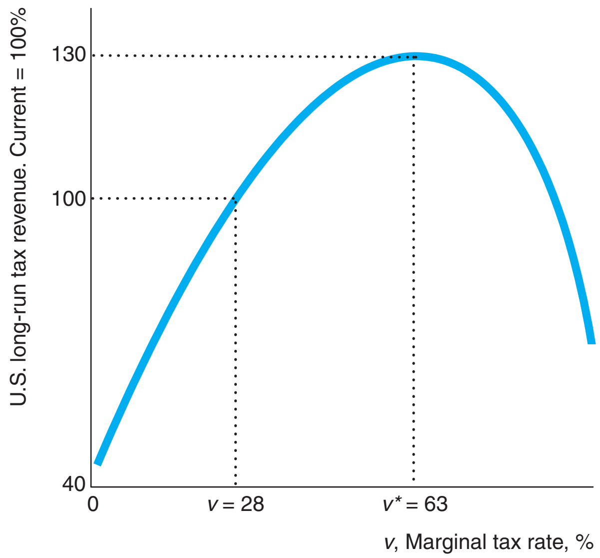
[18]Many other countries' central governments have also lowered their top marginal tax rates in recent years. The top U.K. rate fell sharply during the Thatcher administration from 83% to 60% in 1979 and to 40% in 1988. It rose to 50% in 2010, and fell to 45% in April 2013. Japan's top rate fell from 75% in 1983 to 60% in 1987, 50% in 1988, and to 37% in 1999, but it rose to 40% in 2007, and was 45% in 2015. In 1988, Canada raised the marginal tax rates for the two lowest income groups and lowered them for those falling into the top nine brackets. Canada's top federal rate was 29% in 2015.

[19]Under a progressive income tax system, the marginal tax rate increases with income, and the marginal tax rate is greater than the average tax rate. Suppose that the marginal tax rate is 20% on the first \$10,000 earned and 30% on the second \$10,000. Someone who earns \$20,000 pays a tax of \$2,000 on the first \$10,000 of earnings and \$3,000 on the next \$10,000. That taxpayer's average tax rate is 25% (= [\$2,000 + \$3,000]/\$20,000). In 2012, the U.S. marginal tax rate on a single person with a taxable income of \$50,000 was 25%, while the average rate was 17.06%. For simplicity, in the following analysis, we assume that the marginal tax rate is a constant, v, so the average tax rate is also v.

[20]On average for 14 European Union countries, v is also less than v^*, but raising the rate to v^* would raise European tax revenue by only 8% (Trabandt and Uhlig, 2011).

Figure 5.10 The Relationship of U.S. Tax Revenue to the Marginal Tax Rate

This curve shows how U.S. income tax revenue varies with the marginal income tax rate, v, according to Trabandt and Uhlig (2011, 2013). The typical person pays $v = 28\%$, which corresponds to 100% of the current tax revenue that the government collects. The tax revenue would be maximized at 130% of its current level if the marginal rate were set at $v^* = 63\%$. For rates below v^*, a higher marginal rate results in increased tax revenue. However, at rates above v^*, an increase in the marginal rate decreases tax revenue.



Challenge Solution

Per-Hour Versus Lump-Sum Childcare Subsidies

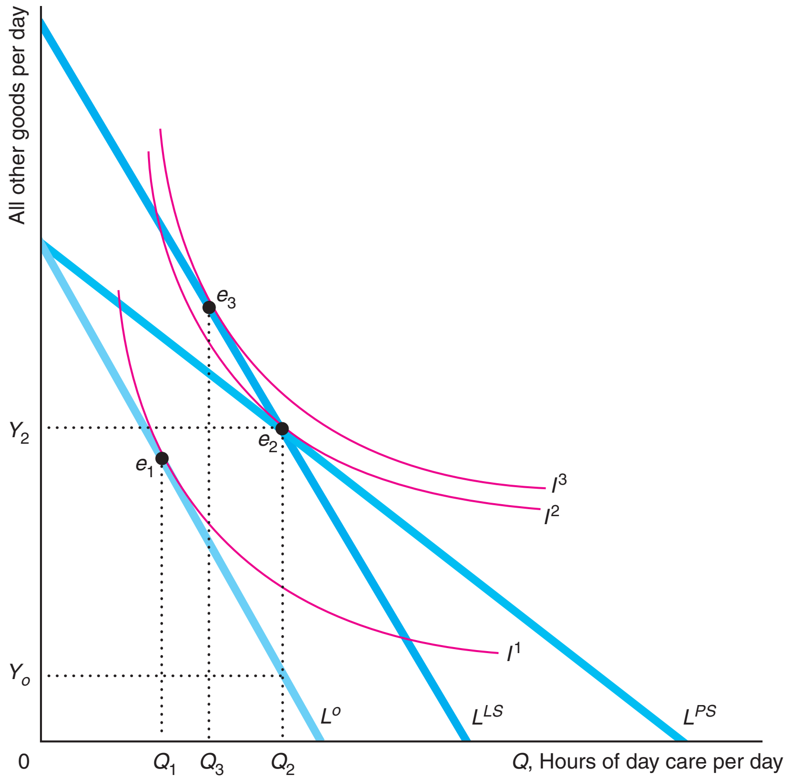
We now return to the questions raised at the beginning of the chapter: For a given government expenditure, does a childcare per-hour subsidy or a lump-sum subsidy provide greater benefit to recipients? Which option increases the demand for childcare services by more? Which one inflicts less cost on other consumers of childcare?

To determine which program benefits recipients more, we employ a model of consumer choice. The figure shows a poor family that chooses between *hours of childcare per day* (Q) and *all other goods per day*. Given that the price of all other goods is $1 per unit, the expenditure on all other goods is the income, Y, not spent on childcare. The family's initial budget constraint is L^o. The family chooses Bundle e_1 on indifference curve I^1, where the family consumes Q_1 hours of childcare services.

If the government gives a childcare price subsidy, the new budget line, L^{PS}, rotates out along the childcare axis. Now the family consumes Bundle e_2 on (higher) indifference curve I^2. The family consumes more hours of childcare, Q_2, because childcare is now less expensive and it is a normal good.

One way to measure the value of the subsidy the family receives is to calculate how many *other goods* the family could buy before and after the subsidy. If the family consumes Q_2 hours of childcare, the family could have consumed Y_o other goods with the original budget constraint and Y_2 with the price-subsidy budget constraint. Given that Y_2 is the family's remaining income after paying for childcare, the family buys Y_2 units of all other goods. Thus, the value to the family of the childcare price subsidy is $Y_2 - Y_o$.

If, instead of receiving a childcare price subsidy, the family were to receive a lump-sum payment of $Y_2 - Y_o$, taxpayers' costs for the two programs would be the same. The family's budget constraint after receiving a lump-sum payment, L^{LS}, has the same slope as the original one, L^o, because the relative prices of childcare and all other goods are the same as originally (see Section 4.3). This budget constraint must go through e_2 because the family has just enough money to buy that bundle. However,



given this budget constraint, the family would be better off if it buys Bundle e_3 on indifference curve I^3 (the reasoning is the same as that in the Consumer Price Index analysis in Figure 5.5). The family consumes less childcare with the lump-sum subsidy: Q_3 rather than Q_2.

Poor families prefer the lump-sum payment to the price subsidy because indifference curve I^3 is above I^2. Taxpayers are indifferent between the two programs because they both cost the same. The childcare industry prefers the price subsidy because the demand curve for its service is farther to the right: At any given price, poor families who receive a price subsidy rather than a lump-sum subsidy demand more childcare.

Given that most of the directly affected groups benefit more from lump-sum payments than price subsidies, why does government more commonly use price subsidies? One possible explanation is that the childcare industry lobbied effectively for price subsidies, but that is not true. Second, politicians might believe that poor families will not make the best choices about childcare, so they might see price subsidies as a way of getting such families to consume relatively more (or better-quality) childcare than they would otherwise choose. Third, politicians may prefer that poor people consume more childcare so that they can work more hours, thereby increasing society's wealth. Fourth, politicians may not understand this analysis.

Summary

1. **Deriving Demand Curves.** We can derive individual demand curves by using the information about tastes contained in a consumer's indifference curve map. By adjusting the price of one good, while holding other prices and income constant, we find how the quantity demanded varies with changes in its price, which is the information we need to draw the demand curve. Indifference curves, which reflect consumers' tastes, determine the shape of the demand curve.

2. **How Changes in Income Shift Demand Curves.** The entire demand curve shifts as a consumer's income rises. By varying income, holding prices constant, we show how quantity demanded shifts with income. An Engel curve summarizes the relationship between income and quantity demanded, holding prices constant.

3. **Effects of a Price Change.** An increase in the price of a good causes both a substitution effect and an income effect. The *substitution effect* is the amount by which a consumer's demand for the good changes as a result of a price increase when we compensate the consumer for the price increase by raising the individual's income by enough that his or her utility does not change. The substitution effect is unambiguous: A compensated rise in a good's price *always* causes consumers to buy less of that good. The *income effect*

shows how a consumer's demand for a good changes as the consumer's income falls. The price rise reduces the consumer's opportunities, because the consumer can buy less than before with the same income. The income effect can be positive or negative. If a good is normal (income elasticity is positive), then the income effect causes the consumer to buy less of the good as income falls.

4. **Cost-of-Living Adjustments.** The government's major index of inflation, the Consumer Price Index, overestimates inflation by ignoring the substitution effect. Though it is small on average, the substitution bias may be substantial for particular individuals or firms.

5. **Deriving Labor Supply Curves.** Using consumer theory, we can derive the daily demand curve for leisure, which is time spent on activities other than work. By subtracting the demand curve for leisure from 24 hours, we obtain the labor supply curve, which shows how the number of hours worked varies with the wage. Depending on whether leisure is an inferior good or a normal good, the labor supply curve may be upward sloping or backward bending. The shape of the supply curve for labor determines the effect of a tax cut. Empirical evidence based on this theory shows that tax cuts do not always increase tax revenue, as commonly predicted by various politicians.

Questions

Select questions are available on MyLab Economics*; * = answer appears at the back of this book ;* **A** *= algebra problem;* **C** *= calculus problem.*

1. Deriving Demand Curves

1.1 Draw diagrams similar to Figure 5.1 showing that the price-consumption curve can be horizontal or downward sloping.

*1.2 In Figure 5.1, how does Mimi's utility at E_1 on D^1 compare to that at E_2?

As we move down from the highest point on an individual's downward-sloping demand curve, must the individual's utility rise?

1.3 Derive and plot Olivia's demand curve for pie if she eats pie only à la mode and does not eat either pie or ice cream alone (pie and ice cream are perfect complements). (*Hint*: See Solved Problem 5.1.)

1.4 There is much controversy surrounding genetically modified (GM) foods in Europe, and the European Union member states have different policies concerning them. On one side, concerns have been expressed about the safety of GM food for human consumption and their effects on the environment and wildlife. On the other side are arguments that GM foods are just as nutritional as conventional foods and, being more resistant to disease and insects, GM crops increase yields and reduce prices. For example, it costs 10% to 15% less to produce GM maize per hectare compared to non-GM maize. Suppose that European consumers change from believing that GM foods, G, are imperfect substitutes for non-GM foods, N, to it being perfect substitutes so that their utility function becomes $U(G, N) = G + N$. What effect will this have on the demand for GM foods? Derive the demand curve for GM foods. **A**

1.5 How would your answer to Question 1.4 change if $U(G, N) = ln(G + N)$ so that consumers have diminishing marginal utility of GM foods? **A**

1.6 Illustrate the effect of cheaper phones in the Philippines that is described in the Application "Smoking Versus Eating and Phoning" using a figure similar to Figure 5.1.

1.7 Derive the demand curve for pizza, Z, if Lisa's utility function is $U = Z^{0.25}B^{0.75}$, where B is burritos. **C**

2. How Changes in Income Shift Demand Curves

2.1 Derive and plot Olivia's Engel curve for pie if she eats pie only à la mode and does not eat either pie or ice cream alone (pie and ice cream are perfect complements). (*Hint*: See Solved Problem 5.2.)

2.2 Sudi spends his income on two goods. His income elasticity of demand for the first good is $\xi_1 = 0.2$, while his income elasticity of demand for the second good is $\xi_1 = 2$. Illustrate in one diagram how a 10% increase in his income would affect the quantity he demands of the two goods that shows an income-consumption curve, and create another diagram for each of the two goods that shows an Engel curve. How do the slopes of the Engel curves compare?

2.3 Hugo views donuts and coffee as perfect complements: He always eats one donut with a cup of coffee and will not eat a donut without coffee or drink coffee without a donut. Derive and plot Hugo's Engel curve for donuts. How much does his weekly budget have to rise for Hugo to buy one more donut per week? (*Hint*: See Solved Problem 5.2.) **A**

*2.4 Don spends his money on food and on operas. Food is an inferior good for Don. Does he view an opera performance as an inferior or a normal good? Why? In a diagram, show a possible income-consumption curve for Don.

*2.5 Using calculus, show that not all goods can be inferior. (*Hint*: Start with the identity that $y = p_1q_1 + p_2q_2 + \ldots + p_nq_n$.) **C**

2.6 Generic private-label brands can cost 25% to 30% less than nationally recognized name brands. Data show that people buy more private-label products during times of economic weakness. In Europe, for example, it is estimated that the market share of private-label products has increased over the past few years to about 30% in Hungary, Poland, Slovakia, and the Czech Republic and over 50% in Switzerland and Spain. If lower-income consumers also tend to prefer private-label products, what can you conclude about the Engel curve for private-label goods? (*Hint*: See the Application "Fast-Food Engel Curve.")

3. Effects of a Price Change

3.1 Michelle spends all her money on food and clothing. When the price of clothing decreases, she buys more clothing.

a. Does the substitution effect cause her to buy more or less clothing? Explain. (If the direction of the effect is ambiguous, say so.)

b. Does the income effect cause her to buy more or less clothing? Explain. (If the direction of the effect is ambiguous, say so.)

3.2 Steve's utility function is $U = BC$, where B = veggie burgers per week and C = packs of cigarettes per week. Here, $MU_B = C$ and $MU_C = B$. What is his marginal rate of substitution if veggie burgers are on the vertical axis and cigarettes are on the horizontal axis? Steve's income is \$120, the price of a veggie burger is \$2, and that of a pack of cigarettes is \$1. How many burgers and how many packs of cigarettes does Steve consume to maximize his utility? When a new tax raises the price of a burger to \$3, what is his new optimal bundle? Illustrate your answers in a graph. In a related graph, show his demand curve for burgers with after-tax price on the vertical axis and show the points on the demand curve corresponding to the before- and after-tax equilibria. (*Hint*: See Appendix 4B.) **C**

3.3 A price increase has two effects on the quantity of a good demanded: consumers switch to relatively less expensive goods (holding utility and other prices constant) and less goods are demanded because purchasing power and the opportunity set has been reduced. What are the two effects on quantity demanded called? Does the quantity demanded of a good typically fall when its price rises? Does the quantity demanded of a good always have to fall when its price rises? Explain. **A**

3.4 A price increase has two effects on the quantity of a good demanded: consumers switch to relatively less expensive goods (holding utility and other prices constant) and less goods are demanded because purchasing power and the opportunity set has been reduced. What are the two effects on quantity demanded called? Does the quantity demanded of a good typically fall when its price rises? Does the quantity demanded of a good always have to fall when its price rises? Explain.

3.5 Using a figure similar to Figure 5.4 or that in Solved Problem 5.3, discuss the substitution, income, and total effects of a price change for Coke for Mahdu who views Coke and Pepsi as perfect substitutes. (*Hint*: See Solved Problem 5.1.)

3.6 India is the top papaya-producing country in the world, accounting for about 40% of the global supply, and most of its production is consumed domestically. Farmers in the state of Andhra Pradesh produce the fruit using both organic and non-organic farming techniques. The price per kilogram for organic papaya is higher than for non-organic papaya. While relatively more of the non-organic papaya grown in Andhra Pradesh is consumed in the local market, relatively more of its organic papaya is consumed in markets of other states within the country. Holding other things constant in the two markets, explain why this might be the case. (*Hint*: See Solved Problem 5.4.)

*3.7 Draw a figure to illustrate the answer given in Solved Problem 5.4. Use math and a figure to show how adding an ad valorem tax changes the analysis.

3.8 Remy views ice cream and fudge sauce as perfect complements. Is it possible that either of these goods or both of them are Giffen goods? (*Hint*: See Solved Problem 5.5.)

3.9 Redraw Figure 5.4 using an equivalent variation rather than a compensating variation approach.

4. Cost-of-Living Adjustments

4.1 *The Wall Street Journal*, a newspaper based in the United States, publishes an index of Starbucks coffee prices in various cities around the world. Would people find this index to be as useful as the Consumer Price Index in measuring how their true cost of living changes over time?

4.2 During his first year at school, Ximing buys eight new college textbooks at a cost of \$50 each. Used books cost \$30 each. When the bookstore announces a 20% price increase in new texts and a 10% increase in used texts for the next year, Ximing's father offers him \$80 extra. Is Ximing better off, the same, or worse off after the price change? Why?

4.3 Ann's only income is her annual college scholarship, which she spends exclusively on gallons of ice cream and books. Last year when ice cream cost \$10 and used books cost \$20, Ann spent her \$250 scholarship on five gallons of ice cream and ten books. This year, the price of ice cream rose to \$15 and the price of books increased to \$25. So that Ann can afford the same bundle of ice cream and books that she bought last year, her college raised her scholarship to \$325. Ann has the usual-shaped indifference curves. Will Ann change the amount of ice cream and books that she buys this year? If so, explain how and why. Will Ann be better off, as well off, or worse off this year than last year? Why?

*4.4 Alix views coffee and cream as perfect complements. In the first period, Alix picks an optimal bundle of coffee and cream, e_1. In the second period, inflation occurs, the prices of coffee and cream change by different amounts, and Alix receives a cost-of-living adjustment (COLA) based on the Consumer Price Index (CPI) for these two goods. After the price changes and she receives the COLA, her new optimal bundle is e_2. Show the two equilibria in a figure. Is she better off, worse off, or equally well off at e_2 compared to e_1? Explain why. By how much will a CPI for these two goods differ from the true cost-of-living index?

4.5 Change Figure 5.5 so that L^2 is steeper than L^1 (but still goes through e_1); that is, food increases by more than clothing in the second year. Show the conclusion that Klaas is still better off after receiving a CPI adjustment. Explain the logic behind the following statement: "The analysis holds as long as the relative prices differ over the two years. Whether both prices, one price, or neither price is higher in the first year than in the second year is irrelevant to the analysis."

4.6 The price of a serving of McDonald's French fries in 1950 was 10¢. Using the Internet, or visiting a McDonald's, determine the price of fries today. The federal government's urban CPI index is available at **http://www.bls.gov/cpi/#data and at https://research.stlouisfed.org/fred2/series/CPIAUCNS**. Based on these data, has the real price of fries increased?

5. Deriving Labor Supply Curves

5.1 On January 1, 2017, the minimum wage in Germany was raised from €8.50 per hour to €8.84 per hour. Frieda's job pays the minimum wage, and she presently works 8 hours per day, while her utility function is $U(Y, N) = Y^{\alpha}N^{1-\alpha}$, where Y is her income and N is leisure. Will she choose to work more or fewer hours per day because of the wage increase? Use calculus to find the answer. (*Hint*: See Appendix 5B.) **C**

5.2 If an individual's labor supply curve slopes upward at low wages and bends backward at high wages, is leisure a Giffen good? If so, at high or low wage rates? (*Hint*: See Solved Problems 5.5 and 5.6.)

5.3 Bessie, who can currently work as many hours as she wants at a wage of w, chooses to work ten hours a day. Her boss decides to limit the number of hours that she can work to eight hours per day. Show how her budget constraint and choice of hours change. Is she unambiguously worse off as a result of this change? Why? (*Hint*: See Solved Problem 5.6.)

5.4 Suppose that Roy could choose how many hours to work at a wage of w and chose to work seven hours a day. The employer now offers him time-and-a-half wages ($1.5w$) for every hour he works beyond a minimum of eight hours per day. Show how his budget constraint changes. Will he choose to work more than seven hours a day? (*Hint*: See Solved Problem 5.6.)

5.5 Jerome moonlights: He holds down two jobs. The higher-paying job pays w, but he can work at most eight hours. The other job pays w^*, but he can work as many hours as he wants. Show how Jerome determines how many hours to work. (*Hint*: See Solved Problem 5.6.)

5.6 Suppose that the job in Question 5.5 that pays w^* and has no restriction on hours is the higher-paying job. How do Jerome's budget constraint and behavior change? (*Hint*: See Solved Problem 5.6.)

5.7 Suppose that Bill's wage varies with the hours he works: $w(H) = aH$, $a > 0$. Show how the number of hours he chooses to work depends on his tastes.

5.8 Joe won \$365,000 a year for life in the state lottery. Use a labor-leisure choice analysis to answer the following questions:

a. Show how Joe's lottery winnings affect the position of his budget line.

b. After winning the lottery, Joe continues to work the same number of hours each day (see the Application "Working After Winning the Lottery"). What is the income effect from Joe's lottery gains?

c. Suppose Joe's employer increases Joe's hourly wage. Use the income effect you derived in part b as well as the substitution effect to analyze whether Joe chooses to work more hours per week. (*Hint*: See Solved Problem 5.6.)

5.9 A flat tax is a tax on income at the same rate for everyone regardless of how much income they earn. Estonia reduced its rate of flat tax from 21% in 2014 to 20% in 2015. What effect can such a tax rate reduction be expected to have on hours worked and tax revenues? Explain your answer in terms of income, substitution, and total effects.

*5.10 Portugal has a progressive personal income tax system. In 2016, tax rates on taxable income were 14.5% on the first €7,035, 21% on the next €13,065, 37% on the next €20,100, 45% on the next €39,800, and 48% on taxable income over €80,000. If Afonso had €30,000 of taxable income in 2016, what was his average tax rate? If he had one more euro of taxable income, how much of that euro would he have kept after tax? What does this tell you about the marginal and average tax rates under a progressive income tax system? **A**

5.11 In its 2016 International Tax Competitiveness Index report, the U.S.-based Tax Foundation ranked Estonia as having the most competitive tax system in the OECD, based in part on its 20% flat tax on individual income. A flat tax is simply a tax on income at the same rate for everyone regardless of how much income they earn. However, Estonia also allows a basic exemption, which equals €2,160 for 2017. Show that the income tax in Estonia is actually a progressive tax because of the basic exemption. For this purpose, assume that Grete, a resident of Estonia, earns €14,000 of income in 2017.

5.12 Under a welfare plan, poor people are given a lump-sum payment of $\$L$. If they accept this welfare payment, they must pay a high tax, $v = \frac{1}{2}$, on anything they earn. If they do not accept the welfare payment, they do not have to pay a tax on their earnings. Show that whether an individual accepts welfare depends on the individual's tastes.

5.13 A lump-sum tax is a fixed amount of tax per person. If a lump-sum tax, T, raises the same amount of revenue for the government as a tax on earnings at the rate, t, then $twH = T$, where w is the wage rate and H is the number of hours of work per day. Assume that Mika works 10 hours per day in the absence of any tax and that her income effect equals her substitution effect when a tax is imposed. Will a lump-sum tax or a tax on earnings affect the number of hours she works?

*5.14 If the wage elasticity of labor supply is negative, what can we say about the slope of the labor supply curve and the relative sizes of the income and substitution effects? Is leisure a normal or inferior good in this case? Will a fall in the tax rate on earnings increase or decrease tax revenues?

5.15 See the Application "Paying Employees to Relocate." A firm transfers its employee Kiki to another country and agrees to pay her enough to buy the same bundle of goods as she currently buys. Kiki does not care where she lives. Draw a figure to illustrate that Kiki is better off in the new country given this compensation. Explain your answer.

6. Challenge

6.1 In Norway, a parent may receive a cash benefit for private childcare for a child between the ages of 1 and 2 years if the child does not attend a government-subsidized childcare facility. Effective August 1, 2017, the maximum amount of this cash-for-care benefit is 625 kr per month for a full-time placement, and it is not reduced if the child does not attend every day. If Berit spends 600 kr for full-time private childcare per month for her 14-month-old child with the cash subsidy of 625 kr, would she be indifferent to receiving the subsidy in the form of a voucher of equal value? Use a diagram to illustrate and explain your answer.

*6.2 Given the figure in the Challenge Solution, could the government set a smaller lump-sum subsidy than shown that would make poor parents as well off as with the price subsidy yet cost the government less? Given the tastes shown in the figure, what would be the effect on the number of hours of childcare service that these parents buy?

*6.3 How do parents who do not receive childcare subsidies feel about the two programs discussed and illustrated in the Challenge Solution? (*Hint*: Use a supply-and-demand analysis from Chapters 2 and 3.)

6 Firms and Production

Hard work never killed anyone, but why take a chance?

Challenge

More Productive Workers During Downturns

Why has a measure of labor productivity—the output produced per worker—risen for many firms during recent recessions, such as the Great Recession (Lazear et al., 2016)? During the Great Recession (fourth quarter of 2007 through the third quarter of 2009), labor productivity rose by 3.2% in nonfarm businesses. In contrast, in the two years before the Great Recession, labor productivity rose by only 2.2%.

Firms produce less output during recessions as demand for their products falls. Consequently, firms typically lay off workers during recessions. Whether output per worker rises depends on whether output falls by more or less than employment. The labor productivity pattern over the business cycle differs across industries.

In the Solution to this Challenge, we're going to examine whether the productivity of a beer bottling plant rises or falls. Given that we know the firm's production process, can we predict whether output produced per worker will rise or fall with each additional layoff?

This chapter looks at the types of decisions that the owners of firms have to make. First, a decision must be made as to how a firm is owned and managed. For example, the American Licorice Co. is a corporation—it is not owned by an individual or by partners—and is run by professional managers. Second, the firm must decide how to produce. American Licorice Co. now uses relatively more machines and robots and fewer workers than in the past. Third, if a firm wants to expand output, it must decide how to do that in both the short run and the long run. In the short run, American Licorice Co. can expand output by extending the workweek to six or seven days and using extra materials. To expand output more in the long run, American Licorice Co. would have to install more equipment (such as extra robotic arms), hire more workers, and eventually build a new plant, all of which take time.

In this chapter, we examine the nature of a firm and how a firm chooses its inputs so as to produce efficiently. In Chapter 7, we examine how the firm chooses the least costly among all possible efficient production processes. In Chapter 8, we combine this information about costs with information about revenues to determine how a firm picks the output level that maximizes profit.

The main lesson of this chapter and the next is that firms are not black boxes that mysteriously transform inputs (such as labor, capital, and material) into outputs. Economic theory explains how firms make decisions about production processes, types of inputs to use, and the volume of output to produce.

In this chapter, we examine six main topics

1. **The Ownership and Management of Firms.** How businesses are organized affects who makes decisions and the firm's objective, such as whether it tries to maximize profit.
2. **Production.** A firm converts inputs into outputs using one of possibly many available technologies.
3. **Short-Run Production.** In the short run, only some inputs can be varied, so the firm changes its output by adjusting its variable inputs.
4. **Long-Run Production.** The firm has more flexibility in how it produces and how it changes its output level in the long run when all factors can be varied.
5. **Returns to Scale.** How the ratio of output to input varies with the size of the firm is an important factor in determining the size of a firm.
6. **Productivity and Technical Change.** The amount of output that can be produced with a given amount of inputs varies across firms and over time.

6.1 The Ownership and Management of Firms

firm
an organization that converts inputs such as labor, materials, energy, and capital into outputs, the goods and services that it sells

A **firm** is an organization that converts *inputs* such as labor, materials, and capital into *outputs*, the goods and services that it sells. U.S. Steel combines iron ore, machinery, and labor to create steel. A local restaurant buys raw food, cooks it, and serves it. A landscape designer hires gardeners and rents machines, buys trees and shrubs, transports them to a customer's home, and supervises the project.

Private, Public, and Nonprofit Firms

Atheism is a non-prophet organization. —George Carlin

Firms operate in the private, public, or nonprofit sectors. The *private sector*, sometimes referred to as the *for-profit private sector*, consists of firms owned by individuals or other nongovernmental entities whose owners try to earn a profit. Throughout this book, we concentrate on these firms. In almost every country, this sector contributes the most to the gross domestic product (GDP, a measure of a country's total output).

The *public sector* consists of firms and organizations that are owned by governments or government agencies. For example, the National Railroad Passenger Corporation (Amtrak) is owned primarily by the U.S. government. The armed forces and the court system are also part of the public sector, as are most schools, colleges, and universities.

The *nonprofit* or *not-for-profit sector* consists of organizations that are neither government-owned nor primarily intended to earn a profit. Organizations in this sector typically pursue social or public interest objectives. Well-known examples include Greenpeace, Alcoholics Anonymous, and the Salvation Army, along with many other charitable, educational, health, and religious organizations. According to the U.S. Bureau of Economic Analysis in 2016, the private sector created 75% of the U.S. gross domestic product, the government sector was responsible for 12%, and nonprofits and households produced the remaining 13%.

Sometimes all three sectors play an important role in the same industry. For example, in the United States, the United Kingdom, Canada, and in many other countries, for-profit, nonprofit, and government-owned hospitals coexist. A single enterprise may be partially owned by a government and partially owned by private interests. For example, during the 2007–2009 Great Recession, the U.S. government took a partial ownership position in many firms in the financial and automobile industries.

Application

Chinese State-Owned Enterprises

Before 1978, virtually all Chinese industrial firms were state-owned enterprises (SOEs). Since then, China has been transitioning to a market-based economy, gradually increasing the role of private-sector firms. China has dramatically reduced the number of SOEs, keeping the largest ones. By 1999, SOEs comprised only about 36% of Chinese industrial firms but still controlled nearly 68% of industrial assets (capital). Since 2000, the Chinese government has allowed many small SOEs to be privatized or to go bankrupt, while it continues to subsidize many large SOEs. By 2014, SOEs accounted for only about 5% of industrial firms, but they still held 40% of industrial assets.

The Ownership of For-Profit Firms

The legal structure of a firm determines who is liable for its debts. Firms in the private sector have three primary legal forms of organization: a sole proprietorship, a general partnership, or a corporation.

Sole proprietorships are firms owned by a single individual.

Partnerships are businesses jointly owned and controlled by two or more people operating under a partnership agreement.

Corporations are owned by *shareholders*, who own the firm's *shares* (also called *stock*). Each share (or unit of stock) is a unit of ownership in the firm. Therefore, shareholders own the firm in proportion to the number of shares they hold. The shareholders elect a board of directors to represent them. In turn, the board of directors usually hires managers who manage the firm's operations. Some corporations are very small and have a single shareholder. Others are very large and have thousands of shareholders. The legal name of a corporation often includes the term Incorporated (Inc.) or Limited (Ltd) to indicate its corporate status.

limited liability
the personal assets of the owners of the corporation cannot be taken to pay a corporation's debts even if it goes into bankruptcy

A fundamental characteristic of corporations is that the owners are not personally liable for the firm's debts; they have **limited liability**: The personal assets of corporate owners cannot be taken to pay a corporation's debts even if it goes into bankruptcy. Because corporations have limited liability, the most that shareholders can lose is the amount they paid for their stock, which typically becomes worthless if the corporation declares bankruptcy.[1]

The purpose of limiting liability was to allow firms to raise funds and grow beyond what was possible when owners risked personal assets on any firm in which they invested. According to the latest available statistics from the Internal Revenue Service as of 2016, U.S. corporations are responsible for 81% of revenue (business receipts) and 61% of net business income even though they are only 18% of all nonfarm firms. Nonfarm sole proprietorships are 72% of firms but make only 4% of the sales revenue and earn 10% of net income. Partnerships are 10% of firms, account for 15% of revenue, and make 29% of net income.

As these statistics illustrate, larger firms tend to be corporations and smaller firms are often sole proprietorships. This pattern reflects a natural evolution in the life cycle of the firm, as an entrepreneur may start a small business as a sole proprietorship and then incorporate as the firm's operations expand. Indeed, successful corporations typically expand, and a relatively small number of corporations account for most of the revenue and income in the U.S. economy.

[1]The United States since 1996, the United Kingdom since 2000, and some other countries now allow any sole proprietorship, partnership, or corporation to register as a *limited liability company* (LLC). Thus, all firms—not just corporations—can now obtain limited liability.

The Management of Firms

In a small firm, the owner usually manages the firm's operations. In larger firms, typically corporations and larger partnerships, a manager or a management team usually runs the company. In such firms, owners, managers, and lower-level supervisors are all decision makers.

Various decision makers may have conflicting objectives. What is in the best interest of the owners may not be in the best interest of managers or other employees. For example, a manager may want a fancy office, a company car, a corporate jet, and other perks, but an owner would likely oppose those drains on profit.

The owner replaces the manager if the manager pursues personal objectives rather than the firm's objectives. In a corporation, the board of directors is responsible for ensuring that the manager stays on track. If the manager and the board of directors are ineffective, the shareholders can fire both or change certain policies through votes at the corporation's annual shareholders' meeting. Until Chapter 20, we'll ignore the potential conflict between managers and owners and assume that the owner *is* the manager of the firm and makes all the decisions.

What Owners Want

profit (π)
the difference between revenue, R, and cost, C: $\pi = R - C$

Economists usually assume that a firm's owners try to maximize profit. Presumably, most people invest in a firm to make money—lots of money, they hope. They want the firm to earn a positive profit rather than make a loss (a negative profit). A firm's **profit** (π), is the difference between its revenue, R, which is what it earns from selling a good, and its cost, C, which is what it pays for labor, materials, and other inputs:

$$\pi = R - C.$$

Typically, revenue is p, the price, times q, the firm's quantity: $R = pq$.

In reality, some owners have other objectives, such as running as large a firm as possible, owning a fancy building, or keeping risks low. However, Chapter 8 shows that a firm in a highly competitive market is likely to be driven out of business if it doesn't maximize its profit.

efficient production
the current level of output cannot be produced with fewer inputs, given existing knowledge about technology and the organization of production

To maximize profits, a firm must produce as efficiently as possible, as we will consider in this chapter. A firm engages in **efficient production** (achieves *technological efficiency*) if it cannot produce its current level of output with fewer inputs, given existing knowledge about technology and the organization of production. Equivalently, the firm produces efficiently if, given the quantity of inputs used, no more output could be produced using existing knowledge.

If the firm does not produce efficiently, it cannot be profit maximizing—so efficient production is a *necessary condition* for profit maximization. Even if a firm produces a given level of output efficiently, it is not maximizing profit if that output level is too high or too low or if it is using excessively expensive inputs. Thus, efficient production alone is not a *sufficient condition* to ensure that a firm's profit is maximized.

A firm may use engineers and other experts to determine the most efficient ways to produce using a known method or technology. However, this knowledge does not indicate which of the many technologies, each of which uses different combinations of inputs, allows for production at the lowest cost or with the highest possible profit. How to produce at the lowest cost is an economic decision typically made by the firm's manager (see Chapter 7).

6.2 Production

A firm uses a *technology* or *production process* to transform *inputs* or *factors of production* into *outputs*. Firms use many types of inputs. Most of these inputs can be grouped into three broad categories:

- **Capital services (*K*):** use of long-lived inputs such as land, buildings (such as factories and stores), and equipment (such as machines and trucks)
- **Labor services (*L*):** hours of work provided by managers, skilled workers (such as architects, economists, engineers, and plumbers), and less-skilled workers (such as custodians, construction laborers, and assembly-line workers)
- **Materials (*M*):** natural resources and raw goods (such as oil, water, and wheat) and processed products (such as aluminum, plastic, paper, and steel) that are typically consumed in producing, or incorporated in making, the final product

For brevity, we typically refer to *capital services* as *capital* and *labor services* as *labor*.

The output can be a *service* such as an automobile tune-up by a mechanic, or a *physical product* such as a computer chip or a potato chip.

Production Functions

Firms can transform inputs into outputs in many different ways. Candy manufacturing companies differ in the skills of their workforce and the amount of equipment they use. While all employ a chef, a manager, and relatively unskilled workers, some candy firms also use skilled technicians and modern equipment. In small candy companies, the relatively unskilled workers shape the candy, decorate it, package it, and box it by hand. In slightly larger firms, these same-level workers use conveyor belts and other industrial equipment. In modern large-scale plants, the relatively unskilled laborers work with robots and other state-of-the-art machines maintained by skilled technicians. Before deciding which production process to use, a firm must consider its options.

production function the relationship between the quantities of inputs used and the maximum quantity of output that can be produced, given current knowledge about technology and organization

The various ways inputs can be transformed into output are summarized in the **production function**: the relationship between the quantities of inputs used and the *maximum* quantity of output that can be produced, given current knowledge about technology and organization. The production function for a firm that uses only labor and capital is

$$q = f(L, K), \tag{6.1}$$

where q units of output (wrapped candy bars) are produced using L units of labor services (days of work by relatively unskilled assembly-line workers) and K units of capital (the number of conveyor belts).

The production function shows only the *maximum* amount of output that can be produced from given levels of labor and capital, because the production function includes efficient production processes only. A profit-maximizing firm is not interested in production processes that are inefficient and wasteful: Why would a firm want to use two workers to do a job that one worker can perform as efficiently?

Varying Inputs over Time

A firm can more easily adjust its inputs in the long run than in the short run. Typically, a firm can vary the amount of materials and of relatively unskilled labor it uses comparatively quickly. However, it needs more time to find and hire skilled workers, order new equipment, or build a new manufacturing plant.

short run
a period so brief that at least one factor of production cannot be varied practically

fixed input
a factor of production that the firm cannot practically vary in the short run

variable input
a factor of production that the firm can easily vary during the relevant period

long run
a lengthy enough period that all factors of production can be varied

The more time a firm has to adjust its inputs, the more factors of production it can alter. The **short run** is a period so brief that at least one factor of production cannot be varied practically. A **fixed input** is a factor of production that the firm cannot practically vary in the short run. In contrast, a **variable input** is a factor of production that the firm can easily vary during the relevant period. The **long run** is a lengthy enough period that all factors of production can be varied. In the long run, the firm has no fixed inputs.

Suppose that one day a painting company has more work than its crew can handle. Even if it wanted to, the firm does not have time to buy or rent an extra truck and buy another compressor to run a power sprayer; these inputs are fixed in the short run. To complete the day's work, the firm uses its only truck to drop off a temporary worker, equipped with only a brush and a can of paint, at the last job. However, in the long run, the firm can adjust all its inputs. If the firm wants to paint more houses every day, it can hire more full-time workers, purchase or lease a second truck, get another compressor to run a power sprayer, and buy a computer program to track its projects.

The time it takes for all inputs to be variable depends on the factors a firm uses. For a janitorial firm whose only major input is workers, the long run is a brief period. In contrast, an automobile manufacturer may need many years to build a new manufacturing plant or design and construct a new type of machine. A pistachio farmer needs about a decade before newly planted trees yield a substantial crop of nuts.

For many firms over a short period, say a month, materials and often labor are variable inputs. However, labor is not always a variable input. Finding additional highly skilled workers may take substantial time. Similarly, capital may be a variable or a fixed input. A firm can rent small capital assets (trucks and personal computers) quickly, but it may take years to obtain larger capital assets (buildings and large specialized pieces of equipment).

To illustrate the greater flexibility that a firm has in the long run than in the short run, we examine the production function in Equation 6.1, in which output is a function of only labor and capital. We look first at the short-run and then at the long-run production processes.

6.3 Short-Run Production

The short run is a period in which at least one input is fixed. We consider a production process with only two inputs in which capital is a fixed input and labor is a variable input. The firm can increase output only by increasing the amount of labor it uses. In the short run, the firm's production function is

$$q = f(L, \overline{K}), \tag{6.2}$$

where q is output, L is the amount of labor, and $\overline{K}$ is the fixed number of units of capital.

To illustrate the short-run production process, we consider a firm that assembles computers for a manufacturing firm that supplies it with the necessary parts, such as computer chips and disk drives. The assembly firm cannot increase its capital—eight workbenches fully equipped with tools, electronic probes, and other equipment for testing computers—in the short run, but it can hire extra workers or pay current workers extra to work overtime so as to increase production.

Total Product

The exact relationship between *output* or *total product* and *labor* can be illustrated by using a particular function, Equation 6.2, a table, or a figure. Table 6.1 shows the relationship between output and labor when capital is fixed for a firm. The first column

Table 6.1 Total Product, Marginal Product, and Average Product of Labor with Fixed Capital

Capital, $\bar{K}$	Labor, L	Output, Total Product, q	Marginal Product of Labor, $MP_L = \Delta q/\Delta L$	Average Product of Labor, $AP_L = q/L$
8	0	0		
8	1	5	5	5
8	2	18	13	9
8	3	36	18	12
8	4	56	20	14
8	5	75	19	15
8	6	90	15	15
8	7	98	8	14
8	8	104	6	13
8	9	108	4	12
8	10	110	2	11
8	11	110	0	10
8	12	108	−2	9
8	13	104	−4	8

lists the fixed amount of capital: eight fully equipped workbenches. As the number of workers (the amount of labor, second column) increases, total output (the number of computers assembled in a day, third column) first increases and then decreases.

With zero workers, no computers are assembled. One worker with access to the firm's equipment assembles five computers in a day. As the number of workers increases, so does output: 1 worker assembles 5 computers in a day, 2 workers assemble 18, 3 workers assemble 36, and so forth. However, the maximum number of computers that can be assembled with the capital on hand is limited to 110 per day. That maximum can be produced with 10 or 11 workers. Adding extra workers beyond 11 lowers production as workers get in each other's way. The dashed line in the table indicates that a firm would not use 11 or more workers, because doing so would be inefficient. That is, the production function—which only includes efficient production—involves fewer than 11 workers.

Marginal Product of Labor

marginal product of labor (MP_L) the change in total output, Δq, resulting from using an extra unit of labor, ΔL, holding other factors constant: $MP_L = \Delta q/\Delta L$

We can show how extra workers affect the total product by using two additional concepts: the marginal product of labor and the average product of labor. Before deciding whether to hire one more worker, a manager wants to determine how much this extra worker, $\Delta L = 1$, will increase output, Δq. That is, the manager wants to know the **marginal product of labor (MP_L)**: the change in total output, Δq, resulting from using an extra unit of labor, ΔL, holding other factors (capital) constant. If output changes by Δq when the number of workers increases by ΔL, the change in output per worker is[2]

[2]With the long-run production function $q = f(L, K)$, the calculus definition of the marginal product of labor is $MP_L = \partial q/\partial L = \partial f(L, K)/\partial L$, where capital is held constant at K. In the short run, with capital fixed at $\bar{K}$, we can write the production function as solely a function of labor $q = f(L, \bar{K}) = \hat{f}(L)$. Thus, in the short run, $MP_L = \mathrm{d}q/\mathrm{d}L = \mathrm{d}\hat{f}/\mathrm{d}L$.

$$MP_L = \frac{\Delta q}{\Delta L}.$$

As Table 6.1 shows, if the number of workers increases from 1 to 2, $\Delta L = 1$, output rises by $\Delta q = 18 - 5 = 13$, so the marginal product of labor is 13.

Solved Problem 6.1

MyLab Economics Solved Problem

For a linear production function $q = f(L, K) = 2L + K$, what is the short-run production function given that capital is fixed at $\overline{K} = 100$? What is the marginal product of labor?

Answer

1. *Obtain the short-run production functions by setting* $\overline{K} = 100$. The short-run production function is $q = 2L + 100$.

2. *Determine the marginal product of labor by showing how q changes as L is increased by* ΔL *units.* The output at $L + \Delta L$ is $q = 2(L + \Delta L) + 100$. Taking the difference between this output and the output with L units of labor, $q = 2L + 100$, we find that $\Delta q = (2[L + \Delta L] + 100) - (2L + 100) = 2\Delta L$. Thus, the marginal product of labor is $MP_L = \Delta q/\Delta L = 2$.[3]

Average Product of Labor

average product of labor (AP_L) the ratio of output, q, to the number of workers, L, used to produce that output: $AP_L = q/L$

Before hiring extra workers, a manager may also want to know whether output will rise in proportion to the extra labor. To answer this question, the firm determines how extra workers affect the **average product of labor (AP_L):** the ratio of output, q, to the number of workers, L, used to produce that output,

$$AP_L = \frac{q}{L}.$$

Table 6.1 shows that 9 workers can assemble 108 computers a day, so the average product of labor for 9 workers is 12 (= 108/9) computers a day. Ten workers can assemble 110 computers in a day, so the average product of labor for 10 workers is 11 (= 110/10) computers. Thus, increasing the labor force from 9 to 10 workers lowers the average product per worker.

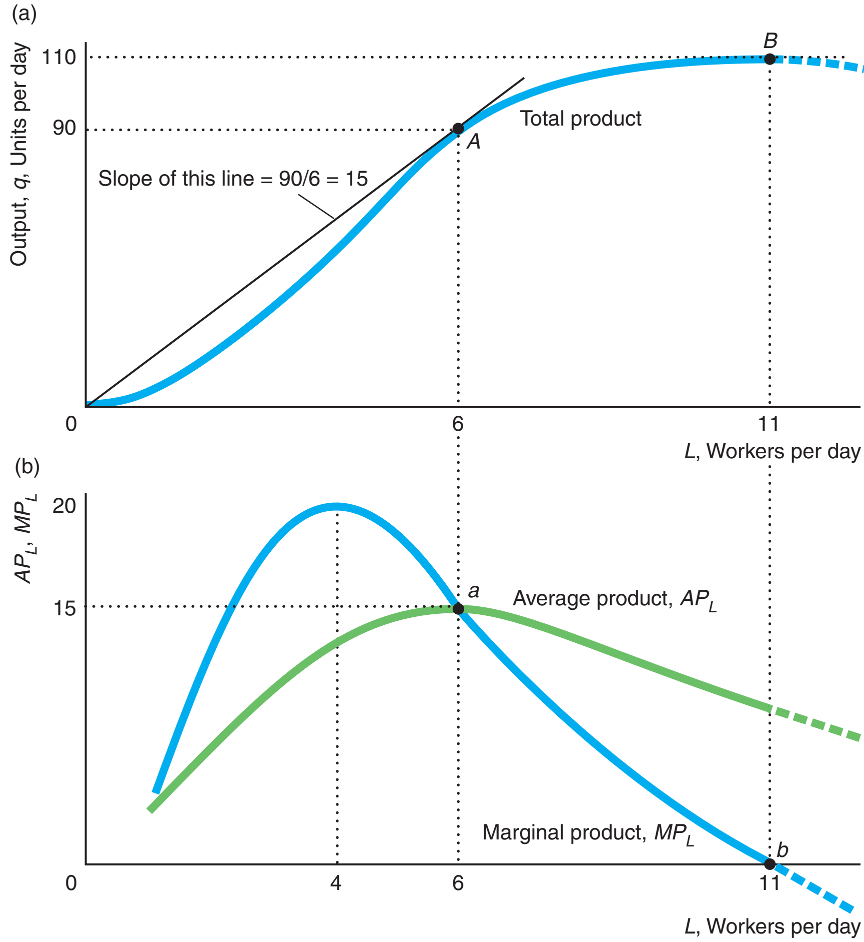
Graphing the Product Curves

Figure 6.1 and Table 6.1 show how output, the average product of labor, and the marginal product of labor vary with the number of workers. (The figure shows smooth curves because the firm can hire a "fraction of a worker" by employing a worker for a fraction of a day.) The curve in panel a of Figure 6.1 shows how a change in labor affects the *total product*, which is the amount of output (or *total product*) that can be produced by a given amount of labor. Output rises with labor until it reaches its maximum of 110 computers at 11 workers, point B; with additional workers beyond that maximum, the number of computers assembled falls.

[3]Using calculus, we can derive the same result by differentiating the short-run production function with respect to labor: $MP_L = d(2L + 100)/dL = 2$.

Figure 6.1 Production Relationships with Variable Labor

(a) The total product curve shows how many computers, q, can be assembled with eight fully equipped workbenches and a varying number of workers, L, who work an eight-hour day (see columns 2 and 3 in Table 6.1). Where additional workers reduce the number of computers assembled, the total product curve is a dashed line, which indicates that such production is inefficient production and not part of the production function. The slope of the line from the origin to point A is the average product of labor for six workers. (b) The marginal product of labor ($MP_L = \Delta q/\Delta L$, column 4 of Table 6.1) equals the average product of labor ($AP_L = q/L$, column 5 of Table 6.1) at the peak of the average product curve.



Panel b of the figure shows how the average product of labor and marginal product of labor vary with the number of workers. We can line up the figures in panels a and b vertically because the units along the horizontal axes of both figures, the number of workers per day, are the same. The vertical axes differ, however. The vertical axes are total product in panel a and the average or marginal product of labor—measure of output per unit of labor—in panel b.

Effect of Extra Labor In most production processes, the average product of labor first rises and then falls as labor increases. One reason the AP_L curve initially rises in Figure 6.1 is that it helps to have more than two hands when assembling a computer. One worker holds a part in place while another one bolts it down. As a result, output increases more than in proportion to labor, so the average product of labor rises. Doubling the number of workers from one to two more than doubles the output from 5 to 18 and causes the average product of labor to rise from 5 to 9, as Table 6.1 shows.

Similarly, output may initially rise more than in proportion to labor because of greater specialization of activities. With greater specialization, workers are assigned to tasks at which they are particularly adept, and time is saved by not having workers move from task to task.

As the number of workers rises further, however, output may not increase by as much per worker because, for example, workers may need to wait to use a particular piece of equipment or they may get in each other's way. In Figure 6.1, as the number of workers exceeds 6, total output increases less than in proportion to labor, so the average product falls.

If more than 11 workers are used, the total product curve falls with each extra worker as the crowding of workers gets worse. Because that much labor is not efficient, that section of the curve is drawn with a dashed line to indicate that it is not part of the production function, which includes only efficient combinations of labor and capital. Similarly, the dashed portions of the average and marginal product curves are irrelevant because no firm would hire additional workers if doing so meant that output would fall.

Relationship of the Product Curves The three curves in Figure 6.1 are geometrically related. First we use panel b to illustrate the relationship between the average and marginal product of labor curves. Then we use panels a and b to show the relationship between the total product curve and the other two curves.

An extra hour of work increases the average product of labor if the marginal product of labor exceeds the average product. Similarly, if an extra hour of work generates less extra output than the average, then the average product falls. Therefore, the average product rises with extra labor if the marginal product curve is above the average product curve, and the average product falls if the marginal product is below the average product curve. Consequently, the average product curve reaches its peak, point *a* in panel b of Figure 6.1, where the marginal product and average product are equal: where the curves cross. (See Appendix 6A for a mathematical proof.)

We can determine the average product of labor using the total product curve. The average product of labor for L workers equals the slope of a straight line from the origin to a point on the total product curve for L workers in panel a. The slope of this line equals output divided by the number of workers, which is the definition of the average product of labor. For example, the slope of the straight line drawn from the origin to point A ($L = 6$, $q = 90$) is 15, which equals the *rise* of $q = 90$ divided by the *run* of $L = 6$. As panel b shows, the average product of labor for 6 workers at point *a* is 15.

The marginal product of labor also has a geometric interpretation in terms of the total product curve. The slope of the total product curve at a given point, $\Delta q/\Delta L$, equals the marginal product of labor. That is, the marginal product of labor equals the slope of a straight line that is tangent to the total output curve at a given point. For example, at point B in panel a where the firm uses 11 workers, the line tangent to the total product curve is flat, so the marginal product of labor is zero: A little extra labor has no effect on output. The total product curve is upward sloping when the firm uses fewer than 11 workers, so the marginal product of labor is positive. If the firm is foolish enough to hire more than 11 workers, the total product curve slopes downward (dashed line), so the MP_L is negative: Extra workers, lower output. Again, this portion of the MP_L curve is not part of the production function.

With 6 workers, the average product of labor equals the marginal product of labor, at the peak of the AP_L curve. The reason is that the line from the origin to point A in panel a is tangent to the total product curve, so the slope of that line, 15, is the marginal product of labor and the average product of labor at point *a* in panel b.

Law of Diminishing Marginal Returns

Next to "supply equals demand," the most commonly used phrase of economic jargon is probably the "law of diminishing marginal returns." This law determines the shapes of the total product and marginal product of labor curves as the firm uses more and more labor.

The *law of diminishing marginal returns* (or *diminishing marginal product*) holds that if a firm keeps increasing an input, holding all other inputs and technology constant, the corresponding increases in output will become smaller eventually. That is, if only one input is increased, *the marginal product of that input will diminish eventually*.

In Table 6.1, if the firm goes from 1 to 2 workers, the marginal product of labor is 13. If 1 or 2 more workers are used, the marginal product rises: The marginal product for 3 workers is 18, and the marginal product for 4 workers is 20. However, if the firm increases the number of workers beyond 4, the marginal product falls: The marginal product of 5 workers is 19, and that for 6 workers is 15. Beyond 4 workers, each extra worker adds less and less extra output, so the total product curve rises by smaller increments. At 11 workers, the marginal product is zero. In short, the law of diminishing marginal returns says that if a firm keeps adding one more unit of an input, the additional output grows smaller and smaller. This diminishing return to extra labor may be due to too many workers sharing too few machines or to crowding, as workers get in each other's way. Thus, as the amount of labor used grows large enough, the marginal product curve approaches zero and the corresponding total product curve becomes nearly flat.

Unfortunately, many people, when attempting to cite this empirical regularity, overstate it. Instead of talking about the law of *diminishing marginal returns*, they talk about *diminishing returns*—leaving out the word *marginal*. The two phrases have different meanings. If as labor increases the marginal returns fall but remain positive, the total return rises. In panel b of Figure 6.1, marginal returns start to diminish when the labor input exceeds 4 but total returns rise, as panel a shows, until the labor input exceeds 11, where the marginal returns become negative. With *diminishing returns*, extra labor causes *output* to fall. The production process has diminishing (total) returns for more than 11 workers—a dashed line in panel a.

Thus, saying that the production process has diminishing returns is much stronger than saying that it has diminishing *marginal* returns. We often observe firms producing where the marginal returns to labor are diminishing, but no well-run firm operates where total returns diminish. Such a firm could produce more output by using fewer inputs.

Due to the influence of Thomas Malthus, many people misstate the law of diminishing marginal returns.

Common Confusion: Marginal product must fall as an input increases.

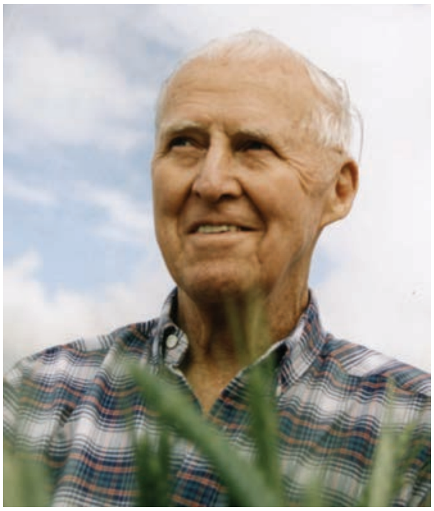
That claim is true only if as we add more of an input, we hold technology and other inputs constant. If we increase labor while simultaneously increasing other factors or adopting superior technologies, the marginal product of labor may rise indefinitely.

Application

Malthus and the Green Revolution

In 1798, Thomas Malthus—a clergyman and professor of modern history and political economy—predicted that population (if unchecked) would grow more rapidly than food production because the quantity of land was fixed. The problem, he believed, was that the fixed amount of land would lead to a diminishing marginal product of labor, so output would rise less than in proportion to the increase in farm workers. Malthus grimly concluded that mass starvation would result. Indeed, Brander and Taylor (1998) argue that such a disaster may have occurred on Easter Island around 500 years ago.

Today, the earth supports a population almost seven times as great as it was when Malthus made his predictions. Why haven't most of us starved to death? The simple



explanation is that fewer workers using less land can produce much more food today than was possible when Malthus was alive. The output of a U.S. farm worker today is more than double that of an average worker just 50 years ago. We do not see diminishing marginal returns to labor because the production function has changed due to substantial technological progress in agriculture and because farmers make greater use of other inputs such as fertilizers, capital, and superior seeds.

Two hundred years ago, most of the population had to work in agriculture to feed themselves. Today, less than 1% of the U.S. population works in agriculture. Over the last century, food production grew substantially faster than the population in most developed countries. For example, since World War II, the U.S. population doubled but U.S. food production tripled.

In 1850 in the United States, it took more than 80 hours of labor to produce 100 bushels of corn. Introducing mechanical power cut the required labor in half. Labor hours were again cut in half by the introduction of hybrid seed and chemical fertilizers, and then in half again by the advent of herbicides and pesticides. Biotechnology, with the introduction of herbicide-tolerant and insect-resistant crops, has reduced the labor required to produce 100 bushels of corn to about two hours—2.5% of the hours of work it took in 1850. Over the past 60 years, the output per work has more than doubled, and the corn yield per acre has increased by 6.2 times.

Of course, the risk of starvation is more severe in developing countries. Nearly all (98%) of the world's hungry people live in developing countries. Luckily, one man decided to defeat the threat of Malthusian disaster personally. Do you know anyone who saved a life? A hundred lives? Do you know the name of the man who probably saved the most lives in history? According to some estimates, during the second half of the twentieth century, Norman Borlaug and his fellow scientists prevented a *billion deaths* with their *Green Revolution*, which included development of drought- and insect-resistant crop varieties, improved irrigation, better use of fertilizer and pesticides, and improved equipment.

However, as Dr. Borlaug noted in his 1970 Nobel Prize speech, superior science is not the complete answer to preventing starvation. A sound economic system and a stable political environment are also needed.

Economic and political failures, such as the breakdown of economic production and distribution systems due to wars, were responsible for one in nine people around the world suffering undernourishment in 2014–2016, one in five in Africa overall, and two in five in Middle Africa. If society cannot solve these economic and political problems, Malthus' prediction may prove to be right for the wrong reason.

6.4 Long-Run Production

We started our analysis of production functions by looking at a short-run production function in which one input, capital, was fixed, and the other, labor, was variable. In the long run, however, both of these inputs are variable. With both factors variable, a firm can usually produce a given level of output by using a great deal of labor and very little capital, a great deal of capital and very little labor, or moderate amounts of both. That is, the firm can substitute one input for another while continuing to produce the same level of output, in much the same way that a consumer can maintain a given level of utility by substituting one good for another.

Typically, a firm can produce in a number of different ways, some of which require more labor than others. For example, a lumberyard can produce 200 planks an hour with 10 workers using hand saws, with 4 workers using handheld power saws, or with 2 workers using bench power saws.

We illustrate a firm's ability to substitute between inputs in Table 6.2, which shows the amount of output per day the firm produces with various combinations of labor and capital per day. The labor inputs, L, are along the top of the table, and the capital inputs, K, are in the first column. The table shows four combinations of labor and capital that the firm can use to produce 24 units of output: The firm may employ (a) 1 worker and 6 units of capital, (b) 2 workers and 3 units of capital, (c) 3 workers and 2 units of capital, or (d) 6 workers and 1 unit of capital.

Isoquants

isoquant
a curve that shows the efficient combinations of labor and capital that can produce the same (*iso*) level of output (*quantity*)

These four combinations of labor and capital are labeled a, b, c, and d on the "$q = 24$" curve in Figure 6.2. We call such a curve an **isoquant**, which is a curve that shows the efficient combinations of labor and capital that can produce the same (*iso*) level of output (*quant*ity). If the production function is $q = f(L, K)$, then the equation for an isoquant where output is held constant at $\overline{q}$ is

$$\overline{q} = f(L, K).$$

An isoquant shows the flexibility that a firm has in producing a given level of output. Figure 6.2 shows three isoquants corresponding to three levels of output. These isoquants are smooth curves because the firm can use fractional units of each input.

We can use these isoquants to illustrate what happens in the short run when capital is fixed and only labor varies. As Table 6.2 shows, if capital is constant at 2 units, 1 worker produces 14 units of output (point e in Figure 6.2), 3 workers produce 24 units (point c), and 6 workers produce 35 units (point f). Thus, if the firm holds one factor constant and varies another factor, it moves from one isoquant to another. In contrast, if the firm increases one input while lowering the other appropriately, the firm stays on a single isoquant.

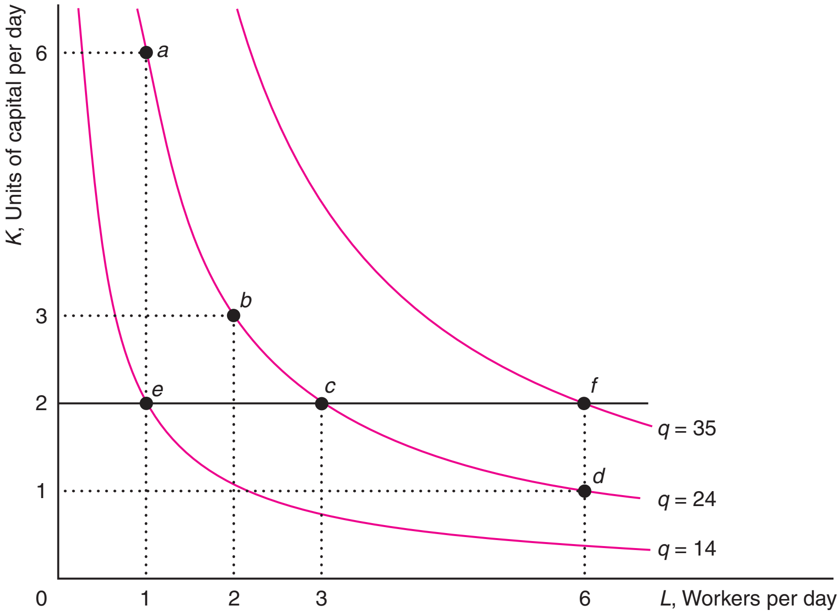
Properties of Isoquants Isoquants have most of the same properties as indifference curves. The biggest difference between indifference curves and isoquants is that an isoquant holds quantity constant, whereas an indifference curve holds utility

Table 6.2 Output Produced with Two Variable Inputs

	Labor, L					
Capital, K	**1**	**2**	**3**	**4**	**5**	**6**
1	10	14	17	20	22	24
2	14	20	24	28	32	35
3	17	24	30	35	39	42
4	20	28	35	40	45	49
5	22	32	39	45	50	55
6	24	35	42	49	55	60

Figure 6.2 Family of Isoquants MyLab Economics Video

These isoquants show the combinations of labor and capital that produce various levels of output. Isoquants farther from the origin correspond to higher levels of output. Points a, b, c, and d are various combinations of labor and capital the firm can use to produce $q = 24$ units of output. If the firm holds capital constant at 2 and increases labor from 1 (point e) to 3 (c) to 6 (f), it shifts from the $q = 14$ isoquant to the $q = 24$ isoquant and then to the $q = 35$ isoquant.



constant. We now discuss three major properties of isoquants. Most of these properties result from firms producing efficiently.

First, *the farther an isoquant is from the origin, the greater the level of output*. That is, the more inputs a firm uses, the more output it gets if it produces efficiently. At point e in Figure 6.2, the firm is producing 14 units of output with 1 worker and 2 units of capital. If the firm holds capital constant and adds 2 more workers, it produces at point c. Point c must be on an isoquant with a higher level of output—here, 24 units—if the firm is producing efficiently and not wasting the extra labor.

Second, *isoquants do not cross*. Such intersections are inconsistent with the requirement that the firm always produces efficiently. For example, if the $q = 15$ and $q = 20$ isoquants crossed, the firm could produce at either output level with the same combination of labor and capital. The firm must be producing inefficiently if it produces $q = 15$ when it could produce $q = 20$. So that labor-capital combination should not lie on the $q = 15$ isoquant, which should include only efficient combinations of inputs. Thus, efficiency requires that isoquants do not cross.

Third, *isoquants slope downward*. If an isoquant sloped upward, the firm could produce the same level of output with relatively few inputs or relatively many inputs. Producing with relatively many inputs would be inefficient. Consequently, because isoquants show only efficient production, an upward-sloping isoquant is impossible. Virtually the same argument can be used to show that isoquants must be thin.

Shape of Isoquants The curvature of an isoquant shows how readily a firm can substitute one input for another. The two extreme cases are production processes in which inputs are perfect substitutes or in which they cannot be substituted for each other.

If the inputs are perfect substitutes, each isoquant is a straight line. Suppose either potatoes from Maine, x, or potatoes from Idaho, y, both of which are measured in pounds per day, can be used to produce potato salad, q, measured in pounds. The production function is

$$q = x + y.$$

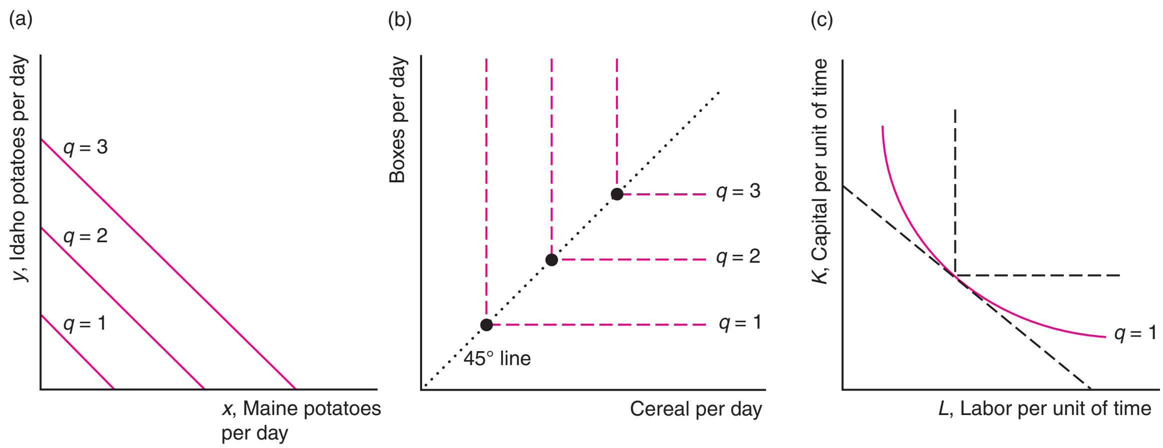
One pound of potato salad can be produced using 1 pound of Idaho potatoes and no Maine potatoes, 1 pound of Maine potatoes and no Idahoes, or $\frac{1}{2}$ pound of each type of potato. Panel a of Figure 6.3 shows the $q = 1$, 2, and 3 isoquants. These isoquants are straight lines with a slope of -1 because we need to use an extra pound of Maine potatoes for every pound fewer of Idaho potatoes used.[4]

Sometimes it is impossible to substitute one input for the other: Inputs must be used in fixed proportions. Such a production function is called a *fixed-proportions production function*. For example, the inputs to produce a 12-ounce box of cereal, q, are cereal (in 12-ounce units per day) and cardboard boxes (boxes per day). If the firm has one unit of cereal and one box, it can produce one box of cereal. If it has one unit of cereal and two boxes, it can still make only one box of cereal. Thus, in panel b, the only efficient points of production are the large dots along the 45° line.[5] Dashed lines show that the isoquants would be right angles if isoquants could include inefficient production processes.

Other production processes allow imperfect substitution between inputs. The isoquants are convex (so the middle of the isoquant is closer to the origin than it would be if the isoquant were a straight line). They do not have the same slope at every point, unlike the straight-line isoquants. Most isoquants are smooth, slope downward, curve away from the origin, and lie between the extreme cases of straight lines (perfect substitutes) and right angles (nonsubstitutes), as panel c illustrates.

Figure 6.3 Substitutability of Inputs

(a) If the inputs are perfect substitutes, each isoquant is a straight line. (b) If the inputs cannot be substituted at all, the isoquants are right angles (the dashed lines show that the isoquants would be right angles if we included inefficient production). (c) Typical isoquants lie between the extreme cases of straight lines and right angles. Along a curved isoquant, the ability to substitute one input for another varies.



[4]The isoquant for $\bar{q} = 1$ pound of potato salad is $1 = x + y$, or $y = 1 - x$. This equation shows that the isoquant is a straight line with a slope of -1.

[5]This fixed-proportions production function is $q = \min(g, b)$, where g is the number of 12-ounce measures of cereal, b is the number of boxes used in a day, and the min function means "the minimum number of g or b." For example, if g is 4 and b is 3, q is 3.

Application

A Semiconductor Integrated Circuit Isoquant

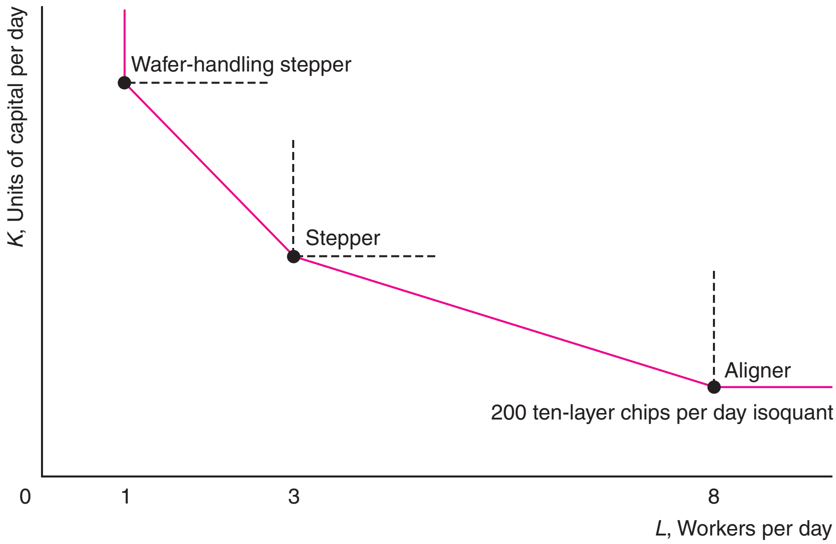
We can show why isoquants curve away from the origin by deriving an isoquant for semiconductor integrated circuits (ICs, or "chips")—the "brains" of computers and other electronic devices. Semiconductor manufacturers buy silicon wafers and then use labor and capital to produce the chips.

A chip consists of multiple layers of silicon wafers. A key step in the production process is to line up these layers. Three alternative alignment technologies are available, using different combinations of labor and capital. In the least capital-intensive technology, employees use machines called *aligners*, which require workers to look through microscopes and line up the layers by hand. To produce 200 ten-layer chips per day takes 8 workers using 8 aligners.

A second, more capital-intensive technology uses machines called *steppers*. The stepper aligns the layers automatically. This technology requires less labor: To produce 200 ten-layer chips per day requires 3 workers and 6 steppers.

A third, even more capital-intensive technology combines steppers with wafer-handling equipment, which further reduces the amount of labor needed. One worker using 4 steppers with wafer-handling capabilities can manufacture 200 ten-layer chips per day.

In the diagram, the vertical axis measures the amount of capital used. An aligner represents less capital than a basic stepper, which in turn is less capital than a stepper with wafer-handling capabilities. All three technologies use labor and capital in fixed proportions. The diagram shows the three right-angle isoquants corresponding to each of these three technologies.

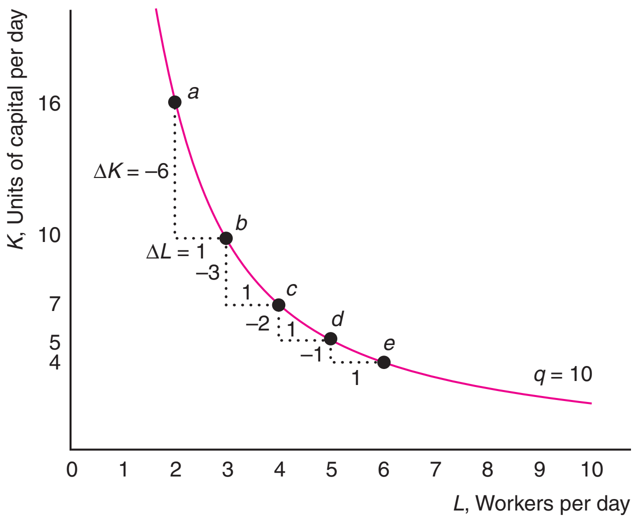


Some plants employ a combination of these technologies, so that some workers use one type of machine while others use different types. By doing so, the plant can produce using intermediate combinations of labor and capital, as the solid-line, kinked isoquant illustrates. The firm does *not* use a combination of the aligner and the wafer-handling stepper technologies because those combinations are less efficient than using the basic stepper: The line connecting the aligner and wafer-handling stepper technologies is farther from the origin than the lines between those technologies and the basic stepper technology.

New processes are constantly being invented. As they are introduced, the isoquant will have more and more kinks (one for each new process) and will begin to resemble the smooth, convex isoquants we've been drawing.

Figure 6.4 How the Marginal Rate of Technical Substitution Varies Along an Isoquant MyLab Economics Video

Moving from point a to b, a service firm can produce the same amount of output, $q = 10$, using six fewer units of capital, $\Delta K = -6$, if it uses one more worker, $\Delta L = 1$ (Devine et al., 2012). Thus, its *MRTS*, $\Delta K/\Delta L$, is -6. Moving from point b to c, its *MRTS* is -3. If it adds yet another worker, moving from c to d, its *MRTS* is -2. Finally, if it moves from d to e, its *MRTS* is -1. Thus, because it curves away from the origin, this isoquant exhibits a diminishing marginal rate of technical substitution. That is, each extra worker allows the firm to reduce capital by a smaller amount as the ratio of capital to labor falls.



Substituting Inputs

The slope of an isoquant shows a firm's ability to replace one input with another while holding output constant. Figure 6.4 illustrates this substitution using an estimated isoquant for a service firm, which uses labor, L, and capital, K, to produce q units of service.[6] The isoquant shows various combinations of L and K that the firm can use to produce 10 units of output.

The firm can produce 10 units of output using the combination of inputs at a or b. At point a, the firm uses 2 workers and 16 units of capital. The firm could produce the same amount of output using six fewer units of capital, $\Delta K = -6$, if it used one more worker, $\Delta L = 1$, point b. If we drew a straight line from a to b, its slope would be $\Delta K/\Delta L = -6$. Thus, this slope tells us how many fewer units of capital (6) the firm can use if it hires one more worker.[7]

The slope of an isoquant is called the *marginal rate of technical substitution* (*MRTS*):

$$MRTS = \frac{\text{change in capital}}{\text{change in labor}} = \frac{\Delta K}{\Delta L}.$$

marginal rate of technical substitution (*MRTS*)
the extra units of one input needed to replace one unit of another input that enables a firm to keep the amount of output it produces constant

The **marginal rate of technical substitution** (***MRTS***) the extra units of one input needed to replace one unit of another input that enables a firm to keep the amount of output it produces constant. Because isoquants slope downward, the *MRTS* is negative. That is, the firm can produce a given level of output by substituting more capital for less labor (or vice versa).

[6]This isoquant for $q = 10$ is based on the estimated "personal and other service" (such as gardening, hairdressers, laundry) production function $q = 2.35L^{0.5}K^{0.4}$ (Devine et al., 2012), where a unit of labor, L, is a worker-day. Because capital, K, includes various types of machines, and output, q, reflects different types of service, their units cannot be described by any common terms.

[7]The slope of the isoquant at a point equals the slope of a straight line that is tangent to the isoquant at that point. Thus, the straight line between two nearby points on an isoquant has nearly the same slope as that of the isoquant.

Substitutability of Inputs Varies Along an Isoquant The marginal rate of technical substitution varies along a curved isoquant, as in Figure 6.4 for the service firm. If the firm is initially at point *a* and it hires one more worker, the firm gives up 6 units of capital and yet remains on the same isoquant at point *b*, so the *MRTS* is −6. If the firm hires another worker, the firm can reduce its capital by 3 units and yet stay on the same isoquant, moving from point *b* to *c*, so the *MRTS* is −3. If the firm moves from point *c* to *d*, the *MRTS* is −2; and if it moves from point *d* to *e*, the *MRTS* is −1. This decline in the *MRTS* (in absolute value) along an isoquant as the firm increases labor illustrates *diminishing marginal rates of technical substitution*. The more labor and less capital the firm has, the harder it is to replace remaining capital with labor and the flatter the isoquant becomes.

In the special case in which isoquants are straight lines, isoquants do not exhibit diminishing marginal rates of technical substitution because neither input becomes more valuable in the production process: The inputs remain perfect substitutes. Solved Problem 6.2 illustrates this result.

Solved Problem 6.2

Does the marginal rate of technical substitution vary along the isoquant for the firm that produced potato salad using Idaho and Maine potatoes? What is the *MRTS* at each point along the isoquant?

Answer

1. *Determine the shape of the isoquant.* As panel a of Figure 6.3 illustrates, the potato salad isoquants are straight lines because the two types of potatoes are perfect substitutes.

2. *On the basis of the shape, conclude whether the MRTS is constant along the isoquant.* Because the isoquant is a straight line, the slope is the same at every point, so the *MRTS* is constant.

3. *Determine the MRTS at each point.* Earlier, we showed that the slope of this isoquant was −1, so the MRTS is −1 at each point along the isoquant. That is, because the two inputs are perfect substitutes, 1 pound of Idaho potatoes can be replaced by 1 pound of Maine potatoes.

Substitutability of Inputs and Marginal Products The marginal rate of technical substitution is equal to the negative of the ratio of marginal products, as we now show.[8] The marginal rate of technical substitution tells us how much a firm can increase one input and lower the other while still staying on the same isoquant. Knowing the marginal products of labor and capital, we can determine how much one input must increase to offset a reduction in the other.

The marginal product of labor, $MP_L = \Delta q/\Delta L$, is the increase in output from an extra unit of labor, holding other inputs fixed. For example, if the MP_L is 2 and the firm hires one extra worker, its output rises by 2 units. Thus, if the firm hires ΔL more workers, its output increases by $MP_L \times \Delta L$.

Similarly, the marginal product of capital, $MP_K = \Delta q/\Delta K$, is the increase in output from an extra unit of capital, holding other inputs fixed. Thus, a decrease in capital, holding labor fixed, causes output to fall by $MP_K \times \Delta K$.

[8]See Appendix 6B for a derivation using calculus.

If the firm increases labor and decreases capital so as to keep output constant, $\Delta q = 0$, the fall in output caused by reducing capital must exactly equal the increase in output resulting from increasing labor:

$$(MP_L \times \Delta L) + (MP_K \times \Delta K) = 0.$$

Rearranging these terms, we find that

$$-\frac{MP_L}{MP_K} = \frac{\Delta K}{\Delta L} = MRTS. \tag{6.3}$$

Thus, the negative of the ratio of the marginal products equals the *MRTS*.

We can use Equation 6.3 to explain why marginal rates of technical substitution diminish as we move to the right along the isoquant in Figure 6.4. The fewer pieces of capital per worker, the more valuable is each piece of capital, so the marginal product of capital rises. Similarly, the more workers per piece of capital, the lower is the marginal product of labor. As we replace capital with labor—moving downward and to the right along the isoquant—the marginal product of capital increases and the marginal product of labor falls. Thus, the $MRTS = -MP_L/MP_K$ falls in absolute value as we move down and to the right along the isoquant.

Cobb-Douglas Production Function We can illustrate how to determine the *MRTS* for a particular production function, the Cobb-Douglas production function:[9]

$$q = AL^a K^b, \tag{6.4}$$

where A, a, and b are all positive constants.

In empirical studies, economists have found that the production processes in a very large number of industries can be accurately summarized by the Cobb-Douglas production function. For the estimated production function of a service firm in Figure 6.4 (Devine et al., 2012), the Cobb-Douglas production function is $q = 2.35L^{0.5}K^{0.4}$, so $A = 2.35$, $a = 0.5$, and $b = 0.4$.

The constants a and b determine the relationships between the marginal and average products of labor and capital (see Appendix 6C). The marginal product of labor is a times the average product of labor, $AP_L = q/L$:

$$MP_L = aq/L = aAP_L. \tag{6.5}$$

Similarly, the marginal product of capital is

$$MP_K = bq/K = bAP_K. \tag{6.6}$$

For a Cobb-Douglas production function, the marginal rate of technical substitution along an isoquant that holds output fixed at $\bar{q}$ is

$$MRTS = -\frac{MP_L}{MP_K} = -\frac{a\bar{q}/L}{b\bar{q}/K} = -\frac{a}{b}\frac{K}{L}. \tag{6.7}$$

For example, for the service firm, the $MRTS = -(0.5/0.4)K/L = -1.25K/L$. As we move down and to the right along the isoquant, the capital/labor ratio, K/L, falls, so the *MRTS* approaches zero.

[9]This production function is named after its discoverers, Charles W. Cobb, a mathematician, and Paul H. Douglas, an economist and U.S. Senator.

6.5 Returns to Scale

So far, we have examined the effects of increasing one input while holding the other input constant (shifting from one isoquant to another) or decreasing the other input by an offsetting amount (the movement along a single isoquant). We now turn to the question of *how much output changes if a firm increases all its inputs proportionately*. The answer helps a firm determine its *scale* or size in the long run.

In the long run, a firm can increase its output by building a second plant and staffing it with the same number of workers as in the first one. Whether the firm chooses to do so depends in part on whether its output increases less than in proportion to, in proportion to, or more than in proportion to its inputs.

Constant, Increasing, and Decreasing Returns to Scale

constant returns to scale
property of a production function whereby when all inputs are increased by a certain percentage, output rises by that same percentage

increasing returns to scale
property of a production function whereby output rises more than in proportion to an equal percentage increase in all inputs

decreasing returns to scale
property of a production function whereby output rises less than in proportion to an equal percentage increase in all inputs

If, when all inputs are increased by a certain percentage, output rises by that same percentage, the production function is said to exhibit **constant returns to scale**. A firm's production process, $q = f(L, K)$, has constant returns to scale if, when the firm doubles its inputs—by, for example, building an identical second plant and using the same amount of labor and equipment as in the first plant—it doubles its output:

$$f(2L, 2K) = 2f(L, K) = 2q.$$

We can check whether the potato salad production function has constant returns to scale. If a firm uses x_1 pounds of Idaho potatoes and y_1 pounds of Maine potatoes, it produces $q_1 = x_1 + y_1$ pounds of potato salad. If it doubles both inputs, using $x_2 = 2x_1$ Idaho and $y_2 = 2y_1$ Maine potatoes, it doubles its output:

$$q_2 = x_2 + y_2 = 2x_1 + 2y_1 = 2q_1.$$

Thus, the potato salad production function exhibits constant returns to scale.

If output rises more than in proportion to an equal percentage increase in all inputs, the production function is said to exhibit **increasing returns to scale**. A technology exhibits increasing returns to scale if doubling inputs more than doubles the output:

$$f(2L, 2K) > 2f(L, K) = 2q.$$

This'll save a lot of time!

Why might a production function have increasing returns to scale? One reason is that, although it could duplicate a small factory and double its output, the firm might be able to more than double its output by building a single large plant, thereby allowing for greater specialization of labor or capital. In the two smaller plants, workers have to perform many unrelated tasks such as operating, maintaining, and fixing the machines they use. In the large plant, some workers may specialize in maintaining and fixing machines, thereby increasing efficiency. Similarly, a firm may use specialized equipment in a large plant but not in a small one.

If output rises less than in proportion to an equal percentage increase in all inputs, the production function exhibits **decreasing returns to scale**. A technology exhibits decreasing returns to scale if doubling inputs causes output to rise less than in proportion:

$$f(2L, 2K) < 2f(L, K) = 2q.$$

One reason for decreasing returns to scale is that the difficulty of organizing, coordinating, and integrating activities increases with firm size. An owner may be able to manage one plant well but may have trouble running two plants. In some sense, the owner's difficulties in running a larger firm may reflect our failure to take into account some factor, such as management, in our production function. When the firm increases the various inputs, it does not increase the management input in proportion. If so, the "decreasing returns to scale" is really due to a fixed input. Another reason is that large teams of workers may not function as well as small teams, in which each individual takes greater personal responsibility.

Solved Problem 6.3

MyLab Economics Solved Problem

Under what conditions does a Cobb-Douglas production function (Equation 6.4, $q = AL^aK^b$) exhibit decreasing, constant, or increasing returns to scale?

Answer

1. *Show how output changes if both inputs are doubled.* If the firm initially uses L and K amounts of inputs, it produces $q_1 = AL^aK^b$. After the firm doubles its inputs of both labor and capital, its output is

$$q_2 = A(2L)^a(2K)^b = 2^{a+b}AL^aK^b = 2^{a+b}q_1. \quad (6.8)$$

That is, q_2 is 2^{a+b} times q_1. If we define $g = a + b$, then Equation 6.8 tells us that

$$q_2 = 2^g q_1. \quad (6.9)$$

Thus, if the inputs double, output increases by 2^g.

2. *Give a rule for determining the returns to scale.* If $g = 1$, we know from Equation 6.9 that $q_2 = 2^1q_1 = 2q_1$. That is, output doubles when the inputs double, so the Cobb-Douglas production function has constant returns to scale. If $g < 1$, then $q_2 = 2^g q_1 < 2q_1$ because $2^g < 2$. That is, when inputs double, output increases less than in proportion, so this Cobb-Douglas production function exhibits decreasing returns to scale. For example, if $g = 0.8$, then $q_2 = 2^{0.8}q_1 \approx 1.74q_1$, so doubling inputs increases output by only 1.74 times. Finally, the Cobb-Douglas production function has increasing returns to scale if $g > 1$, so that $q_2 > 2q_1$. For example, if $g = 1.2$, then $q_2 = 2^{1.2}q_1 \approx 2.3q_1$. Thus, the rule for determining returns to scale for a Cobb-Douglas production function is that the returns to scale are decreasing if $g < 1$, constant if $g = 1$, and increasing if $g > 1$.

Comment: One interpretation of g is that it is the elasticity of output with respect to all inputs. When all inputs increase by 1%, output increases by g%. For example, if $g = 1$, a 1% increase in all inputs increases output by 1%, so the elasticity equals one.

Application

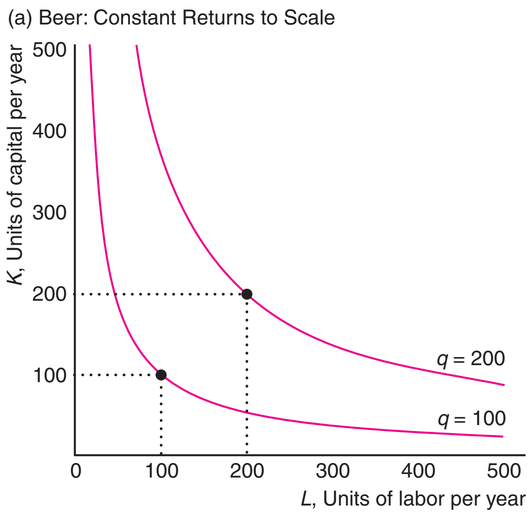
Returns to Scale in Various Industries

Increasing, constant, and decreasing returns to scale are commonly observed. The table shows estimates of Cobb-Douglas production functions and returns to scale in various industries.

The graphs use isoquants to illustrate the returns to scale for three firms: a Japanese beer firm, a U.S. tobacco firm, and a Bangladesh fabricated metal firm. We measure the units of labor, capital, and output so that, for all three firms, 100 units of labor and 100 units of capital produce 100 units of output on the $q = 100$ isoquant in the three panels. These graphs illustrate that the spacing of the isoquant reflects the returns to scale. The closer together the $q = 100$ and $q = 200$ isoquants, the greater the returns to scale.

	Labor, a	Capital, b	Scale, $g = a + b$
Decreasing Returns to Scale			
U.S. tobacco products[1]	0.18	0.33	0.51
Bangladesh glass[2]	0.27	0.45	0.72
Danish food and beverages[3]	0.69	0.18	0.87
Chinese high technology[4]	0.28	0.66	0.94
Constant Returns to Scale			
Japanese synthetic rubber[5]	0.50	0.50	1.00
Japanese beer[5]	0.60	0.40	1.00
New Zealand wholesale trade[6]	0.60	0.42	1.02
Danish publishing and printing[3]	0.89	0.14	1.03
Increasing Returns to Scale			
New Zealand mining[6]	0.69	0.45	1.14
Bangladesh leather products[2]	0.86	0.27	1.13
Bangladesh fabricated metal[2]	0.98	0.28	1.26

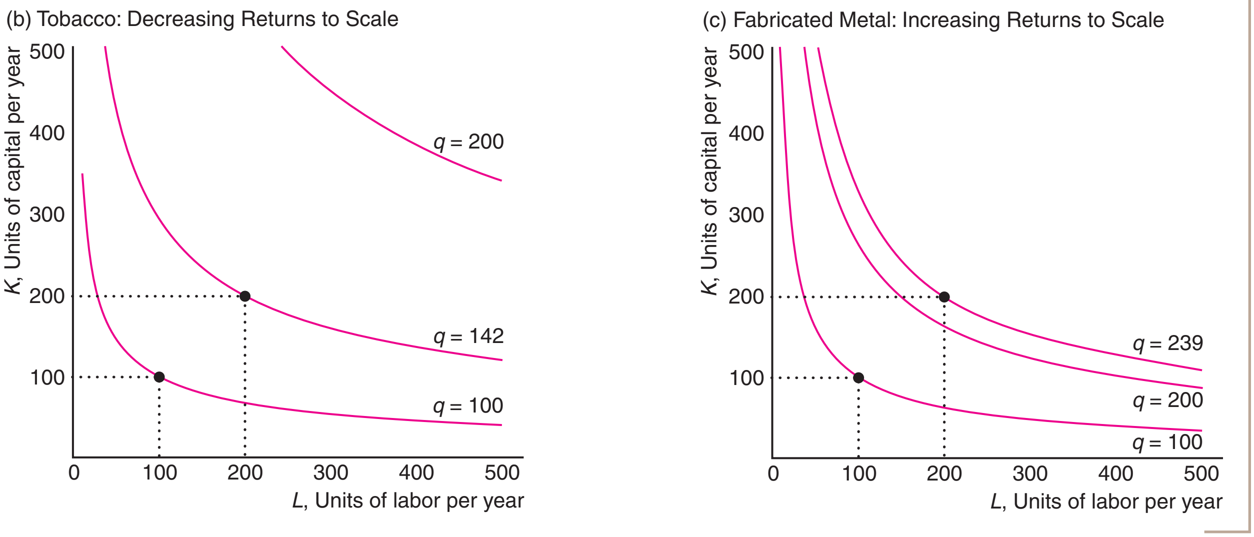
[1]Hsieh (1995); [2]Hossain et al. (2012); [3]Fox and Smeets (2011); [4]Zhang et al. (2012); [5]Flath (2011); [6]Devine et al. (2012).



In panel a, the beer firm has constant returns to scale because $g = 1$: A 1% increase in the inputs causes output to rise by 1%. If both its labor and capital are doubled from 100 to 200 units, output doubles to 200 ($= 100 \times 2^1$, multiplying the original output by the rate of increase using Equation 6.9).

In panel b, the tobacco firm has decreasing returns to scale, $g < 1$, because $g = 0.51$. The same doubling of inputs causes output to rise to only 142 ($\approx 100 \times 2^{0.51}$) for the tobacco firm: output rises less than in proportion to inputs.

In panel c, the fabricated metal firm exhibits increasing returns to scale, $g > 1$, because $g = 1.26$. If it doubles its inputs, its output more than doubles, to 239 ($\approx 100 \times 2^{1.26}$), so the production function has increasing returns to scale.



Varying Returns to Scale

In the case of a Cobb-Douglas production function, the returns to scale are the same at all levels of output. However, in other industries, a production function's returns to scale may vary as the output level changes. A firm might, for example, have increasing returns to scale at low levels of output, constant returns to scale for some range of output, and decreasing returns to scale at higher levels of output.

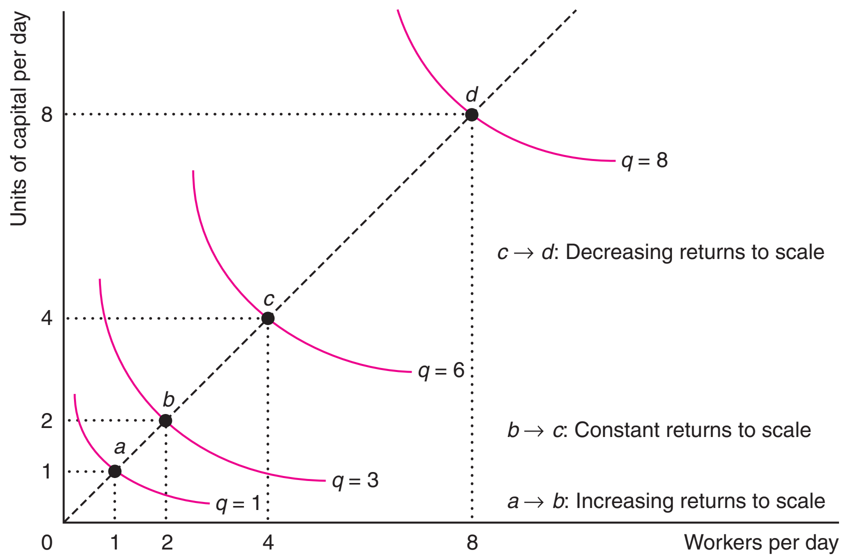
Sato and Söderbom (2012) found that the returns to scale fell with firm size for Swedish firms. The returns to scale were 1.156 for micro firms (fewer than 10 employees), 1.081 for small firms (10 to 49 employees), 1.010 for medium-sized firms (50 to 249 employees), and 0.934 for large firms.

Many production functions have increasing returns to scale for small amounts of output, constant returns for moderate amounts of output, and decreasing returns for large amounts of output. When a firm is small, increasing labor and capital allows for gains from cooperation between workers and greater specialization of workers and equipment—*returns to specialization*—so the production function exhibits increasing returns to scale. As the firm grows, returns to scale are eventually exhausted. After the firm grows to the point where it has no more returns to specialization, the production process exhibits constant returns to scale. If the firm continues to grow, the owner starts having difficulty managing everyone, so the firm suffers from decreasing returns to scale.

We show such a pattern in Figure 6.5. Again, the spacing of the isoquants reflects the returns to scale. Initially, the firm has one worker and one piece of equipment, point *a*, and produces 1 unit of output on the $q = 1$ isoquant. If the firm doubles its inputs, it produces at *b*, where $L = 2$ and $K = 2$, which lies on the dashed line through the origin and point *a*. Output more than doubles to $q = 3$, so the production function exhibits increasing returns to scale in this range. Another doubling of inputs to *c* causes output to double to 6 units, so the production function has constant returns to scale in this range. Another doubling of inputs to *d* causes output to increase by only a third, to $q = 8$, so the production function has decreasing returns to scale in this range.

Figure 6.5 Varying Scale Economies MyLab Economics Video

This production function exhibits varying returns to scale. Initially, the firm uses one worker and one unit of capital, point *a*. It repeatedly doubles these inputs to points *b*, *c*, and *d*, which lie along the dashed line. The first time the inputs are doubled, from *a* to *b*, output more than doubles from $q = 1$ to $q = 3$, so the production function has increasing returns to scale. The next doubling, from *b* to *c*, causes a proportionate increase in output, constant returns to scale. At the last doubling, from *c* to *d*, the production function exhibits decreasing returns to scale.



6.6 Productivity and Technical Change

Because firms may use different technologies and different methods of organizing production, the amount of output that one firm produces from a given amount of inputs may differ from that produced by another firm. Further, after a technical or managerial innovation, a firm can produce more today from a given amount of inputs than it could in the past.

Relative Productivity

This chapter has assumed that firms produce efficiently. A firm must produce efficiently to maximize its profit. However, even if each firm in a market produces as efficiently as possible, firms may not be equally *productive*—one firm may be able to produce more than another from a given amount of inputs.

One firm may be more productive than another if its management knows a better way to organize production or if it has access to a new invention. Union-mandated work rules, racial or gender discrimination, government regulations, or institutional restrictions that affect only certain firms may lower the relative productivity of those firms.

Differences in productivity across markets may be due to differences in the degree of competition. In competitive markets, where many firms can enter and exit easily, less productive firms lose money and are driven out of business, so the firms that continue to produce are equally productive (see Chapter 8). In a less competitive market with few firms and no possibility of entry by new ones, a less productive firm may be able to survive, so firms with varying levels of productivity are observed.[10]

Innovations

In its production process, a firm tries to use the best available technological and managerial knowledge. An advance in knowledge that allows more output to be produced with the same level of inputs is called **technical progress**. The invention of new products, a *technical innovation*, causes technical progress. The use of robotic arms increases the number of automobiles produced with a given amount of labor and raw materials. Better *management* or *organization of the production process* similarly allows the firm to produce more output from given levels of inputs.

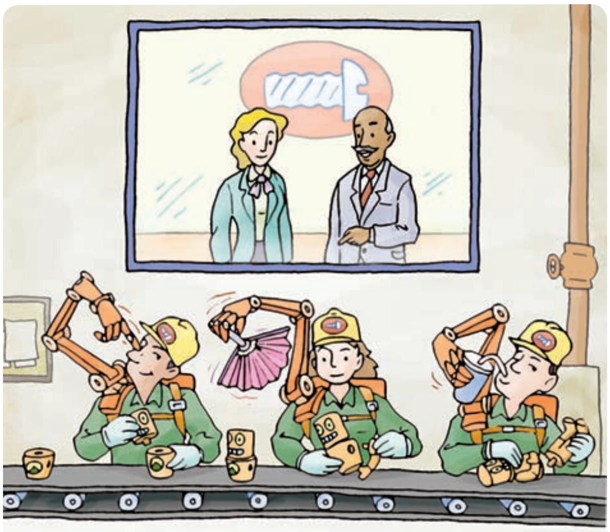
technical progress
an advance in knowledge that allows more output to be produced with the same level of inputs

Technical Progress A technological innovation changes the production process. Last year a firm produced

$$q_1 = f(L, K)$$

units of output using L units of labor services and K units of capital service. Due to a new invention, the firm's production function this year differs from last year's, so the firm produces 10% more output with the same inputs:

$$q_2 = 1.1\, f(L, K).$$



Surprisingly, those robo-arms increase productivity substantially.

For example, Flath (2011) estimated the annual rate of technical innovation in Japanese manufacturing firms to be 0.91% for electric copper, 0.87% for medicine, 0.33% for steel pipes and tubes, 0.19% for cement, and 0.08% for beer. Shao and Lin (2016) estimated that the annual average productivity growth rate was 7.4% in information technology (IT) goods and service industries in 12 developed countries.

This type of technological progress reflects *neutral technical change*, in which a firm can produce more output using the same ratio of inputs. For example, a technical innovation in the form of a new printing press may allow more output to be produced using the same ratio of inputs as before: one worker to one printing press. In our neutral technical change example, the firm's rate of growth of output was $10\% = \Delta q/q_1 = [1.1f(L, K) - f(L, K)]/f(L, K)$ in one year due to the technical progress.

Neutral technical progress leaves the shapes of the isoquants unchanged. However, each isoquant is now associated with more output. For example, in Figure 6.5, if neutral technical progress occurred that doubled output for any combination of inputs, we would relabel the isoquants from lowest to highest as $q = 2$, $q = 6$, $q = 12$, and $q = 16$.

---

[10]See MyLab Economics, Chapter 6, "German Versus British Productivity" and "U.S. Electric Generation Efficiency."

Non-neutral technical progress results from innovations that alter the proportion in which inputs are used. Technological progress could be *capital saving*, where relatively less capital is used relative to other inputs. For example, the development of cell phones allowed firms to eliminate enough landline phones, fax machines, and computers to lower the capital-labor ratio for its sales or repair workers while increasing output.

Alternatively, technological progress may be *labor saving*. Basker (2012) found that the introduction of barcode scanners in grocery stores increased the average product of labor by 4.5% on average across stores. By 2015, Amazon was using at least 15,000 robots at 10 of its U.S. warehouses to move items in its warehouses, replacing some workers. Today, robots help doctors perform surgery quicker and reduce patients' recovery times.

Application

Robots and the Food You Eat

Robots have been used in manufacturing for many years, and they are now gaining a foothold in agriculture. A strawberry-picking robot called the Agrobot costs about $100,000 and, despite the expense, is attracting buyers in California. The Hackney Nursery in Florida uses robots to assess whether flowers have adequate room to grow optimally and to move the flowers around accordingly. And fully autonomous cow-milking robots are widely used.

It is not just the farming end of the food business that is using robots. Restaurant robots date back to at least 2010. The Dalu Robot Restaurant in Jinan, China, uses robots to wait on tables, greet customers, and provide entertainment. Each robot serving food has a motion sensor that tells it to stop when someone is in its path so customers can reach for dishes they want. The customers' favorite employee is a female robot, complete with batting eyelashes, who greets people with an electronic "Welcome." First-time customer Li Xiaomei praised the robots, claiming that "they have a better service attitude than humans."

In 2016, KFC opened the world's first human-free fast food restaurant in Shanghai. However, restaurant jobs may not be on their way out just yet. In 2016, China's *Workers' Daily* newspaper reported that three restaurants in the Chinese city of Guangzhou fired their robotic staff for incompetence.

Organizational Change Organizational change may also alter the production function and increase the amount of output produced by a given amount of inputs. In the early 1900s, Henry Ford revolutionized mass production through two organizational innovations.

First, he introduced interchangeable parts, which cut the time required to install parts because workers no longer had to file or machine individually made parts to get them to fit.

Second, Ford introduced a conveyor belt and an assembly line to his production process. Before this innovation, workers walked around the car, and each worker performed many assembly activities. In Ford's plant, each worker specialized in a single activity such as attaching the right rear fender to the chassis. A conveyor belt moved the car at a constant speed from worker to worker along the assembly line. Because his workers gained proficiency from specializing in only a few activities, and because the conveyor belts reduced the number of movements workers had to make, Ford could produce more automobiles with the same number of workers.

These innovations reduced the ratio of labor to capital used. In 1908, the Ford Model T sold for \$850, when rival vehicles sold for \$2,000. By the early 1920s, Ford had increased production from fewer than a thousand cars per year to two million cars per year.

Application

A Good Boss Raises Productivity

Does a good supervisor make workers more productive? To answer this question, Lazear et al. (2015) looked at a large service-oriented company. Supervisor quality varied substantially as measured by the boss' effect on worker productivity. Replacing one of the 10% worst bosses with one of the 10% best ones raised a team's output by about the same amount as adding one worker to a nine-member team. Thus, differences in managers can cause one firm to be more productive than another.

Challenge Solution

Labor Productivity During Downturns

We can use what we've learned to answer the question posed at the beginning of the chapter about how labor productivity, as measured by the average product of labor, changes during a recession when a firm reduces its output by reducing the number of workers it employs. How much will the output produced per worker rise or fall with each additional layoff?

In the short run when the firm holds its capital constant, layoffs have the positive effect of freeing up machines to be used by the remaining workers. However, if layoffs mean that the remaining workers might have to "multitask" to replace departed colleagues, the firm will lose the benefits from specialization. When a firm has many workers, the advantage of freeing up machines is important and increased multitasking is unlikely to be a problem. With only a few workers, freeing up more machines does not help much—some machines might stand idle part of the time—while multitasking becomes a more serious problem. As a result, laying off a worker might raise the average product of labor if the firm has many workers relative to the available capital, but might reduce average product if it has only a few workers.

For example, in panel b of Figure 6.1, the average product of labor rises with the number of workers up to six workers and then falls as the number of workers increases beyond six. As a result, the average product of labor falls if the firm initially has two to six workers and lays one off, but rises if the firm initially has seven or more workers and lays off a worker.

However, for some production functions, layoffs always raise labor productivity because the AP_L curve is downward sloping everywhere. For such a production function, the positive effect of freeing up capital always dominates any negative effect of layoffs on the average product of labor.

Consider a Cobb-Douglas production function, $q = AL^aK^b$, where $AP_L = q/L = q = AL^{a-1}K^b$. As Appendix 6C shows, if we increase labor slightly, the change in the AP_L is $(a - 1)AL^{a-2}K^b$. Thus, if a is less than one, so that $a - 1$ is negative, the AP_L falls with extra labor. This condition holds for all of the estimated Cobb-Douglas production functions listed in the Application "Returns to Scale in Various Industries" (though not necessarily in all industries).

For example, for the beer firm's estimated Cobb-Douglas production function (Flath, 2011), $q = AL^{0.6}K^{0.4}$, $a = 0.6$ is less than 1, so the AP_L curve slopes downward at every quantity. We can illustrate how much the AP_L rises with a layoff for this particular production function. If $A = 1$ and $L = K = 10$ initially, then the firm's output is $q = 10^{0.6} \times 10^{0.4} = 10$, and its average product of labor is $AP_L = q/L = 10/10 = 1$. If the number of workers is reduced by one, then output falls to $q = 9^{0.6} \times 10^{0.4} \approx 9.39$, and the average product of labor rises to $AP_L \approx 9.39/9 \approx 1.04$. That is, a 10% reduction in labor causes output to *fall* by 6.1%, but causes the average product of labor to *rise* by 4%. The firm's output falls less than 10% because each remaining worker is more productive.

This increase in labor productivity in many industries reduces the impact of the recession on output in the United States. However, this increase in labor productivity is not always observed in other countries that are less likely to lay off workers during a downturn. Until recently, most large Japanese firms did not lay off workers during recessions. Thus, in contrast to U.S. firms, their average product of labor decreased substantially during recessions because their output fell while labor remained constant.

Similarly, European firms show 30% less employment volatility over time than do U.S. firms, at least in part because European firms that fire workers are subject to a tax (Veracierto, 2008).[11] Consequently, with other factors held constant in the short run, recessions might be more damaging to the profit and output of a Japanese or European firm than to the profit and output of a comparable U.S. firm. However, retaining good workers over short-run downturns might be a good long-run policy both for the firm and for its workers.

Summary

1. **The Ownership and Management of Firms.** The three types of firms are private, public, and nonprofit. Private firms may be sole proprietorships, partnerships, or corporations. In small firms (particularly sole proprietorships and partnerships), the owners usually run the company. In large firms (such as most corporations), the owners hire a manager to run the company. Owners want to maximize profits. If managers have different objectives than owners, owners must keep a close watch over managers to ensure that profits are maximized.

2. **Production.** Inputs, or factors of production—labor, capital, and materials—are combined to produce output using the current state of knowledge about technology and management. To maximize profits, a firm must produce as efficiently as possible: It must get the maximum amount of output from the inputs it uses, given existing knowledge. A firm may have access to many efficient production processes that use different combinations of inputs to produce a given level of output. New technologies or new forms of organization can increase the amount of output that can be produced from a given combination of inputs. A production function shows how much output can be produced efficiently from various levels of inputs. A firm can vary all its inputs in the long run but only some of its inputs in the short run.

3. **Short-Run Production.** In the short run, a firm cannot adjust the quantity of some inputs, such as capital. The firm varies its output by adjusting its variable inputs, such as labor. If all factors are fixed except labor, and a firm that was using very little labor increases its use of labor, its output may rise more than in proportion to the increase in labor because of greater specialization of workers. Eventually, however, as more workers are hired, the workers get in each other's way or wait to share equipment, so output increases by smaller and smaller amounts. This latter phenomenon is described by the law of diminishing marginal returns: The marginal product of an input—the extra output from the last unit of input—eventually decreases as more of that input is used, holding other inputs fixed.

4. **Long-Run Production.** In the long run, when all inputs are variable, firms can substitute between inputs. An isoquant shows the combinations of inputs

[11]Severance payments for blue-collar workers with ten years of experience may exceed one year of wages in some European countries, unlike in the United States.

that can produce a given level of output. The marginal rate of technical substitution is the absolute value of the slope of the isoquant and indicates how easily the firm can substitute one factor of production for another. Usually, the more of one input the firm uses, the more difficult it is to substitute that input for another input, so the marginal rates of technical substitution decreases.

5. **Returns to Scale.** If, when a firm increases all inputs in proportion, its output increases by the same proportion, the production process is said to exhibit constant returns to scale. If output increases less than in proportion to inputs, the production process has decreasing returns to scale; if it increases more than in proportion to inputs, it has increasing returns to scale. All three types of returns to scale are commonly seen in actual industries. Many production processes exhibit first increasing, then constant, and finally decreasing returns to scale as the size of the firm increases.

6. **Productivity and Technical Change.** Although all firms in an industry produce efficiently, given what they know and the institutional and other constraints they face, some firms may be more productive than others: They can produce more output from a given bundle of inputs. Technological progress allows a firm to produce a given level of output using fewer inputs than it did previously. Technological progress changes the production function.

Questions

Select questions are available on MyLab Economics*; * = answer appears at the back of this book;* **A** *= algebra problem;* **C** *= calculus problem.*

1. The Ownership and Management of Firms

1.1 Are firms with limited liability likely to be larger than other firms? Why?

1.2 What types of firms would not normally maximize profit?

1.3 What types of organizations allow owners of a firm to obtain the advantages of limited liability?

2. Production Function

2.1 With respect to production functions, how long is the *short run*?

2.2 China, Japan, and South Korea have the top three shipbuilding industries in the world. Each of these countries also has a vibrant tourism industry. Of the two, which industry has the longer short run? Which has the shorter long run? Explain.

*2.3 Suppose that for the production function $q = f(L, K)$, if $L = 3$ and $K = 5$, then $q = 10$. Is it possible that $L = 3$ and $K = 6$ also produces $q = 10$ for this production function? Why or why not?

3. Short-Run Production

*3.1 Suppose that a firm's production function is $q = 2KL$, where L is labor services and K is capital services, and that $K = 3$. What are the total product, average product of labor, and marginal product of labor curves?

3.2 If the marginal product of labor is 2 units of output per day per worker for the first five workers and 0 units of output for the sixth worker onward, draw the total product curve, the average product of labor curve, and the marginal product of labor curve.

3.3 If the production function is $q = f(L, K) = 3L + 2K$, and capital is fixed at $\overline{K} = 50$, what is the short-run production function? What is the marginal product of labor? (*Hint*: See Solved Problem 6.1.)

*3.4 Suppose that the production function is $q = L^{0.75}K^{0.25}$.

a. What is the average product of labor, holding capital fixed at $\overline{K}$? **A**

b. What is the marginal product of labor? (*Hints*: See Solved Problem 6.1. Calculate how much q changes as L increases by 1 unit for a particular pair of K and L or use calculus [see Appendix 6C].) **C**

3.5 Why do we expect the law of diminishing marginal returns to hold?

3.6 Ben swims 50,000 yards per week in his practices. Given this amount of training, he will swim the 100-yard butterfly in 52.6 seconds and place tenth in a big upcoming meet. Ben's coach calculates that if Ben increases his practice to 60,000 yards per week, his time will decrease to 50.7 seconds and he will place eighth in the meet. If Ben practices 70,000 yards per week, his time will be 49.9 and he will win the meet.

a. In terms of Ben's *time* in the big meet, what is his marginal productivity of the number of yards he practices? Does the marginal product diminish as the practice yards increase?

b. In terms of Ben's *place* in the big meet, what is his marginal productivity of the number of yards he practices? Does the marginal product diminish as the practice yards increase?

c. Does Ben's marginal productivity of the number of yards he practices depend on how he measures his productivity, either place or time, in the big meet?

3.7 What is the law of diminishing marginal returns? Will a firm experience diminishing marginal returns in the short run if its production function is:

a. $q = K+L$?

b. $q = KL$?

c. $q = KL^{0.5}$? **C**

3.8 Based on the information in the Application "Malthus and the Green Revolution," how did the average product of labor in corn production change over time?

4. Long-Run Production

4.1 In the long run, a firm can usually substitute one input for another while continuing to produce the same level of output. If such a firm produces a given level of output using the same amount of capital but a greater amount of labor, what can we say about how the firm is producing? Why?

4.2 What is the difference between an isoquant and an indifference curve?

4.3 Why must isoquants be thin? (*Hint*: See the explanation of why indifference curves are thin in Chapter 4.)

4.4 Suppose that a firm cannot substitute capital for labor in production, and to produce every two units of output, it must use one unit of capital and two units of labor. Draw the isoquants for this production function as well as the total product curve, the average product of labor curve, and the marginal product of labor curve.

4.5 According to Card (2009), (a) workers with less than a high school education are perfect substitutes for those with a high school education, (b) "high school equivalent" and "college equivalent" workers are imperfect substitutes, and (c) within education groups, immigrants and natives are imperfect substitutes. For each of these comparisons, draw the isoquants for a production function that uses two types of workers. For example, in part (a), production is a function of workers with a high school diploma and workers with less education.

4.6 Suppose that capital and labor are perfect substitutes in production, and a firm is presently using two units of capital and two units of labor to produce four units of output. What is the firm's production function?

*4.7 At $L = 4$ and $K = 4$, the marginal product of labor is 2 and the marginal product of capital is 3. What is the marginal rate of technical substitution? **A**

*4.8 Mark launders his white clothes using the production function $q = B + 0.5G$, where B is the number of cups of Clorox bleach and G is the number of cups of generic bleach that is half as potent. Draw an isoquant. What are the marginal products of B and G? If B is on the vertical axis, what is the marginal rate of technical substitution at each point on an isoquant? (*Hint*: See Solved Problem 6.2.)

*4.9 To produce a recorded CD, $q = 1$, a firm uses one blank disk, $D = 1$, and the services of a recording machine, $M = 1$, for one hour. Draw an isoquant for this production process. Explain the reason for its shape.

4.10 The production function at Ginko's Copy Shop is $q = 1{,}000 \times \min(L, 3K)$, where q is the number of copies per hour, L is the number of workers, and K is the number of copy machines. As an example, if $L = 4$ and $K = 1$, then $\min(L, 3K) = 3$, and $q = 3{,}000$.

a. Draw the isoquants for this production function.

b. Draw the total product, average product, and marginal product of labor curves for this production function for some fixed level of capital.

4.11 Using the figure in the Application "A Semiconductor Integrated Circuit Isoquant," show that as the firm employs additional fixed-proportion technologies, the firm's overall isoquant approaches a smooth curve.

4.12 By studying, Will can produce a higher grade, G_W, on an upcoming economics exam. His production function depends on the number of hours he studies marginal analysis problems, A, and the number of hours he studies supply-and-demand problems, R. Specifically, $G_W = 2.5A^{0.36}R^{0.64}$. His roommate David's grade-production function is $G_D = 2.5A^{0.25}R^{0.75}$.

a. What is Will's marginal product of studying supply-and-demand problems? What is David's? (*Hint*: See Appendix 6C.)

b. What is Will's marginal rate of technical substitution between studying the two types of problems? What is David's?

c. Is it possible that Will and David have different marginal productivity functions but the same marginal rate of technical substitution functions? Explain. (*Hint*: See the Section "Cobb-Douglas Production Function.") **C**

4.13 Electric power is often generated by burning oil or gas to create steam. That steam is used to drive the turbines and produce electricity. One barrel of crude oil produces about 5.6 million BTUs of energy, while 1,000 cubic feet of natural gas produces 1,027,000 BTUs (**http://www.physics.uci.edu/~silverma/units.html**). Thus, an electric generating company can substitute 1 barrel of crude oil with 5,648 cubic feet of natural gas. Draw a few isoquants for this production process. What is the marginal rate of technical substitution? **A**

5. Returns to Scale

5.1 Haiti was hit by a devastating earthquake on January 12, 2010. A quarter of a million lives were lost and millions more were left injured and homeless. Many countries sent medical teams and medical supplies, among other items, to assist in the relief efforts. Identify the outputs and inputs, and describe the production process and the shape of the isoquants.

5.2 Michelle's business produces ceramic cups using labor, clay, and a kiln. She can manufacture 25 cups a day with one worker and 35 with two workers. Does her production process necessarily illustrate *decreasing returns to scale* or *diminishing marginal returns to labor*? What is the likely explanation for why output doesn't increase proportionately with the number of workers?

5.3 Show in a diagram that a production function can have diminishing marginal returns to a factor and constant returns to scale.

5.4 Under what conditions do the following production functions exhibit decreasing, constant, or increasing returns to scale?

a. $q = L + K$

b. $q = L^a K^b$

c. $q = L + L^a K^b + K$ (*Hint*: See Solved Problem 6.3.) **A**

*5.5 What are returns to scale? Under what conditions (that is, for what values of the parameters a and b) does the Cobb-Douglas production function, $q = K^a L^b$, exhibit constant and increasing returns to scale? (*Hint*: See Solved Problem 6.3.) **A**

5.6 A production function is said to be *homogeneous of degree g* if $f(xL, xK) = x^g f(L, K)$, where x is a positive constant. That is, the production function has the same returns to scale for every combination of inputs. For such a production function, show that the marginal product of labor and marginal product of capital functions are homogeneous of degree $g - 1$. **C**

5.7 The Application "Returns to Scale in Various Industries" provides estimated parameter values for the Cobb-Douglas production function, $q = L^a K^b$, for firms in various industries. Use the estimates for the Danish food and beverages firm, the Japanese rubber firm, and the New Zealand mining firm to explore the relationship between returns to scale and diminishing marginal returns. If $a = 1$ and $b = 1$, does the law of diminishing marginal returns still hold? **A**

6. Productivity and Technical Change

6.1 If the average product of labor is higher for Firm 1 than for Firm 2, must Firm 1 be more productive in the sense that it can produce more output from a given amount of inputs? Why?

*6.2 Firm 1 and Firm 2 use the same type of production function, but Firm 1 is only 90% as productive as Firm 2. That is, the production function of Firm 2 is $q_2 = f(L, K)$, and the production function of Firm 1 is $q_1 = 0.9f(L, K)$. At a particular level of inputs, how does the marginal product of labor differ between the firms? **C**

6.3 After a technical or managerial innovation, a firm can produce more today from a given amount of inputs than it could in the past. In farming, for example, new seed varieties, fertilizers, machines, and livestock feeds have greatly increased agricultural output over the past century even though the number of farmers has fallen in many countries. Was this technical progress neutral? In a figure, show how these technical changes affected isoquants.

6.4 If a new type of capital is introduced into a production process such that a firm can use fewer workers and still produce the same level of output, what type of technological progress is this? Suppose that the firm uses the Cobb-Douglas production function, $q = (\beta K)^{0.5} L^{0.5}$, where β represents the level of this type of technological progress. Before the innovation, the firm used 16 units of labor and 4 units of capital, and $\beta = 1$. With the innovation, it uses 4 units of both labor and capital to produce the same level of output, and $\beta = 4$. What is the effect of this invention on the average and marginal product of labor?

6.5 Are the robots in the Application "Robots and the Food You Eat" an example of neutral, labor-saving, or capital-saving innovation? Explain.

6.6 Is a boss a fixed or variable input in the Application "A Good Boss Raises Productivity"? How does having a good boss affect the marginal product of labor curve for this firm? Assuming that the production process also includes a capital input, what effect does a good boss have on a typical isoquant?

7. Challenge

7.1 How would the answer in the Challenge Solution change if we used the marginal product of labor rather than the average product of labor as our measure of labor productivity?

*7.2 Different firms respond to recessions in different ways. Some will reduce their workforce and produce less in response to a fall in the demand for their product. Others may continue to produce at prerecession levels of output but store output they cannot sell immediately. Consider two identical firms for which capital and labor are perfect substitutes in production. During a recession, Firm 1 lays off half of its workforce while Firm 2 does not and continues to produce the same level of output. Would you expect the average product of labor to be higher for Firm 1 or Firm 2? Why?

7

Costs

People want economy and they will pay any price to get it.
—Lee Iacocca (former CEO of Chrysler)

Challenge

Technology Choice at Home Versus Abroad

A manager of a semiconductor-manufacturing firm, who can choose from many different production technologies, must determine whether to use the same technology in its foreign plant that it uses in its domestic plant. U.S. semiconductor manufacturing firms have been moving much of their production abroad since 1961, when Fairchild Semiconductor built a plant in Hong Kong. According to the Semiconductor Industry Association, worldwide semiconductor billings from the Americas dropped from 66% in 1976, to 34% in 1998, to 17% in 2011, and then rose to 18% by early 2016.

Semiconductor firms are moving their production abroad because of lower taxes, lower labor costs, and capital grant benefits. Capital grants are funds provided by a foreign government to a firm to induce them to produce in that country. Such grants can reduce the cost of owning and operating an overseas semiconductor fabrication facility by as much as 25% compared to the costs of a U.S.-based plant.

The semiconductor manufacturer can produce a chip using sophisticated equipment and relatively few workers or many workers and less complex equipment. In the United States, firms use a relatively capital-intensive technology, because doing so minimizes their cost of producing a given level of output. Will that same technology be cost minimizing if they move their production abroad?

A firm uses a two-step procedure in determining how to produce a certain amount of output efficiently. It first determines which production processes are *technologically efficient* so that it can produce the desired level of output with the least amount of inputs. As we saw in Chapter 6, the firm uses engineering and other information to determine its production function, which summarizes the many technologically efficient production processes available.

economically efficient minimizing the cost of producing a specified amount of output

The firm's second step is to pick from these technologically efficient production processes the one that is also **economically efficient**, minimizing the cost of producing a specified amount of output. To determine which process minimizes its cost of production, the firm uses information about the production function and the cost of inputs.

By reducing its cost of producing a given level of output, a firm can increase its profit. Therefore, any profit-maximizing competitive, monopolistic, or oligopolistic firm minimizes its cost of production.

In this chapter, we examine five main topics

1. **The Nature of Costs.** When considering the cost of a proposed action, a good manager of a firm takes account of forgone alternative opportunities.
2. **Short-Run Costs.** To minimize its costs in the short run, a firm adjusts its variable factors (such as labor), but it cannot adjust its fixed factors (such as capital).
3. **Long-Run Costs.** In the long run, a firm adjusts all its inputs because usually all inputs are variable.
4. **Lower Costs in the Long Run.** Long-run cost is as low as or lower than short-run cost because in the long run, the firm has more flexibility, technological progress occurs, and workers and managers learn from experience.
5. **Cost of Producing Multiple Goods.** If the firm produces several goods simultaneously, the cost of each may depend on the quantity of all the goods produced.

Businesspeople and economists need to understand the relationship between costs of inputs and production to determine the least costly way to produce. Economists have an additional reason for wanting to know about costs. As we'll see in later chapters, the relationship between output and costs plays an important role in determining the nature of a market—the number of firms in the market and the level of price relative to cost.

7.1 The Nature of Costs

How much would it cost you to stand at the wrong end of a shooting gallery?
—S. J. Perelman

To show how a firm's cost varies with its output, we must first measure costs. Businesspeople and economists often measure costs differently. Economists include all relevant costs. To run a firm profitably, a manager should think like an economist and consider all relevant costs. Although a skilled manager acknowledges this point, the manager may direct the firm's accountant or bookkeeper to measure costs in ways that are more consistent with tax laws and other laws and to make the firm's financial statements look better to stockholders.

To produce a particular amount of output, a firm incurs costs for the required inputs such as labor, capital, energy, and materials. A firm's manager (or accountant) determines the cost of labor, energy, and materials by multiplying the factor's price by the number of units used. If workers earn \$20 per hour and work a total of 100 hours per day, then the firm's cost of labor is $\$20 \times 100 = \$2,000$ per day. The manager can easily calculate these *explicit costs*, which are its direct, out-of-pocket payments for inputs to its production process within a given period. While calculating explicit costs is straightforward, some costs are *implicit* in that they reflect only a forgone opportunity rather than an explicit, current expenditure. Properly taking account of forgone opportunities requires particularly careful attention when dealing with durable capital goods, as past expenditures for an input may be irrelevant to current cost calculations if that input has no current, alternative use.

Opportunity Costs

An economist is a person who, when invited to give a talk at a banquet, tells the audience there's no such thing as a free lunch.

economic cost or opportunity cost the value of the best alternative use of a resource

The **economic cost or opportunity cost** is the value of the best alternative use of a resource. The economic or opportunity cost includes both explicit and implicit costs. If a firm purchases and uses an input immediately, that input's opportunity cost is the amount the firm pays for it. However, if the firm does not use the input in its production process, its best alternative would be to sell it to someone else at the market price.

The concept of an opportunity cost becomes particularly useful when the firm uses an input that is not available for purchase in a market or that it purchased in a market in the past. A key example of such an opportunity cost is the value of a manager's time. For instance, Maoyong owns and manages a firm. He pays himself only a small monthly salary of $1,000 because he also receives the firm's profit. However, Maoyong could work for another firm and earn $11,000 a month. Thus, the opportunity cost of his time is $11,000—from his best alternative use of his time—not the $1,000 he actually pays himself.

The phrase "there's no such thing as a free lunch" provides the classic example of an implicit opportunity cost. Suppose that your parents offer to take you to lunch tomorrow. You know that they will pay for the meal, but you also know that this lunch will not truly be free. Your opportunity cost for the lunch is the best alternative use of your time.

Presumably, the best alternative use of your time is reading this chapter, but other possible alternatives include working at a job or watching TV. Often, such an opportunity cost is substantial. (What are you giving up to study opportunity costs?)

At one point or another, most of us have held the following false belief:

> **Common Confusion:** I can save money by doing things myself rather than buying goods and services from firms.

The fallacy in this belief is that we have ignored the opportunity cost of our time. Have you ever tried to fix a plumbing problem and ended up taking hours where a professional plumber could have repaired it in a few minutes? Doing that only makes sense if the opportunity cost of your time is very low (or the plumber's fee is very high). Similarly, growing our own food would cost most of us much more than buying it from a store once we take into account the value of our time.

Application

The Opportunity Cost of an MBA

During the sharp economic downturn in 2008–2010, did applications to MBA programs fall, hold steady, or take off like tech stocks during the 2016 Internet bubble? We can answer this question using the opportunity cost concept.

For many potential MBA students, the biggest cost of attending an MBA program is the opportunity cost of giving up a well-paying job. Someone who leaves a job that pays $5,000 per month to attend an MBA program has a $5,000-per-month implicit cost from forgone earnings, in addition to the explicit cost of tuition and textbooks.

Thus, it is not surprising that MBA applications rise in bad economic times when outside opportunities decline. People thinking of going back to school face a reduced opportunity cost of entering an MBA program if they think their firms may lay them off or not promote them during an economic downturn. As Stacey Kole, deputy dean for the MBA program at the University of Chicago Graduate School of Business observed, "When there's a go-go economy, fewer people decide to go back to school. When things go south, the opportunity cost of leaving work is lower."

In 2008, when U.S. unemployment rose sharply and the economy was in poor shape, the number of people seeking admission to MBA programs shot up substantially. The number of applicants to MBA programs for the class of 2008–2009 increased over the previous year by 79% in the United States, 77% in the United Kingdom, and 69% in other European countries. Applicants also increased substantially for 2009–2010 in Canada and Europe. However, as economic conditions improved, global applications were unchanged in 2010, fell slightly in 2011, and remained relatively low through 2015.

Solved Problem 7.1

MyLab Economics Solved Problem

Meredith's firm sends her to a conference for managers and has paid her registration fee. Included in the registration fee is free admission to a class on how to price derivative securities such as options. She is considering attending, but her most attractive alternative opportunity is to attend a talk by Warren Buffett about his investment strategies, which will take place at the same time. Although she would be willing to pay $100 to hear his talk, the cost of a ticket is only $40. Given she incurs no other costs to attend either event, what is Meredith's opportunity cost of attending the derivatives talk?

Answer

To calculate her opportunity cost, determine the benefit that Meredith would forgo by attending the derivatives class. Because she incurs no additional fee to attend the derivatives talk, Meredith's opportunity cost is the forgone benefit of hearing the Buffett speech. Because she values hearing the Buffett speech at $100, but only has to pay $40, her net benefit from hearing that talk is $60 (= $100 − $40). Thus, her opportunity cost of attending the derivatives talk is $60.

Opportunity Cost of Capital

Capital: *Something—like a car, refrigerator, factory, or airplane—that is blown up in an action flick.*

durable good
a product that is usable for years

Determining the opportunity cost of capital, such as land or equipment, requires special considerations. Capital is a **durable good**: a product that is usable for years. Two problems may arise in measuring the cost of capital. The first is how to allocate the initial purchase cost over time. The second is what to do if the value of the capital changes over time.

We can avoid these two measurement problems if the firm rents capital rather than purchasing it. For example, suppose a firm can rent a small pickup truck for $400 a month or buy it outright for $20,000. If the firm rents the truck, the rental payment is the relevant opportunity cost per month. Because the firm rents the truck month to month, the firm does not have to worry about how to allocate the purchase cost of a truck over time. Moreover, the rental rate will adjust if the cost of trucks changes over time. Thus, if the firm can rent capital for short periods, it calculates the cost of this capital in the same way that it calculates the cost of nondurable inputs such as labor services or materials.

The firm faces a more complex problem in determining the opportunity cost of the truck if it purchases the truck. The firm's accountant may *expense* the truck's purchase price by treating the full $20,000 as a cost at the time of purchase.

Alternatively, the accountant may *amortize* the cost by spreading the $20,000 over the life of the truck, following rules set by an accounting organization or by a relevant government authority such as the Internal Revenue Service (IRS).

A manager who wants to make sound decisions does not expense or amortize the truck using such rules. The true opportunity cost of using a truck that the firm owns is the amount that the firm could earn if it rented the truck to others. That is, regardless of whether the firm rents or buys the truck, the manager views the opportunity cost of this capital good as the rental rate for a given period. If the value of an older truck is less than that of a newer one, the rental rate for the truck falls over time.

But what is the opportunity cost if no one rents trucks? Suppose that the firm has two choices: It can choose not to buy the truck and keep the truck's purchase price of $20,000, or it can use the truck for a year and sell it for $17,000 at the end of the year. If the firm does not purchase the truck, it will deposit the $20,000 in a bank account that pays 5% per year, so the firm will have $21,000 at the end of the year. Thus, the opportunity cost of capital of using the truck for a year is $\$21,000 - \$17,000 = \$4,000$.[1] This $4,000 opportunity cost equals the $3,000 depreciation of the truck ($= \$20,000 - \$17,000$) plus the $1,000 in forgone interest that the firm could have earned over the year if the firm had invested the $20,000.

Because the values of trucks, machines, and other equipment decline over time, their rental rates fall, so the firm's opportunity costs decline. In contrast, the value of some land, buildings, and other forms of capital may rise over time. To maximize profit, a firm must properly measure the opportunity cost of a piece of capital even if its value rises over time. If a beauty parlor buys a building when similar buildings in the area rent for $1,000 per month, the opportunity cost of using the building is $1,000 a month. If property values increase so that rents in the area rise to $2,000 per month, the beauty parlor's opportunity cost of its building rises to $2,000 per month.

Sunk Costs

There's no use in crying over spilt milk.

sunk cost
a past expenditure that cannot be recovered

An opportunity cost is not always easy to observe but should always be taken into account when deciding how much to produce. In contrast, a **sunk cost**—a past expenditure that cannot be recovered—though easily observed, is not relevant to a firm when deciding how much to produce now. If an expenditure on a fixed input is sunk, it is not an opportunity cost.

If a firm buys a forklift for $25,000 and can resell it for the same price, it is not a sunk expenditure, and the opportunity cost of the forklift is $25,000. If instead the firm buys a specialized piece of equipment for $25,000 and cannot resell it, then the original expenditure is a sunk cost. Because this equipment has no alternative use, its opportunity cost is zero, and it should not be included in the firm's current cost calculations. If the firm can resell the specialized equipment that originally cost $25,000 for $10,000, then only $15,000 of the original expenditure is a sunk cost, and the opportunity cost is $10,000.

To illustrate why a sunk cost should not influence a manager's current decisions, consider a firm that paid $300,000 for a piece of land for which the market value has fallen to $200,000. Now, the land's true opportunity cost is $200,000. The $100,000 difference between the $300,000 purchase price and the current market value of

[1]The firm would also pay for gasoline, insurance, licensing fees, and other operating costs, but these items would all be expensed as operating costs and would not appear in the firm's accounts as capital costs.

\$200,000 is a sunk cost that the firm cannot recover. The land is worth \$240,000 to the firm if it builds a plant on this parcel. Is it worth carrying out production on this land, or should the land be sold for its market value of \$200,000? If the firm uses the original purchase price in its decision-making process, the firm will incorrectly conclude that using the land for production will result in a \$60,000 loss: the \$240,000 value of using the land minus the purchase price of \$300,000. Instead, the firm should use the land because it is worth \$40,000 more as a production facility than if the firm sells the land for \$200,000, which is its next best alternative. Thus, in making its decisions, the firm should use the land's opportunity cost and ignore the land's sunk cost.

7.2 Short-Run Costs

To make profit-maximizing decisions, a firm needs to know how its cost varies with output. A firm's cost rises as it increases its output. A firm cannot vary some of its inputs, such as capital, in the short run (Chapter 6). As a result, it is usually more costly for a firm to increase output in the short run than in the long run, when all inputs can be varied. In this section, we look at the cost of increasing output in the short run.

Short-Run Cost Measures

We start by using a numerical example to illustrate the basic cost concepts. We then examine the graphic relationship between these concepts.

fixed cost (F)
a production expense that does not vary with output

variable cost (VC)
a production expense that changes with the quantity of output produced

cost (total cost, C)
the sum of a firm's variable cost and fixed cost: $C = VC + F$

Fixed Cost, Variable Cost, and Total Cost To produce a given level of output in the short run, a firm incurs costs for both its fixed and variable inputs. A firm's **fixed cost** (F) is its production expense that does not vary with output. The fixed cost includes the cost of inputs that the firm cannot practically adjust in the short run, such as land, a plant, large machines, and other capital goods. The fixed cost for a capital good a firm owns and uses is the opportunity cost of not renting it to someone else. The fixed cost is \$48 per day for the firm in Table 7.1.

A firm's **variable cost** (VC) is the production expense that changes with the quantity of output produced. The variable cost is the cost of the variable inputs—the inputs the firm can adjust to alter its output level, such as labor and materials. Table 7.1 shows that the firm's variable cost changes with output. Variable cost goes from \$25 a day when 1 unit is produced to \$46 a day when 2 units are produced.

A firm's **cost** (or **total cost**, C) is the sum of a firm's variable cost and fixed cost:

$$C = VC + F.$$

The firm's total cost of producing 2 units of output is \$94 per day, which is the sum of the fixed cost, \$48, and the variable cost, \$46. Because variable cost changes with the level of output, total cost also varies with the level of output, as the table illustrates.

To decide how much to produce, a firm uses several measures of how its cost varies with the level of output. Table 7.1 shows four such measures that we derive using the fixed cost, the variable cost, and the total cost. These measures are marginal cost, average fixed cost, average variable cost, and average cost.

Fixed

Variable

Table 7.1 Variation of Short-Run Cost with Output

Output, q	Fixed Cost, F	Variable Cost, VC	Total Cost, C	Marginal Cost, MC	Average Fixed Cost, $AFC = F/q$	Average Variable Cost, $AVC = VC/q$	Average Cost, $AC = C/q$
0	48	0	48				
1	48	25	73	25	48	25	73
2	48	46	94	21	24	23	47
3	48	66	114	20	16	22	38
4	48	82	130	16	12	20.5	32.5
5	48	100	148	18	9.6	20	29.6
6	48	120	168	20	8	20	28
7	48	141	189	21	6.9	20.1	27
8	48	168	216	27	6	21	27
9	48	198	246	30	5.3	22	27.3
10	48	230	278	32	4.8	23	27.8
11	48	272	320	42	4.4	24.7	29.1
12	48	321	369	49	4.0	26.8	30.8

Application

The Sharing Economy and the Short Run

A construction company views workers' earnings as a variable cost and the capital that the firm owns—particularly heavy equipment such as bulldozers—as a fixed cost. The sharing economy is changing that.

When Platinum Pipeline Inc., a firm that installs water and sewer lines, won a new job, it needed a third bulldozer. Rather than buy one, the firm's president, Manuel de Freitas, merely called up an app on his phone and found a Caterpillar D6T dozer that he could rent for two months at $7,500 a month. The rental firm, Yard Club Inc., finds idle heavy equipment and rents it—much like Airbnb Inc. does with spare bedrooms. Often, rental companies own this equipment.

Renting construction equipment is catching on. In 2014, rental companies owned 54% of U.S. construction equipment, up from 40% a decade earlier. According to one forecast, the share could top 60% within the next 5 to 10 years.

If construction companies can rely on renting heavy equipment rather than owning it, all their inputs are variable. As a result, they face no distinction between the short run and the long run.

marginal cost (MC) the amount by which a firm's cost changes if the firm produces one more unit of output

Marginal Cost A firm's **marginal cost** (MC) is the amount by which a firm's cost changes if the firm produces one more unit of output. The marginal cost is[2]

$$MC = \frac{\Delta C}{\Delta q},$$

[2]If we use calculus, the marginal cost is $MC = dC(q)/dq$, where $C(q)$ is the cost function that shows how cost varies with output. The calculus definition shows how the cost changes for an infinitesimal change in output. To illustrate the idea, however, we use larger changes in Table 7.1.

where ΔC is the change in cost when output changes by Δq. Table 7.1 shows that, if the firm increases its output from 2 to 3 units, $\Delta q = 1$, its total cost rises from \$94 to \$114, $\Delta C = \$20$, so its marginal cost is $\$20 = \Delta C/\Delta q$.

Because only variable cost changes with output, the marginal cost is also equal to the change in variable cost from a one-unit increase in output:

$$MC = \frac{\Delta VC}{\Delta q}.$$

As the firm increases output from 2 to 3 units, its variable cost increases by $\Delta VC = \$20 = \$66 - \$46$, so its marginal cost is $MC = \Delta VC/\Delta q = \$20$. A firm uses marginal cost in deciding whether it pays to change its output level.

average fixed cost (*AFC*) the fixed cost divided by the units of output produced: $AFC = F/q$

average variable cost (*AVC*) the variable cost divided by the units of output produced: $AVC = VC/q$

average cost (*AC*) the total cost divided by the units of output produced: $AC = C/q$

Average Costs Firms use three average cost measures. The **average fixed cost** (***AFC***) is the fixed cost divided by the units of output produced: $AFC = F/q$. The average fixed cost falls as output rises because the fixed cost is spread over more units. The average fixed cost falls from \$48 for 1 unit of output to \$4 for 12 units of output in Table 7.1.

The **average variable cost** (***AVC***) is the variable cost divided by the units of output produced: $AVC = VC/q$. Because the variable cost increases with output, the average variable cost may either increase or decrease as output rises. The average variable cost is \$25 at 1 unit, falls until it reaches a minimum of \$20 at 6 units, and then rises. As we show in Chapter 8, a firm uses the average variable cost to determine whether to shut down operations when demand is low.

The **average cost** (***AC***)—or *average total cost*—is the total cost divided by the units of output produced: $AC = C/q$. The average cost is the sum of the average fixed cost and the average variable cost:[3]

$$AC = AFC + AVC.$$

In Table 7.1, as output increases, average cost falls until output is 8 units and then rises. The firm makes a profit if its average cost is below its price, which is the firm's average revenue.

Short-Run Cost Curves

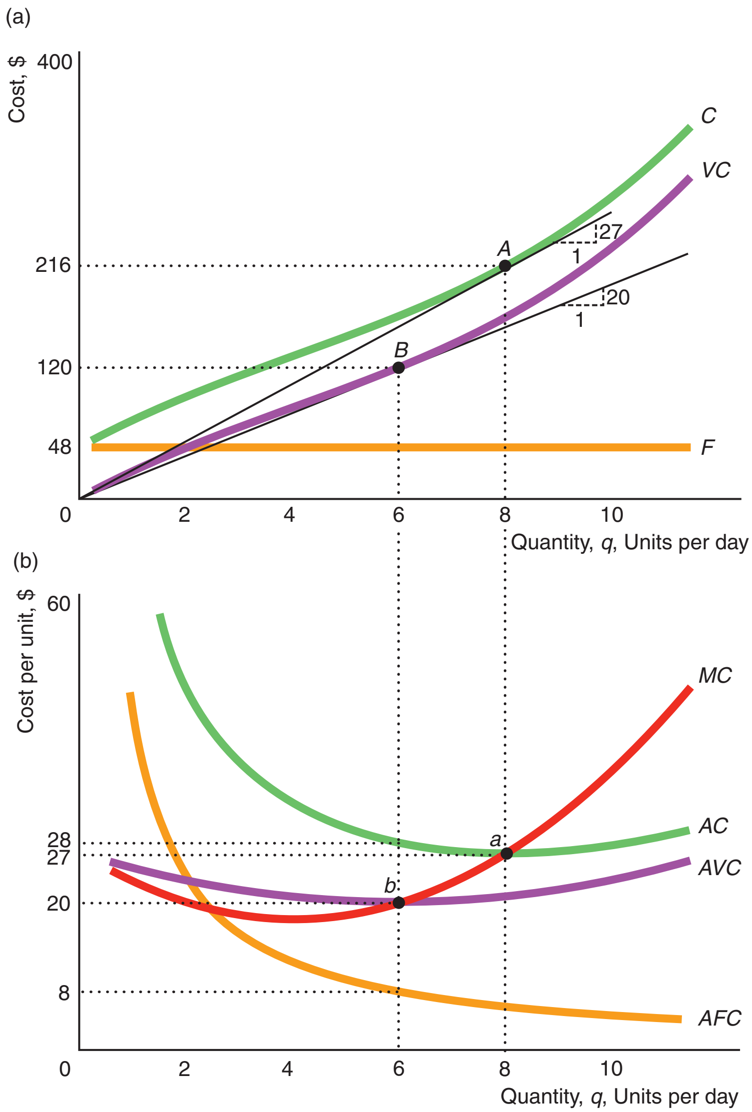
We illustrate the relationship between output and the various cost measures using curves in Figure 7.1. Panel a shows the variable cost, fixed cost, and total cost curves that correspond to Table 7.1. The fixed cost, which does not vary with output, is a horizontal line at \$48. The variable cost curve is zero at zero units of output and rises with output. The total cost curve, which is the vertical sum of the variable cost curve and the fixed cost line, is \$48 higher than the variable cost curve at every output level, so the variable cost and total cost curves are parallel.

Panel b shows the average fixed cost, average variable cost, average cost, and marginal cost curves. The average fixed cost curve falls as output increases. It approaches zero as output gets large because the fixed cost is spread over many units of output. The average cost curve is the vertical sum of the average fixed cost and average variable cost curves. For example, at 6 units of output, the average variable cost is 20 and the average fixed cost is 8, so the average cost is 28.

[3]Because $C = VC + F$, if we divide both sides of the equation by q, we obtain $AC = C/q = F/q + VC/q = AFC + AVC$.

Figure 7.1 Short-Run Cost Curves MyLab Economics Video

(a) Because the total cost differs from the variable cost by the fixed cost, *F*, of \$48, the total cost curve, *C*, is parallel to the variable cost curve, *VC*. (b) The marginal cost curve, *MC*, intersects the average variable cost, *AVC*, and average cost, *AC*, curves at their minimums. The height of the *AC* curve at point *a* equals the slope of the line from the origin to the cost curve at *A*. The height of the *AVC* at *b* equals the slope of the line from the origin to the variable cost curve at *B*. The height of the marginal cost is the slope of either the *C* or *VC* curve at that quantity.



The relationships of the total cost curves to the average and marginal cost curves are similar to those between the total product, marginal product, and average product curves, which we discussed in Chapter 6. The average cost at a particular output level is the slope of a line from the origin to the corresponding point on the cost curve. The slope of that line is the rise—the cost at that output level—divided by the run—the output level—which is the definition of the average cost. In panel a, the slope of the line from the origin to point *A* is the average cost for 8 units of output. The height of the cost curve at *A* is 216, so the slope is $216/8 = 27$, which is the height of the average cost curve at the corresponding point *a* in panel b.

Similarly, the average variable cost is the slope of a line from the origin to a point on the variable cost curve. The slope of the line from the origin to *B* in panel a is

20—the height of the variable cost curve, 120, divided by the number of units of output, 6—which is the height of the average variable cost at 6 units of output, point *b* in panel b.

The marginal cost is the slope of either the cost curve or the variable cost curve at a given output level. As the cost and variable cost curves are parallel, they have the same slope at any given output. The difference between cost and variable cost is fixed cost, which does not affect marginal cost.

Panel a of Figure 7.1 shows a line from the origin that is tangent to the cost curve at *A*. Thus, the slope of this line equals both the average cost and the marginal cost at 8 units of output. This equality occurs at the corresponding point *a* in panel b, where the marginal cost curve intersects the average cost. (See Appendix 7A for a mathematical proof.)

Where the marginal cost curve is below the average cost, the average cost curve slopes downward. Because the average cost of 47 for 2 units is greater than the marginal cost of the third unit, 20, the average cost for 3 units falls to 38. Where the marginal cost is above the average cost, the average cost curve rises with output. At 8 units, the marginal cost equals the average cost, so the average is unchanging, which is the minimum point, *a*, of the average cost curve.

We can show the same results using the graph. Because the line from the origin is tangent to the variable cost curve at *B* in panel a, the marginal cost equals the average variable cost at the corresponding point *b* in panel b. Again, where marginal cost is above average variable cost, the average variable cost curve rises with output; where marginal cost is below average variable cost, the average variable cost curve falls with output. Because the average cost curve is above the average variable cost curve everywhere and the marginal cost curve is rising where it crosses both average curves, the minimum of the average variable cost curve, *b*, is at a lower output level than the minimum of the average cost curve, *a*.

Production Functions and the Shape of Cost Curves

The production function determines the shape of a firm's cost curves. The production function shows the amount of inputs needed to produce a given level of output. The firm calculates its cost by multiplying the quantity of each input by its price and then summing the costs of the those inputs.

If a firm produces output using capital and labor, the firm's variable cost is its cost of labor in the short run because the firm cannot vary its capital. Its labor cost is the wage per hour, w, times the number of hours of labor, L, employed by the firm: $VC = wL$.

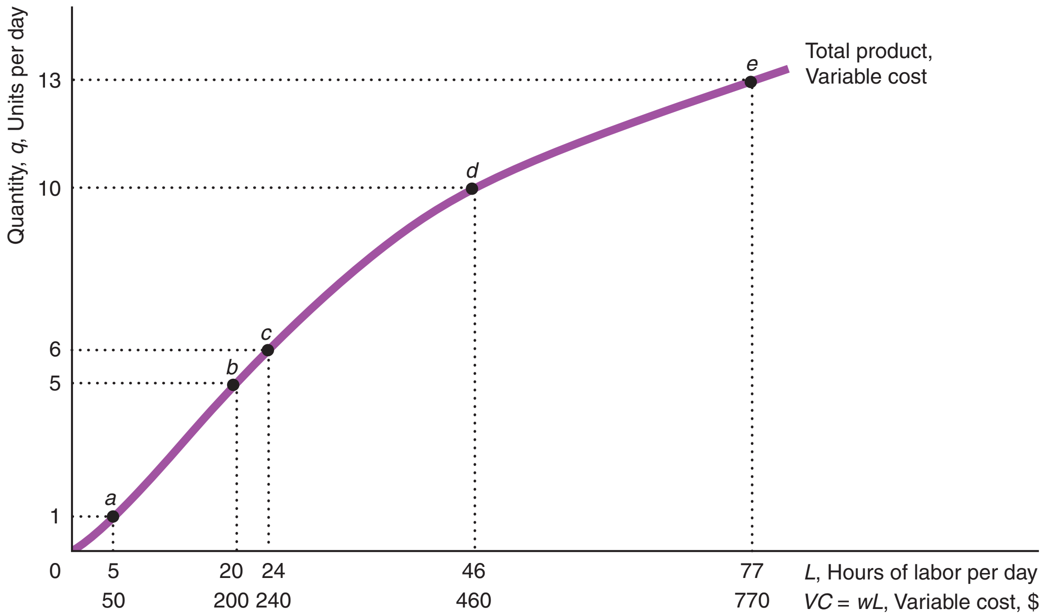
In the short run, when the firm's capital is fixed, the only way the firm can increase its output is to use more labor. If the firm increases its labor enough, it reaches the point of *diminishing marginal returns to labor*, at which each extra worker increases output by a smaller amount. We can use this information about the relationship between labor and output—the production function—to determine the shape of the variable cost curve and its related curves.

Shape of the Variable Cost Curve If input prices are constant, the production function determines the shape of the variable cost curve. We illustrate this relationship for the firm in Figure 7.2. The firm faces a constant input price for labor, the wage, of $10 per hour.

The total product of labor curve in Figure 7.2 shows the firm's short-run production function relationship between output and labor holding capital fixed. At point *a*,

Figure 7.2 Variable Cost and Total Product of Labor

The firm's short-run variable cost curve and its total product of labor curve have the same shape. The total product of labor curve uses the horizontal axis measuring hours of work. The variable cost curve uses the horizontal axis measuring labor cost, which is the only variable cost.



the firm uses 5 hours of labor to produce 1 unit of output. At point *b*, it takes 20 hours of labor to produce 5 units of output. Here, output increases more than in proportion to labor: Output rises 5 times when labor increases 4 times. In contrast, as the firm moves from *b* to *d*, output increases less than in proportion to labor. When labor increases 2.3 times from 20 to 46, output doubles to 10. The movement from *d* to *e* results in an even smaller increase in output relative to labor. This flattening of the total product curve at higher levels of labor reflects diminishing marginal returns to labor.

This curve shows both the production relationship between output and labor and the variable cost relationship between output and cost. Because each hour of work costs the firm $10, we can use an alternative measure for the horizontal axis in Figure 7.2: the firm's variable cost, which is its cost of labor. To produce 5 units of output takes 20 hours of labor, so the firm's variable cost is $200. By using the variable cost labels on the horizontal axis, the total product of labor curve becomes the variable cost curve.

As output increases, the variable cost increases more than proportionally due to the diminishing marginal returns. Because the production function determines the shape of the variable cost curve, it also determines the shape of the marginal, average variable, and average cost curves. When making decisions, firms rely more on these per-unit cost measures than on total variable cost. We now examine the shape of each of these cost curves in detail.

Shape of the Marginal Cost Curve The marginal cost is the change in variable cost as output increases by one unit: $MC = \Delta VC/\Delta q$. If the firm cannot vary capital in the short run, the only way it can increase output is by using extra labor. The extra labor required to produce one more unit of output is $\Delta L/\Delta q$. The extra labor costs the firm w per unit, so the firm's cost rises by $w(\Delta L/\Delta q)$. As a result, the firm's marginal cost is

$$MC = \frac{VC}{q} = w\frac{L}{q}.$$

The marginal cost equals the wage times the extra labor necessary to produce one more unit of output. To increase output by one unit from 5 to 6 units takes 4 extra hours of work in Figure 7.2. If the wage is \$10 per hour, the marginal cost of the last unit of output is \$40.

How do we know how much extra labor we need to produce one more unit of output? That information comes from the production function. The marginal product of labor—the amount of extra output produced by another unit of labor, holding other inputs fixed—is $MP_L = \Delta q/\Delta L$. Thus, the extra labor we need to produce one more unit of output, $\Delta L/\Delta q$, is $1/MP_L$, so the firm's marginal cost is

$$MC = \frac{w}{MP_L}. \qquad (7.1)$$

Equation 7.1 says that the marginal cost equals the wage divided by the marginal product of labor. In Figure 7.2 to increase its output from 5 to 6 units, the firm must use 4 extra hours of labor, so its marginal product of an hour of labor is $\frac{1}{4}$. Given a wage of \$10 an hour, the marginal cost of the sixth unit is \$10 divided by $\frac{1}{4}$, or \$40.

Equation 7.1 shows that the marginal cost moves in the opposite direction of the marginal product of labor. At low levels of labor, the marginal product of labor commonly rises with additional labor because extra workers help the original workers and they can collectively make better use of the firm's equipment (Chapter 6). As the marginal product of labor rises, the marginal cost falls.

Eventually, however, as the number of workers increases, workers must share the fixed amount of equipment and may get in each other's way, so the marginal cost curve slopes upward because of diminishing marginal returns to labor. Thus, the marginal cost first falls and then rises, as panel b of Figure 7.1 illustrates.

Shape of the Average Cost Curves Diminishing marginal returns to labor, which determine the shape of the variable cost curve, also determine the shape of the average variable cost curve. The average variable cost is the variable cost divided by output: $AVC = VC/q$. For a firm that can vary labor but not capital in the short run, its variable cost is wL, so its average variable cost is

$$AVC = \frac{VC}{q} = \frac{wL}{q}.$$

Because the average product of labor is q/L, average variable cost is the wage divided by the average product of labor:

$$AVC = \frac{w}{AP_L}. \qquad (7.2)$$

Using the example in Figure 7.2, at 5 units of output, the average variable cost is \$40, which is the wage, \$10, divided by the average product of labor, which is $\frac{1}{4}$ $(= q/L = 5/20)$.

With a constant wage, the average variable cost moves in the opposite direction of the average product of labor in Equation 7.2. As we discussed in Chapter 6, the average product of labor tends to rise and then fall, so the average cost tends to fall and then rise, as in panel b of Figure 7.1.

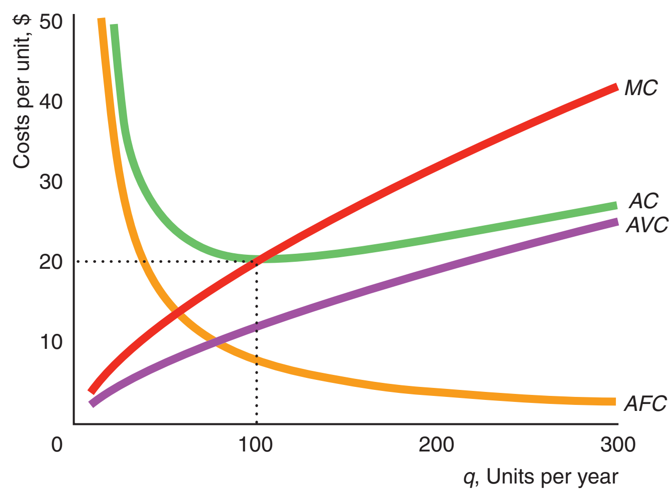
The average cost curve is the vertical sum of the average variable cost curve and the average fixed cost curve, as in panel b. If the average variable cost curve is U-shaped, adding the strictly falling average fixed cost makes the average cost fall more steeply than the average variable cost curve at low output levels. As output increases, the gap between the average cost and the average variable cost curves shrinks, as the average fixed cost, F/q, approaches zero. Thus, the average cost curve is also U-shaped.

Application

A Beer Manufacturer's Short-Run Cost Curves

The short-run average cost curve for a Japanese beer manufacturer is U-shaped, even though its average variable cost is strictly upward sloping. The graph (based on the estimates of Flath, 2011) shows the firm's various short-run cost curves, where the firm's capital is fixed at $\overline{K} = 100$. Appendix 7B derives the firm's short-run cost curves mathematically.

The firm's average fixed cost (AFC) falls as output increases. The firm's average variable cost curve is strictly increasing. The average cost (AC) curve is the vertical sum of the average variable cost (AVC) and average fixed cost curves. Because the average fixed cost curve falls with output and the average variable cost curve rises with output, the average cost curve is U-shaped. The firm's marginal cost (MC) lies above the rising average variable cost curve for all positive quantities of output and crosses the average cost curve at its minimum.



Effects of Taxes on Costs

Taxes applied to a firm shift some or all of the marginal and average cost curves. For example, suppose that the government collects a specific tax of $10 per unit of output from the firm. This tax, which varies with output, affects the firm's variable cost but not its fixed cost. As a result, it affects the firm's average cost, average variable cost, and marginal cost curves but not its average fixed cost curve.

At every quantity, the average variable cost, average cost, and marginal cost rise by the full amount of the tax. In the example in Table 7.1, the firm's average variable cost before the tax, AVC^b, is $21 if it produces 8 units of output. After the tax, the firm must pay the government $10 per unit, so the firm's after-tax average variable cost rises to $31. More generally, the firm's after-tax average variable cost, AVC^a, is its average variable cost of production—the before-tax average variable cost—plus the tax per unit, $10: $AVC^a = AVC^b + \$10$.

The average cost equals the average variable cost plus the average fixed cost. Because the tax increases average variable cost by \$10 and does not affect the average fixed cost, the tax increases average cost from $AC^b = \$27$ to $AC^a = \$37$.

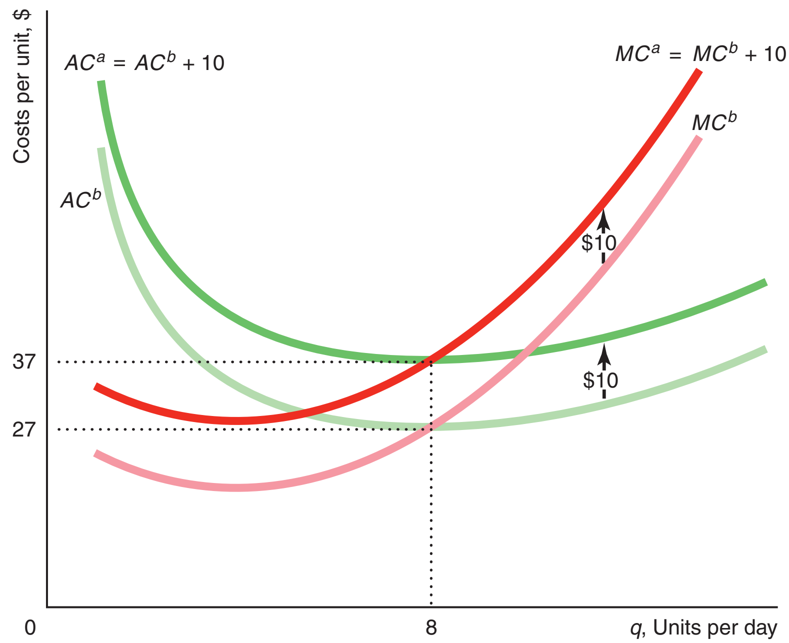
The tax also increases the firm's marginal cost. For example, in Table 7.1, its before-tax marginal cost, MC^b, is \$27 at 8 units of output. To produce this last unit of output, the cost to the firm is the before-tax marginal cost of producing the extra unit, $MC^b = \$27$, plus \$10, so its after-tax marginal cost is $MC^a = MC^b + \$10 = \$37$.

Figure 7.3 shows that the \$10 specific tax shifts the firm's average cost and marginal cost curves up by the amount of the tax. The after-tax marginal cost intersects the after-tax average cost at its minimum. Because both the marginal and average cost curves shift upward by exactly the same amount, both the before-tax and after-tax average cost curves reach their minimums at 8 units of output, as Figure 7.3 shows. So even though a specific tax increases a firm's average cost, it does not affect the output at which it minimizes its average cost.

Similarly, we can analyze the effect of a franchise tax on costs. A *franchise tax*—also called a *business license fee*—is a lump sum that a firm pays for the right to operate a business. For example, a four-year license for a cart used to sell hot dogs outside the former Tavern on the Green restaurant in New York City's Central Park cost \$1.39 million in 2012. These taxes do not vary with output, so they affect firms' fixed costs only—not their variable costs.

Figure 7.3 Effect of a Specific Tax on Cost Curves MyLab Economics Video

A specific tax of \$10 per unit shifts both the marginal cost and average cost curves upward by \$10. Because of the parallel upward shift of the average cost curve, the minimums of both the before-tax average cost curve, AC^b, and the after-tax average cost curve, AC^a, occur at the same output, 8 units.

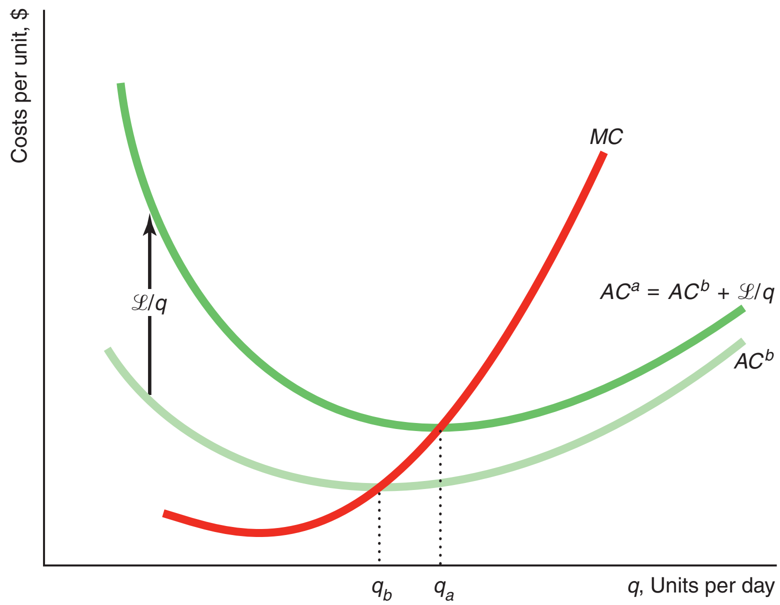


Solved Problem 7.2

What is the effect of a lump-sum franchise tax $\mathcal{L}$ on the quantity at which a firm's after-tax average cost curve reaches its minimum? (Assume that the firm's before-tax average cost curve is U-shaped.)

Answer

1. *Determine the average tax per unit of output.* Because the franchise tax is a lump-sum payment that does not vary with output, the more the firm produces, the less tax it pays per unit. The tax per unit is $\mathcal{L}/q$. If the firm sells only 1 unit, its cost is $\mathcal{L}$; however, if it sells 100 units, its tax payment per unit is only $\mathcal{L}/100$.



2. *Show how the tax per unit affects the average cost.* The firm's after-tax average cost, AC^a, is the sum of its before-tax average cost, AC^b, and its average tax payment per unit, $\mathcal{L}/q$. Because the average tax payment per unit falls with output, the gap between the after-tax average cost curve and the before-tax average cost curve also falls with output on the graph.

3. *Determine the effect of the tax on the marginal cost curve.* Because the franchise tax does not vary with output, it does not affect the marginal cost curve.

4. *Compare the minimum points of the two average cost curves.* The marginal cost curve crosses from below both average cost curves at their minimum points. Because the after-tax average cost curve lies above the before-tax average cost curve, the quantity at which the after-tax average cost curve reaches its minimum, q_a, is larger than the quantity, q_b, at which the before-tax average cost curve achieves a minimum.

Short-Run Cost Summary

We have discussed three cost-level curves—total cost, fixed cost, and variable cost—and four cost-per-unit curves—average cost, average fixed cost, average variable cost, and marginal cost. Understanding the shapes of these curves and the relationships between them is crucial to understanding the analysis of firm behavior in the rest of this book. Fortunately, we can derive most of what we need to know about the shapes of and the relationships between the curves using four basic concepts:

- In the short run, the cost associated with inputs that the firm cannot adjust is fixed, while the cost from inputs that it can adjust is variable.
- Given that input prices are constant, the production function determines the shapes of the variable cost and cost curves.
- Where a variable input exhibits diminishing marginal returns, the variable cost and cost curves become relatively steep as output increases, so the average cost, average variable cost, and marginal cost curves rise with output.
- Because of the relationship between a marginal and an average, both the average cost and average variable cost curves fall when marginal cost is below them and rise when marginal cost is above them. Thus, the marginal cost curve crosses both these average cost curves at their minimum points.

7.3 Long-Run Costs

In the long run, the firm adjusts all its inputs so that its cost of production is as low as possible. The firm can change its plant size, design and build new machines, and otherwise adjust inputs that were fixed in the short run.

All Costs Are Avoidable in the Long Run

Although firms may incur fixed costs in the long run, these fixed costs are *avoidable* (rather than *sunk*, as in the short run). The rent of F per month that a restaurant pays is a fixed cost because it does not vary with the number of meals (output) served. In the short run, this fixed cost is sunk: The firm must pay F even if the restaurant does not operate. In the long run, the firm can avoid this fixed cost by shutting down. The length of the rental contract determines the length of the short run.

In the examples throughout this chapter, we assume that all inputs can be varied in the long run so that the firm has no long-run fixed costs ($F = 0$). As a result, the long-run total cost equals the long-run variable cost: $C = VC$. Thus, our firm is concerned about only three cost concepts in the long run—total cost, average cost, and marginal cost—instead of the seven cost concepts that it considers in the short run.

Minimizing Cost

To produce a given quantity of output at minimum cost, our firm uses information about the production function and the price of labor and capital. The firm chooses how much labor and capital to use in the long run, whereas the firm chooses only how much labor to use in the short run when capital is fixed. Consequently, the firm's long-run cost is lower than its short-run cost of production if it has to use the "wrong" level of capital in the short run. In this section, we show how a firm picks the cost-minimizing combinations of inputs in the long run.

A firm can produce a given level of output using many different *technologically efficient* combinations of inputs, as summarized by an isoquant (Chapter 6). From among the technologically efficient combinations of inputs, a firm wants to choose the particular bundle with the lowest cost of production, which is the *economically efficient* combination of inputs. To do so, the firm combines information about technology from the isoquant with information about the cost of labor and capital.

We now show how an *isocost line* summarizes information about cost. Then, we show how a firm can combine the information in an isoquant and isocost lines to determine the economically efficient combination of inputs.

Isocost Line

The cost of producing a given level of output depends on the price of labor and capital. The firm hires L hours of labor services at a wage of w per hour, so its labor cost is wL. The firm rents K hours of machine services at a rental rate of r per hour, so its capital cost is rK. (If the firm owns the capital, r is the implicit rental rate.) The firm's total cost is the sum of its labor and capital costs:

$$C = wL + rK. \tag{7.3}$$

The firm can hire as much labor and capital as it wants at these constant input prices.

The firm can use many combinations of labor and capital that cost the same amount. Suppose that the wage rate, w, is \$10 an hour and the rental rate of capital, r, is \$20. Table 7.2 lists five of the many combinations of labor and capital the firm can use that cost \$200. We plot these combinations of labor and capital on an **isocost line**, which is all the combinations of inputs that require the same (*iso*) total expenditure (*cost*). Figure 7.4 shows three isocost lines. The \$200 isocost line represents all the combinations of labor and capital that the firm can buy for \$200, including the combinations a through e in Table 7.2.

isocost line
all the combinations of inputs that require the same (*iso*) total expenditure (*cost*)

Along an isocost line, cost is fixed at a particular level, $\overline{C}$, so by setting cost at $\overline{C}$ in Equation 7.3, we can write the equation for the $\overline{C}$ isocost line as

$$\overline{C} = wL + rK.$$

Using algebra, we can rewrite this equation to show how much capital the firm can buy if it spends a total of $\overline{C}$ and purchases L units of labor:

$$K = \frac{\overline{C}}{r} - \frac{w}{r}L. \tag{7.4}$$

By substituting $\overline{C} = \$200$, $w = \$10$, and $r = \$20$ in Equation 7.4, we find that the \$200 isocost line is $K = 10 - \frac{1}{2}L$. We can use Equation 7.4 to derive three properties of isocost lines.

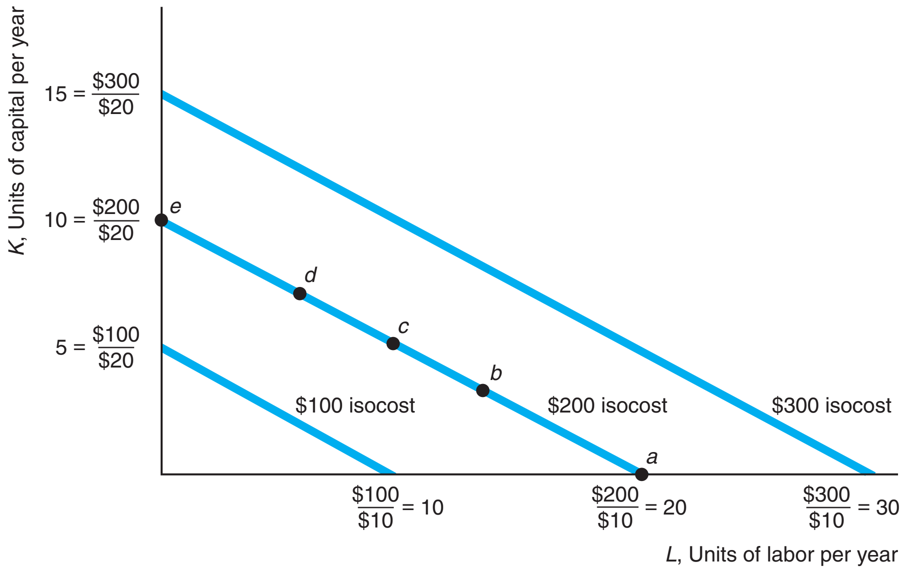
First, where the isocost lines hit the capital and labor axes depends on the firm's cost, $\overline{C}$, and on the input prices. The $\overline{C}$ isocost line intersects the capital axis where the firm is using only capital. Setting $L = 0$ in Equation 7.4, we find that the firm buys $K = \overline{C}/r$ units of capital. In the figure, the \$200 isocost line intersects the capital axis at \$200/\$20 = 10 units of capital. Similarly, the intersection of the isocost line

Table 7.2 Bundles of Labor and Capital That Cost the Firm \$200

Bundle	Labor, L	Capital, K	Labor Cost, $wL = \$10L$	Capital Cost, $rK = \$20K$	Total Cost, $wL + rK$
a	20	0	\$200	\$0	\$200
b	14	3	\$140	\$60	\$200
c	10	5	\$100	\$100	\$200
d	6	7	\$60	\$140	\$200
e	0	10	\$0	\$200	\$200

Figure 7.4 A Family of Isocost Lines

An isocost line shows all the combinations of labor and capital that cost the firm the same amount. The greater the total cost, the farther from the origin the isocost line lies. All the isocost lines have the same slope, $-w/r = -\frac{1}{2}$. The slope shows the rate at which the firm can substitute capital for labor holding total cost constant: For each extra unit of capital it uses, the firm must use two fewer units of labor to hold its cost constant.



with the labor axis is at $\overline{C}/w$, which is the amount of labor the firm hires if it uses only labor. In the figure, the intersection of the \$200 isocost line with the labor axis occurs at $L = 20$, where $K = 10 - \frac{1}{2} \times 20 = 0$.

Second, isocost lines that are farther from the origin have higher costs than those that are closer to the origin. Because the isocost lines intersect the capital axis at $\overline{C}/r$ and the labor axis at $\overline{C}/w$, an increase in the cost shifts these intersections with the axes proportionately outward. The \$100 isocost line hits the capital axis at 5 and the labor axis at 10, whereas the \$200 isocost line intersects at 10 and 20.

Third, the slope of each isocost line is the same. From Equation 7.4, if the firm increases labor by ΔL, it must decrease capital by

$$\Delta K = -\frac{w}{r}\Delta L.$$

Dividing both sides of this expression by ΔL, we find that the slope of an isocost line, $\Delta K/\Delta L$, is $-w/r$. Thus, the slope of the isocost line depends on the relative prices of the inputs. The slope of the isocost lines in the figure is $-w/r = -\$10/\$20 = -\frac{1}{2}$. If the firm uses two more units of labor, $\Delta L = 2$, it must reduce capital by one unit, $\Delta K = -\frac{1}{2}\Delta L = -1$, to keep its total cost constant. Because all the isocost lines use the same prices, they all have the same slope, so they are parallel.

The isocost line plays a similar role in the firm's decision making as the budget line does in consumer decision making. Both an isocost line and a budget line are straight lines, the slopes of which depend on relative prices. However, the isocost and budget lines have an important difference. The consumer has a single budget line determined by the consumer's income. The firm faces many isocost lines, each of which corresponds to a different level of expenditures the firm might make. A firm may incur a relatively low cost by producing relatively little output with few inputs, or it may incur a relatively high cost by producing a relatively large quantity.

Combining Cost and Production Information

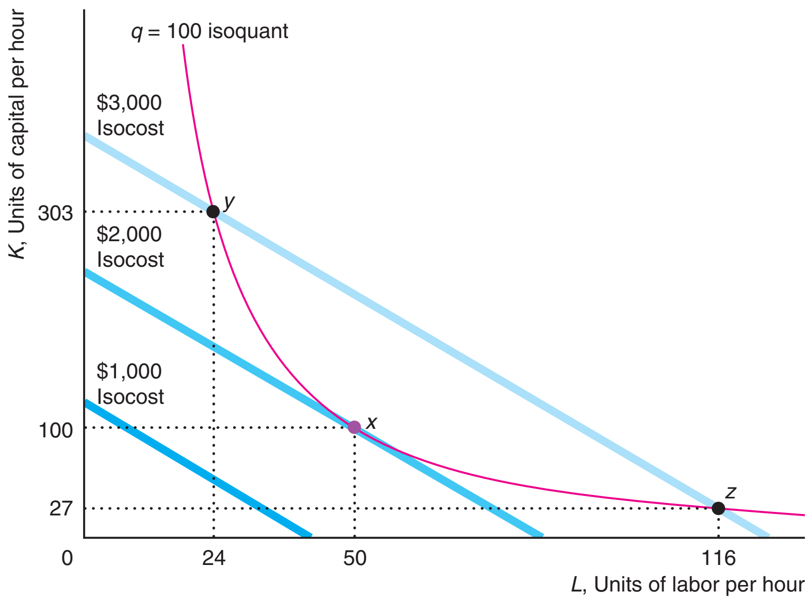
By combining the information about costs contained in isocost lines with information about efficient production summarized by an isoquant, a firm chooses the lowest-cost way to produce a given level of output. We examine how a beer manufacturer picks the combination of labor and capital that minimizes its cost of producing 100 units of output. Figure 7.5 shows the isoquant for 100 units of output (based on Flath, 2011) and three isocost lines where the rental rate of a unit of capital is \$8 per hour and the wage rate is \$24 per hour.

The firm can choose any of three equivalent approaches to minimize its cost:

- **Lowest-isocost rule.** Pick the bundle of inputs where the lowest isocost line touches the isoquant.
- **Tangency rule.** Pick the bundle of inputs where the isoquant is tangent to the isocost line.
- **Last-dollar rule.** Pick the bundle of inputs where the last dollar spent on one input gives as much extra output as the last dollar spent on any other input.

Figure 7.5 Cost Minimization MyLab Economics Video

The beer manufacturer minimizes its cost of producing 100 units of output by producing at x ($L = 50$ and $K = 100$). This cost-minimizing combination of inputs is determined by the tangency between the $q = 100$ isoquant and the lowest isocost line, \$2,000, that touches that isoquant. At x, the isocost line is tangent to the isoquant, so the slope of the isocost line, $-w/r = -3$, equals the slope of the isoquant, which is the negative of the marginal rate of technical substitution. That is, the rate at which the firm can trade capital for labor in the input markets equals the rate at which it can substitute capital for labor in the production process.



Lowest-Isocost Rule Using the *lowest-isocost rule*, the firm minimizes its cost by using the combination of inputs on the isoquant that is on the lowest isocost line that touches the 100-unit isoquant. In Figure 7.5, the \$2,000 isocost line touches the isoquant at input Bundle x, where the firm uses $L = 50$ workers and $K = 100$ units of capital. We want to show that Bundle x is the least costly way to produce 100 units of output. We need to demonstrate that other combinations of inputs either produce less than 100 units or produce 100 units at greater cost.

Because the \$2,000 isocost line just touches the isoquant, any lower isocost line, such as the \$1,000 isocost line, does not touch the isoquant. Each combination of inputs on the \$1,000 isocost line lies below the isoquant, so the firm cannot produce 100 units of output for \$1,000.

The firm can produce 100 units of output using combinations of inputs other than Bundle x; however, using these other bundles of inputs is more expensive. For example, the firm can produce 100 units of output using the combinations y ($L = 24$, $K = 303$) or z ($L = 116$, $K = 28$). However, both these combinations lie on the \$3,000 isocost line, so producing 100 units of output with these bundles costs more than using Bundle x.

Tangency Rule If an isocost line crosses the isoquant twice, as the \$3,000 isocost line does, another lower isocost line must also touch the isoquant. The lowest possible isocost line that touches the isoquant, the \$2,000 isocost line, is tangent to the isoquant at a single bundle, x. Consequently, the firm can identify the low-cost bundle using the *tangency rule*: The firm chooses the input bundle where the relevant isoquant is tangent to an isocost line to produce a given level of output at the lowest cost.

At the point of tangency, the slope of the isoquant equals the slope of the isocost line. As we showed in Chapter 6, the slope of the isoquant is the marginal rate of technical substitution ($MRTS$). The isocost line's slope is the negative of the ratio of the wage to the cost of capital, $-w/r$. Thus, to minimize its cost of producing a given level of output, a firm chooses its inputs so that the marginal rate of technical substitution equals the negative of the relative input prices:[4]

$$MRTS = -\frac{w}{r}. \tag{7.5}$$

The firm picks inputs so that the rate at which it can substitute capital for labor in the production process, the $MRTS$, exactly equals the rate at which it can trade capital for labor in input markets, $-w/r$.

In Figure 7.5, the slope of the isocost line is $-w/r = -24/8 = -3$. At Bundle x, the slope of the isoquant line, its $MRTS$, is also -3. In contrast, at y, the $MRTS$ is -18.9375, which is steeper than the isocost's slope of $-w/r = -3$. Because the slopes are not equal at y, the firm can produce 100 units of output at lower cost.

Solved Problem 7.3

MyLab Economics Solved Problem

Use the tangency rule to determine the cost-minimizing bundles of labor and capital for a general Cobb-Douglas production function, $q = AL^aK^b$, and for the specific beer production function that underlies Figure 7.5, $q = 1.516L^{0.6}K^{0.4}$, where $w = 24$ and $r = 8$.

[4]Appendix 7C derives Equation 7.7 using calculus.

Answer

1. *Use the general Cobb-Douglas production function and Equation 7.5 to determine the tangency condition for the general Cobb-Douglas*. According to Equation 6.7, the slope of a Cobb-Douglas isoquant is $MRTS = -(a/b)(K/L)$. According to Equation 7.5, the firm minimizes its cost if its $MRTS$ equals the slope of its isocost line, $-w/r$:

$$-\frac{a}{b}\frac{K}{L} = -\frac{w}{r}.$$

   Rearranging this expression, we find that, at the cost-minimizing bundle,

$$K = \frac{wb}{ra}L. \tag{7.6}$$

2. *Substitute the specific values for our beer example into Equation 7.6*. For the beer production function, $a = 0.6$ and $b = 0.4$. Substituting those values and $w = 24$ and $r = 8$ into Equation 7.6, we find that

$$K = \frac{24 \times 0.4}{8 \times 0.6}L = 2L. \tag{7.7}$$

   That is, at a cost-minimizing input bundle, the firm uses twice as many units of K as of L.

Last-Dollar Rule We can interpret the tangency rule, Equation 7.5, in another way. We showed in Chapter 6, Equation 6.3, that the marginal rate of technical substitution equals the negative of the ratio of the marginal product of labor to that of capital: $MRTS = -MP_L/MP_K$. Thus, the cost-minimizing condition in Equation 7.5 is equivalent to $-MP_L/MP_K = -w/r$. We can rewrite this expression as

$$\frac{MP_L}{w} = \frac{MP_K}{r}. \tag{7.8}$$

Equation 7.8 is the *last-dollar rule*: The firm minimizes its cost if the last dollar it spends on labor adds as much extra output, MP_L/w, as the last dollar it spends on capital, MP_K/r.

Using Equations 6.5 and 6.6, we know that the beer firm's marginal product of labor is $MP_L = 0.6q/L$ and its marginal product of capital is $MP_K = 0.4q/K$. At x ($L = 50$, $K = 100$), where $q = 100$, the beer firm's marginal product of labor is 1.2 (= 0.6 × 100/50) and its marginal product of capital is 0.4 (= 0.4 × 100/100). The last dollar spent on labor produces

$$\frac{MP_L}{w} = \frac{1.2}{24} = 0.05$$

more units of output. The last dollar spent on capital leads to

$$\frac{MP_K}{r} = \frac{0.4}{8} = 0.05$$

extra output. Thus, spending one more dollar on labor at x gets the firm as much extra output as spending the same amount on capital. Equation 7.8 holds, so the firm is minimizing its cost of producing 100 units of output.

If instead the firm produced at y ($L = 24$, $K = 303$), where it is using more capital and less labor, its MP_L is 2.5 (= 0.6 × 100/24) and the MP_K is approximately

0.13 ($\approx 0.4 \times 100/303$). As a result, the last dollar spent on labor generates $MP_L/w \approx 0.1$ more units of output, whereas the last dollar spent on capital gets only a fourth as much extra output, $MP_K/r \approx 0.017$. At y, if the firm shifts one dollar from capital to labor, output falls by 0.017 because of the reduced capital, but increases by 0.1 because of the extra labor, so the firm has a net gain of 0.083 more output at the same cost. The firm should shift even more resources from capital to labor—which increases the marginal product of capital and decreases the marginal product of labor—until Equation 7.8 holds with equality at x.

To summarize, a firm has three equivalent rules that it can use to pick the lowest-cost combination of inputs to produce a given level of output when isoquants are smooth: the lowest-isocost rule, the tangency rule (Equations 7.5), and the last-dollar rule (Equation 7.8). If the isoquant is not smooth, we cannot determine the lowest-cost method of production using the tangency rule or the last-dollar rule. The lowest-isocost rule always works—even when isoquants are not smooth (as in Chapter 6's Application "A Semiconductor Integrated Circuit Isoquant").

Factor Price Changes

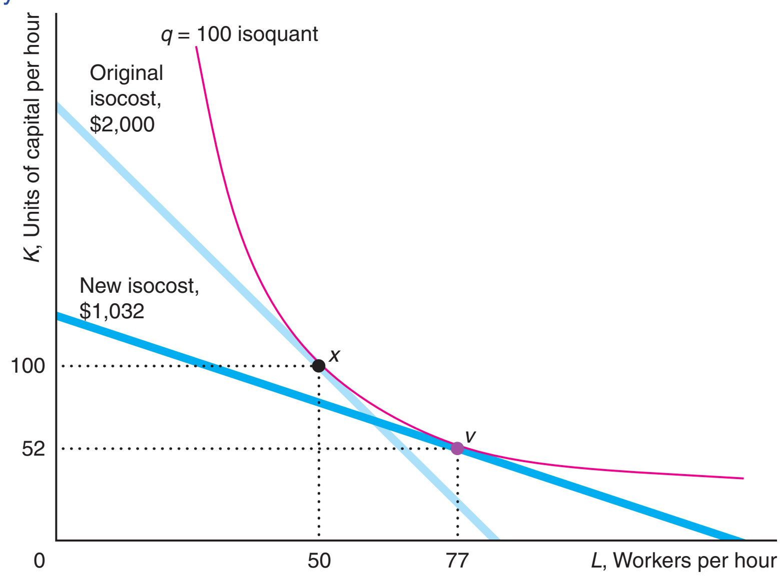
Once the beer manufacturer determines the lowest-cost combination of inputs to produce a given level of output, it uses that combination as long as the input prices remain constant. How should the firm change its behavior if the cost of one of the factors changes? Suppose that the wage falls from \$24 to \$8 but the rental rate of capital stays constant at \$8.

The firm minimizes its new cost by substituting away from the now relatively more expensive input, capital, toward the now relatively less expensive input, labor. The change in the wage does not affect technological efficiency, so it does not affect the isoquant in Figure 7.6. Because of the wage decrease, the new isocost lines have a flatter slope, $-w/r = -8/8 = -1$, than the original isocost lines, $-w/r = -24/8 = -3$.

The relatively steep original isocost line is tangent to the 100-unit isoquant at Bundle x ($L = 50$, $K = 100$). The new, flatter isocost line is tangent to the isoquant

Figure 7.6 Change in Factor Price MyLab Economics Video

Originally, the wage was \$24 and the rental rate of capital was \$8, so the lowest isocost line (\$2,000) was tangent to the $q = 100$ isoquant at x ($L = 50$, $K = 100$). When the wage fell to \$8, the isocost lines became flatter: Labor became relatively less expensive than capital. The slope of the isocost lines falls from $-w/r = -24/8 = -3$ to $-8/8 = -1$. The new lowest isocost line (\$1,032) is tangent at v ($L = 77$, $K = 52$). Thus, when the wage falls, the firm uses more labor and less capital to produce a given level of output, and the cost of production falls from \$2,000 to \$1,032.



at Bundle v ($L = 77$, $K = 52$). Thus, the firm uses more labor and less capital as labor becomes relatively less expensive. Moreover, the firm's cost of producing 100 units falls from \$2,000 to \$1,032 because of the decrease in the wage. This example illustrates that a change in the relative prices of inputs affects the mix of inputs that a firm uses.

Solved Problem 7.4

MyLab Economics Solved Problem

If a firm manufactures in its home country, it faces input prices for labor and capital of $\hat{w}$ and $\hat{r}$ and produces $\hat{q}$ units of output using $\hat{L}$ units of labor and $\hat{K}$ units of capital. Abroad, the wage and cost of capital are half as much as at home. If the firm manufactures abroad, will it change the amount of labor and capital it uses to produce $\hat{q}$? What happens to its cost of producing $\hat{q}$?

Answer

1. *Determine whether the change in factor prices affects the slopes of the isoquant or the isocost lines.* The change in input prices does not affect the isoquant, which depends only on technology (the production function). Moreover, cutting all input prices in half does not affect the slope of the isocost lines. The original slope was $-\hat{w}/\hat{r}$, and the new slope is $-(\hat{w}/2)/(\hat{r}/2) = -\hat{w}/\hat{r}$.

2. *Using a rule for cost minimization, determine whether the firm changes its input mix.* A firm minimizes its cost by producing where its isoquant is tangent to the lowest possible isocost line. That is, the firm produces where the slope of its isoquant, *MRTS*, equals the slope of its isocost line, $-w/r$. Because the slopes of the isoquant and the isocost lines are unchanged after input prices are cut in half, the firm continues to produce $\hat{q}$ using the same amount of labor, $\hat{L}$, and capital, $\hat{K}$, as originally.

3. *Calculate the original cost and the new cost and compare them.* The firm's original cost of producing $\hat{q}$ units of output was $\hat{w}\hat{L} + \hat{r}\hat{K} = \hat{C}$. Its new cost of producing the same amount of output is $(\hat{w}/2)\hat{L} + (\hat{r}/2)\hat{K} = \hat{C}/2$. Thus, its cost of producing $\hat{q}$ falls by half when the input prices fall by half. The isocost lines have the same slope as before, but the cost associated with each isocost line is cut in half.

The Long-Run Expansion Path and the Long-Run Cost Function

We now know how a firm determines the cost-minimizing output for any given level of output. By repeating this analysis for different output levels, the firm determines how its cost varies with output.

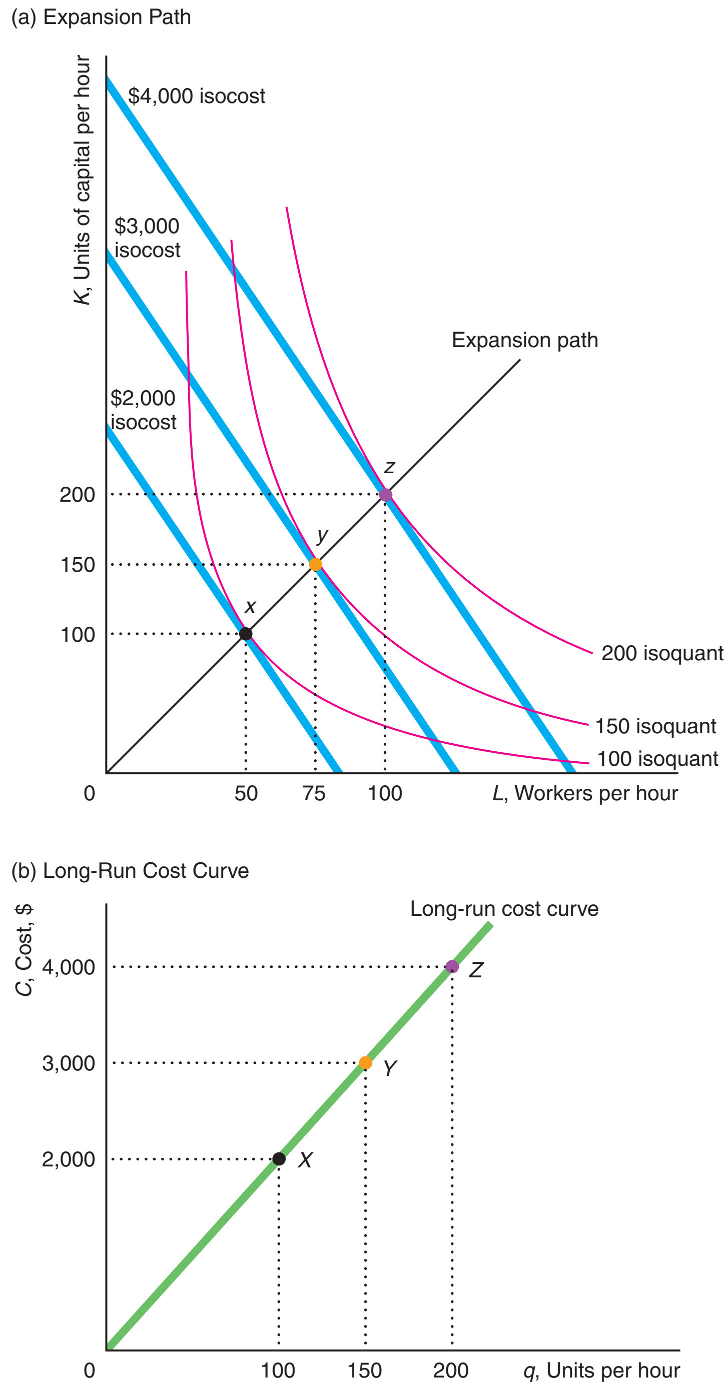
Panel a of Figure 7.7 shows the relationship between the lowest-cost factor combinations and various levels of output for the beer manufacturer when input prices are held constant at $w = \$24$ and $r = \$8$. The curve through the tangency points is the long-run **expansion path**: the cost-minimizing combination of labor and capital for each output level. The lowest-cost way to produce 100 units of output is to use the labor and capital combination x ($L = 50$ and $K = 100$), which lies on the \$2,000 isocost line. Similarly, the lowest-cost way to produce 200 units is to use z, which is on the \$4,000 isocost line. The expansion path goes through x and z.

expansion path
the cost-minimizing combination of labor and capital for each output level

We derived this expansion path for the beer manufacturer in Solved Problem 7.3. According to Equation 7.7, at the cost-minimizing bundles, the firm uses twice as much capital as labor: $K = 2L$. Thus, the expansion path is a straight line through the origin with a slope of 2. To double its output from 100 to 200 units, the firm

Figure 7.7 Expansion Path and Long-Run Cost Curve

(a) The curve through the tangency points between isocost lines and isoquants, such as x, y, and z, is called the expansion path. The points on the expansion path are the cost-minimizing combinations of labor and capital for each output level. The beer manufacturer's expansion path is a straight line. (b) The beer manufacturer's expansion path shows the same relationship between long-run cost and output as the long-run cost curve.



doubles the amount of labor from 50 to 100 workers and doubles the amount of capital from 100 to 200 units. Because both inputs double when output doubles from 100 to 200, cost also doubles.

The beer manufacturer's expansion path contains the same information as its long-run cost function, $C(q)$, which shows the relationship between the cost of production and output. From inspection of the expansion path, to produce q units of output takes $K = q$ units of capital and $L = q/2$ units of labor.[5] Thus, the long-run cost of producing q units of output is

$$C(q) = wL + rK = wq/2 + rq = (w/2 + r)q = (24/2 + 8)q = 20q.$$

That is, the long-run cost function corresponding to this expansion path is $C(q) = 20q$. This cost function is consistent with the expansion path in panel a: $C(100) = \$2,000$ at x on the expansion path, $C(150) = \$3,000$ at y, and $C(200) = \$4,000$ at z.

Panel b plots this long-run cost curve. Points X, Y, and Z on the cost curve correspond to points x, y, and z on the expansion path. For example, the \$2,000 isocost line goes through x, which is the lowest-cost combination of labor and capital that can produce 100 units of output. Similarly, X on the long-run cost curve is at \$2,000 and 100 units of output. Consistent with the expansion path, the cost curve shows that as output doubles, cost doubles.

Solved Problem 7.5

What is the long-run cost function for a fixed-proportions production function (Chapter 6) when it takes one unit of labor and one unit of capital to produce one unit of output? Describe the long-run cost curve.

Answer

Multiply the inputs by their prices, and sum to determine total cost. The long-run cost of producing q units of output is $C(q) = wL + rK = wq + rq = (w + r)q$. Cost rises in proportion to output. The long-run cost curve is a straight line with a slope of $w + r$.

The Shape of Long-Run Cost Curves

The shapes of the average cost and marginal cost curves depend on the shape of the long-run cost curve. To illustrate these relationships, we examine the long-run cost curves of a typical firm that has a U-shaped long-run average cost curve.

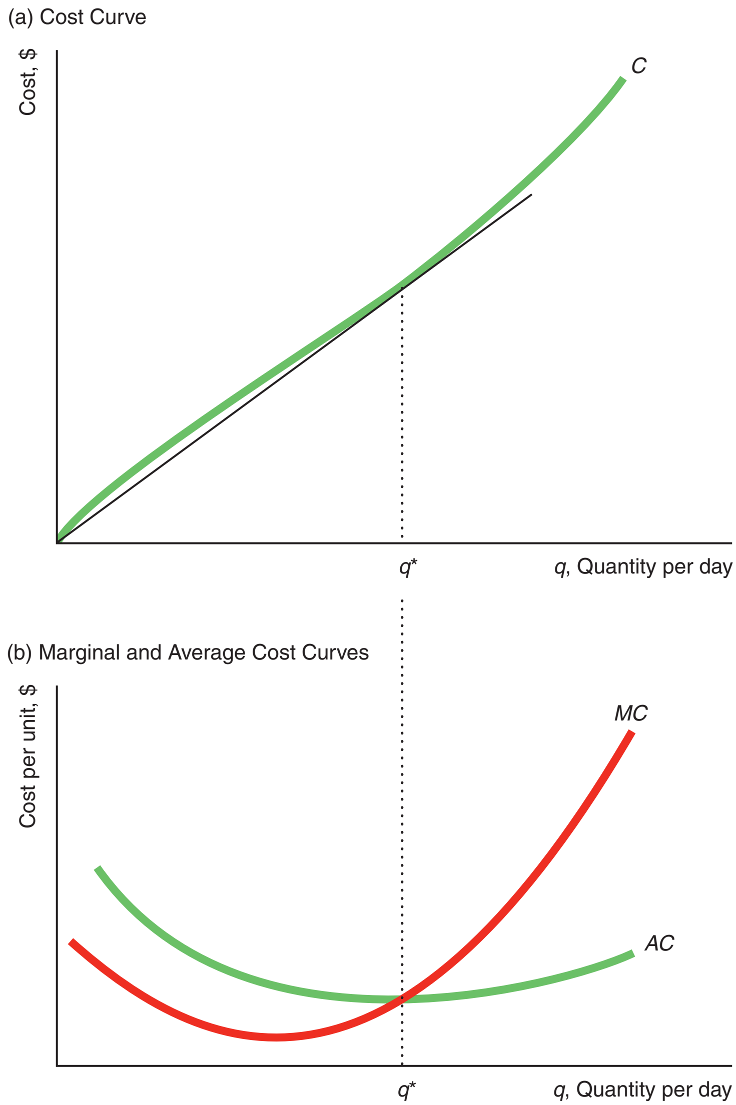
The long-run cost curve in panel a of Figure 7.8 corresponds to the long-run average and marginal cost curves in panel b. Unlike the straight-line long-run cost curves of the beer firm in Figure 7.7 and the firm with fixed-proportions production in Solved Problem 7.5, the long-run cost curve of this firm rises less than in proportion to output at outputs below q^* and then rises more rapidly.

We can apply the same type of analysis that we used to study short-run cost curves to look at the geometric relationship between long-run total, average, and marginal cost curves. A line from the origin is tangent to the long-run cost curve at q^*, where the marginal cost curve crosses the average cost curve, because the slope of that line equals the marginal and average costs at that output. The long-run average cost curve

[5]We can derive this result formally. As the expansion path shows, cost is minimized when $2L = K$. Substituting this expression into the production function, $q = 1.516L^{0.6}K^{0.4}$, we find that $q = 1.516L^{0.6}(2L)^{0.4} = 1.516 \times 2^{0.4}L = 2L$, or $L = q/2$. Thus, $K = 2L = q$.

Figure 7.8 Long-Run Cost Curves

(a) The long-run cost curve rises less rapidly than output at output levels below q^* and more rapidly at higher output levels. (b) As a consequence, the marginal cost and average cost curves are U-shaped. The marginal cost curve crosses the average cost curve at its minimum at q^*.



falls when the long-run marginal cost curve is below it and rises when the long-run marginal cost curve is above it. Thus, the marginal cost curve crosses the average cost curve at the lowest point on the average cost curve.

Why does the average cost curve first fall and then rise, as in panel b? The explanation differs from those given for why short-run average cost curves are U-shaped.

A key reason why the short-run average cost curve is initially downward sloping is that the average fixed cost curve is downward sloping: Spreading the fixed cost over more units of output lowers the average fixed cost per unit. Because a firm has no fixed costs in the long run, the initial downward slope of the long-run average cost curve is not due to fixed costs.

A major reason why the short-run average cost curve slopes upward at higher levels of output is diminishing marginal returns. In the long run, however, all factors can be varied, so diminishing marginal returns do not explain the upward slope of a long-run average cost curve.

Ultimately, as with the short-run curves, the production function's relationship between output and inputs determines the shape of the long-run curves. In the long run, returns to scale play a major role in determining the shape of the average cost curve and other cost curves. As we discussed in Chapter 6, increasing all inputs in proportion may cause output to increase more than in proportion (increasing returns to scale) at low levels of output, in proportion (constant returns to scale) at intermediate levels of output, and less than in proportion (decreasing returns to scale) at high levels of output. If a production function has this returns-to-scale pattern and the prices of inputs are constant, a long-run average cost curve must be U-shaped.

To illustrate the relationship between returns to scale and long-run average cost, we use the returns-to-scale example of Figure 6.5, the data for which are reproduced in Table 7.3. The firm produces one unit of output using a unit each of labor and capital. Given a wage and rental cost of capital of $12 per unit, the total cost and average cost of producing this unit are both $24. Doubling both inputs ($L = K = 2$) causes output to increase more than in proportion to 3 units, reflecting increasing returns to scale. Because cost only doubles and output triples, the average cost falls. A cost function exhibits **economies of scale** if the average cost of production falls as output increases. Thus, returns to scale is a sufficient condition for economies of scale.

Table 7.3 Returns to Scale and Long-Run Costs

Output, Q	Labor, L	Capital, K	Cost, $C = wL + rK$	Average Cost, $AC = C/q$	Returns to Scale
1	1	1	24	24	
3	2	2	48	16	Increasing
6	4	4	96	16	Constant
8	8	8	192	24	Decreasing

$w = r = \$12$ per unit.

economies of scale property of a cost function whereby the average cost of production falls as output increases

Doubling the inputs again ($L = K = 4$) causes output to double ($q = 6$) as well—constant returns to scale—so the average cost remains constant. If the average cost curve is flat, so that an increase in output has no effect on average cost, the cost function has *no economies of scale* in this range.

Doubling the inputs once more ($L = K = 8$) causes only a small increase in output ($q = 8$)—decreasing returns to scale—so average cost increases. A firm suffers from **diseconomies of scale** if average cost rises when output increases.

diseconomies of scale property of a cost function whereby the average cost of production rises when output increases

Average cost curves can have many different shapes. Competitive firms typically have U-shaped average cost curves. Average cost curves in noncompetitive markets may be U-shaped, L-shaped (average cost at first falls rapidly and then levels off as output increases), everywhere downward sloping, or everywhere upward sloping or have other shapes. The shapes of the average cost curves indicate whether the production process has economies or diseconomies of scale.

Table 7.4 summarizes the shapes of average cost curves of firms in various Canadian manufacturing industries (as estimated by Robidoux and Lester, 1992). The table shows that U-shaped average cost curves are the exception rather than the rule in Canadian manufacturing and that nearly one-third of these average cost curves are L-shaped. Some of these apparently L-shaped average cost curves may be part of a U-shaped curve with long, flat bottoms, where we don't observe any firm producing enough to exhibit diseconomies of scale.

Table 7.4 Shape of Average Cost Curves in Canadian Manufacturing

Scale Economies	Share of Manufacturing Industries, %	
Economies of scale: Initially downward-sloping AC	57	
Everywhere downward-sloping AC		18
L-shaped AC (downward-sloping, then flat)		31
U-shaped AC		8
No economies of scale: Flat AC	23	
Diseconomies of scale: Upward-sloping AC	14	

Source: Data from Robidoux and Lester (1992).

Application

3D Printing

Over time, firms increase the size of their factories to take advantage of economies of scale, which reduces costs. However, three-dimensional (3D) printing may reverse this trend by making it equally inexpensive to manufacture one item as it is to manufacture a thousand.

With 3D printing, an employee gives instructions—essentially a blueprint—to the machine, presses *Print*, and the machine builds the object from the ground up, either by depositing material from a nozzle, or by selectively solidifying a thin layer of plastic or metal dust using drops of glue or a tightly focused beam.

Until recently, firms primarily used 3D printers to create prototypes in the aerospace, medical, and automotive industries. Then, they manufactured the final products using conventional techniques. However, costs have fallen to the point where manufacturing using 3D printers is cost-effective in industries that need small numbers of customized parts.

Biomedical and aerospace companies are using 3D printing to reduce costs by engaging in just-in-time manufacturing. It has proven especially useful to biomedical firms in producing customized prostheses. In the aerospace industry, Airbus printed 45,000 to 60,000 parts for its aircraft in 2014 (compared to 20 in 2013). Airbus prints parts that may be 30% to 50% lighter than in the past, which reduces the weight of planes and saves fuel. By one estimate, Boeing had 100,000 printed parts in its aircraft in 2015. Perhaps most striking, Airbus introduced the world's first 3D-printed aircraft, a drone, in 2016.

Some scientists and firms believe that 3D printing will eventually eliminate the need for many factories and may eliminate the manufacturing advantage of low-wage countries. As the cost of 3D printing drops, these machines may be used to produce small, highly customized batches of products as end-users need them.

Estimating Cost Curves Versus Introspection

Economists use statistical methods to estimate a cost function. Sometimes, however, we can infer the shape by casual observation and deductive reasoning.

For example, starting in 1920, the Good Humor company has sent out ice-cream trucks to purvey its products. It seems likely that the company's production process has fixed proportions and constant returns to scale: If it wants to sell more, Good Humor dispatches one more truck and one more driver. Drivers and trucks are almost certainly nonsubstitutable inputs (the isoquants are right angles). If the cost of a driver is w per day, the rental cost is r per day, and q quantity of ice cream is sold in a day, then the cost function is $C = (w + r)q$.

Such deductive reasoning can lead one astray, as I once discovered. A water heater manufacturing firm provided me with many years of data on the inputs it used and the amount of output it produced. I also talked to the company's engineers about the production process and toured the plant (which resembled a scene from Dante's *Inferno*, with staggering noise levels and flames everywhere).

A water heater consists of an outside cylinder of metal, a liner, an electronic control unit, hundreds of tiny parts (screws, washers, etc.), and a couple of rods that slow corrosion. Workers cut out the metal for the cylinder, weld it together, and add the other parts. "Okay," I said to myself, "this production process must be one of fixed proportions because the firm needs one of everything to produce a water heater. How could you substitute a cylinder for an electronic control unit? Or how can you substitute labor for metal?"

I then used statistical techniques to estimate the production and cost functions. Following the usual procedure, however, I did not assume that I knew the exact form of the functions. Rather, I allowed the data to "tell" me the type of production and cost functions. To my surprise, the estimates indicated that the production process was not one of fixed proportions. Rather, the firm could readily substitute between labor and capital.

"Surely I've made a mistake," I said to the plant manager after describing these results. "No," he said, "that's correct. There's a great deal of substitutability between labor and metal."

"How can they be substitutes?"

"Easy," he said. "We can use a lot of labor and waste very little metal by cutting out exactly what we want and being very careful. Or we can use relatively little labor, cut quickly, and waste more metal. When the cost of labor is relatively high, we waste more metal. When the cost of metal is relatively high, we cut more carefully." This practice minimizes the firm's cost.

7.4 Lower Costs in the Long Run

In its long-run planning, a firm chooses a plant size and makes other investments so as to minimize its long-run cost on the basis of how many units it produces. Once it chooses its plant size and equipment, these inputs are fixed in the short run. Thus, the firm's long-run decision determines its short-run cost. Because the firm cannot vary its capital in the short run but can vary it in the long run, short-run cost is at least as high as long-run cost and is higher if the "wrong" level of capital is used in the short run.

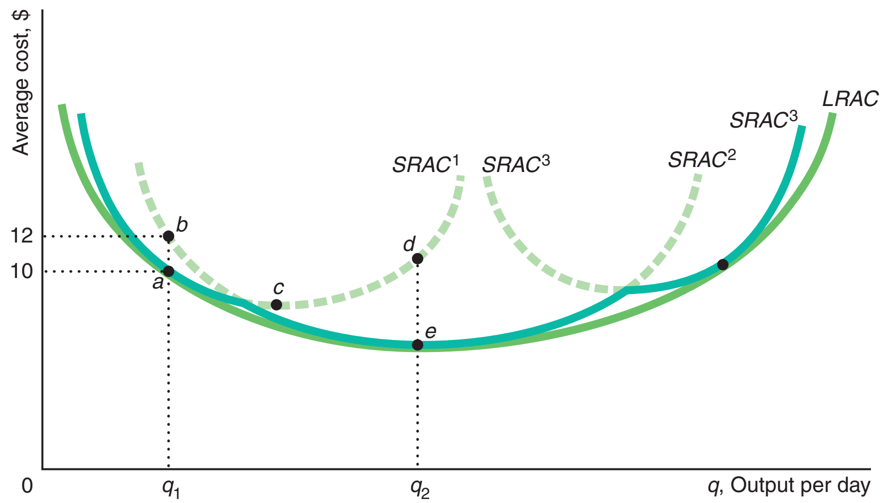
Long-Run Average Cost as the Envelope of Short-Run Average Cost Curves

As a result, the long-run average cost is always equal to or below the short-run average cost. Suppose, initially, that the firm in Figure 7.9 has only three possible plant sizes. The firm's short-run average cost curve is $SRAC^1$ for the smallest possible plant. The average cost of producing q_1 units of output using this plant, point *a* on $SRAC^1$, is \$10. If instead the plant used the next larger plant size, its cost of producing q_1 units of output, point *b* on $SRAC^2$, would be \$12. Thus, if the firm knows that it will produce only q_1 units of output, it minimizes its average cost by using the smaller plant size. If it expects to be producing q_2, its average cost is lower on the $SRAC^2$ curve, point *e*, than on the $SRAC^1$ curve, point *d*.

In the long run, the firm chooses the plant size that minimizes its cost of production, so it picks the plant size that has the lowest average cost for each possible output

Figure 7.9 Long-Run Average Cost as the Envelope of Short-Run Average Cost Curves
MyLab Economics Video

If a firm has only three possible plant sizes, with short-run average costs $SRAC^1$, $SRAC^2$, and $SRAC^3$, the long-run average cost curve is the solid, scalloped portion of the three short-run curves. $LRAC$ is the smooth and U-shaped long-run average cost curve if the firm has many possible short-run average cost curves.



level. At q_1, it opts for the small plant size, whereas at q_2, it uses the medium plant size. Thus, the long-run average cost curve is the blue-green, scalloped section of the three short-run cost curves.

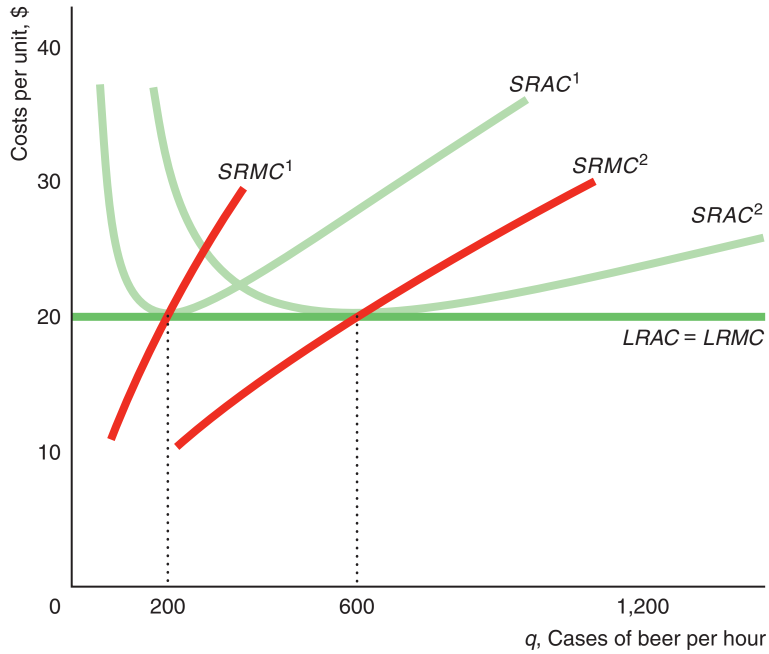
If a firm may choose any plant size it wants, the long-run average curve, *LRAC*, is smooth and U-shaped. The *LRAC* is tangent at one point to each possible short-run average cost curve, making it the *envelope* of the short-run curves. These tangency points are not necessarily the minimum points on the short-run curves. For example, the *LRAC* includes *a* on $SRAC^1$ and not its minimum point, *c*. A small plant operating at minimum average cost cannot produce at as low an average cost as a slightly larger plant that is taking advantage of economies of scale.

Application

A Beer Manufacturer's Long-Run Cost Curves

The graph shows the relationship between short-run and long-run average cost curves for the Japanese beer manufacturer (based on the estimates of Flath, 2011). Because this production function has constant returns to scale, doubling both inputs doubles output, so the long-run average cost, *LRAC*, is constant. If capital is fixed at 200 units, the firm's short-run average cost curve is $SRAC^1$. If the firm produces 200 units of output, its short-run and long-run average costs are equal. At any other output, its short-run cost is higher than its long-run cost.

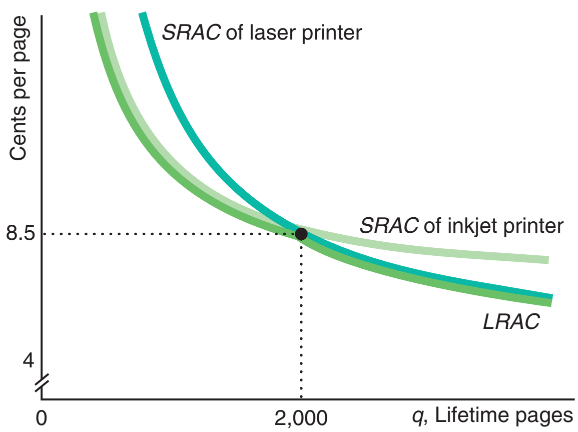
The short-run marginal cost curves, $SRMC^1$ and $SRMC^2$, are upward sloping and equal the corresponding U-shaped short-run average cost curves, $SRAC^1$ and $SRAC^2$, only at their minimum points, \$20. In contrast, because the long-run average cost is horizontal at \$20, the long-run marginal cost curve, *LRMC*, is horizontal at \$20. Thus, the long-run marginal cost curve is *not* the envelope of the short-run marginal cost curves.



Application

Should You Buy an Inkjet or a Laser Printer?

In 2016, you can buy a personal laser printer for \$90 or an inkjet printer for \$30. If you buy the inkjet printer, you immediately save \$60. However, the laser printer costs less per page to operate. The cost of paper and ink or toner is about 4¢ per page for a laser printer compared to about 7¢ per page for an inkjet.



Thus, the average cost per page of operating a laser printer is $\$90/q + 0.04$, where q is the number of pages, while the average cost for an inkjet is $\$30/q + 0.07$. The graph shows the short-run average cost curves for the laser and inkjet printers. The average cost per page is lower with the inkjet printer until q reaches 2,000 pages where the average cost of both is about 8.5¢ per page. For larger quantities, the laser printer is less expensive per page. If the printers last two years and you print 20 or more pages per week, then the laser printer has a lower average cost.

So, should you buy the laser printer? If you buy the printer before you know how much you'll need to print, you may be stuck with the wrong type of printer in the short run. For example, if you buy the inkjet printer thinking that you won't have to print much and then find out that several of your courses require a paper a week, you will incur a higher short-run average cost per page than if you had bought a laser printer.

In the long run, you can adjust the type of printer that you use. Thus, your long-run average cost may be lower than your short-run average cost. The figure shows that the long-run average cost curve corresponds to whichever short-run average cost curve is lowest at a given quantity. Along all of these average cost curves, the more you print, the lower your average cost per page.

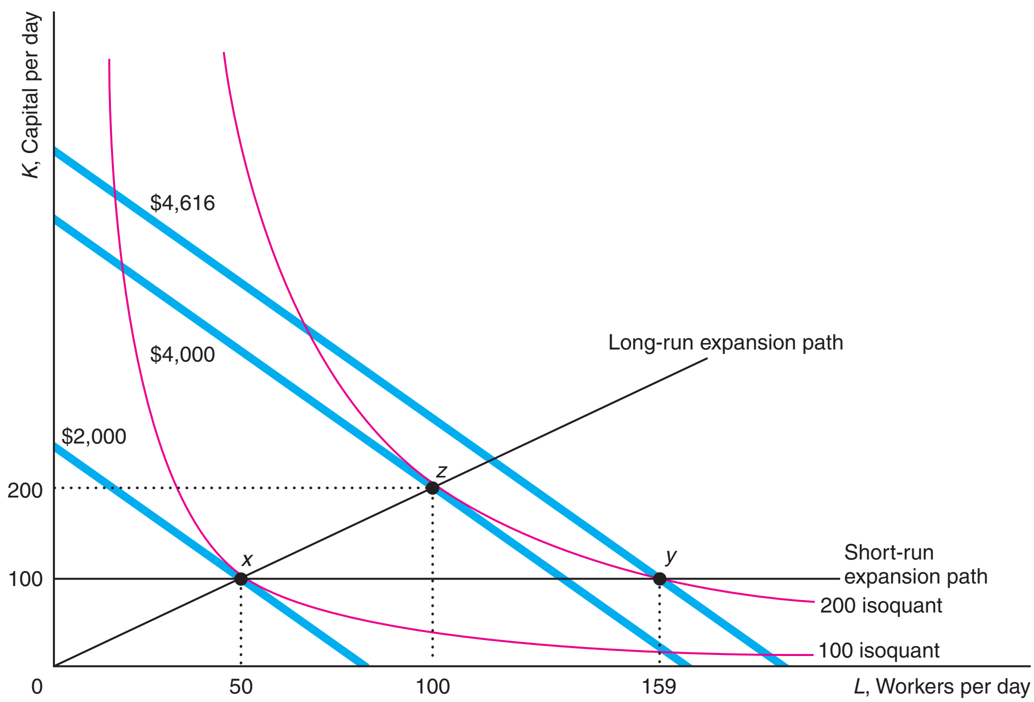
Short-Run and Long-Run Expansion Paths

Long-run cost is lower than short-run cost because the firm has more flexibility in the long run. To show the advantage of flexibility, we can compare the short-run and long-run expansion paths, which correspond to the short-run and long-run cost curves.

The beer manufacturer has greater flexibility in the long run. The tangency of the firm's isoquants and isocost lines determines the long-run expansion path in Figure 7.10. The firm expands output by increasing both its labor and its capital, so

Figure 7.10 Long-Run and Short-Run Expansion Paths MyLab Economics Video

In the long run, the beer manufacturer increases its output by using more of both inputs, so its long-run expansion path is upward sloping. In the short run, the firm cannot vary its capital, so its short-run expansion path is horizontal at the fixed level of output. That is, it increases its output by increasing the amount of labor it uses. Expanding output from 100 to 200 raises the beer firm's long-run cost from \$2,000 to \$4,000 but raises its short-run cost from \$2,000 to \$4,616.



its long-run expansion path is upward sloping. To increase its output from 100 to 200 units (move from x to z), it doubles its capital from 100 to 200 units and its labor from 50 to 100 workers. Its cost increases from $2,000 to $4,000.

In the short run, the firm cannot increase its capital, which is fixed at 100 units. The firm can increase its output only by using more labor, so its short-run expansion path is horizontal at $K = 100$. To expand its output from 100 to 200 units (move from x to y), the firm must increase its labor from 50 to 159 workers, and its cost rises from $2,000 to $4,616. Doubling output increases long-run cost by a factor of 2 and short-run cost by approximately 2.3.

The Learning Curve

What we have to learn to do, we learn by doing. —Aristotle

A firm's average cost may fall over time for three reasons. First, operating at a larger scale in the long run may lower average cost due to increasing returns to scale (IRS). Second, technological progress (see Chapter 6) may increase productivity and thereby lower average cost. Third, a firm may benefit from **learning by doing**: the productive skills and knowledge that workers and managers gain from experience.

learning by doing
the productive skills and knowledge that workers and managers gain from experience

Workers given a new task may perform it slowly the first few times they try, but their speed increases with practice. Managers may learn how to organize production more efficiently, discover which workers to assign to which tasks, and determine where more inventories are needed and where they can be reduced. Engineers may optimize product designs by experimenting with various production methods. For these and other reasons, the average cost of production tends to fall over time, and the effect is particularly strong with new products.

In some firms, learning by doing is a function of the time elapsed since the start of production of a particular product. However, more commonly, learning is a function of *cumulative output*: the total number of units of output produced since the firm started production.

learning curve
the relationship between average costs and cumulative output

The **learning curve** is the relationship between average costs and cumulative output. The learning curve for Intel central processing units (CPUs) in panel a of Figure 7.11 shows that Intel's average cost fell very rapidly with the first few million units of cumulative output, but then dropped relatively slowly with additional units (Salgado, 2008).

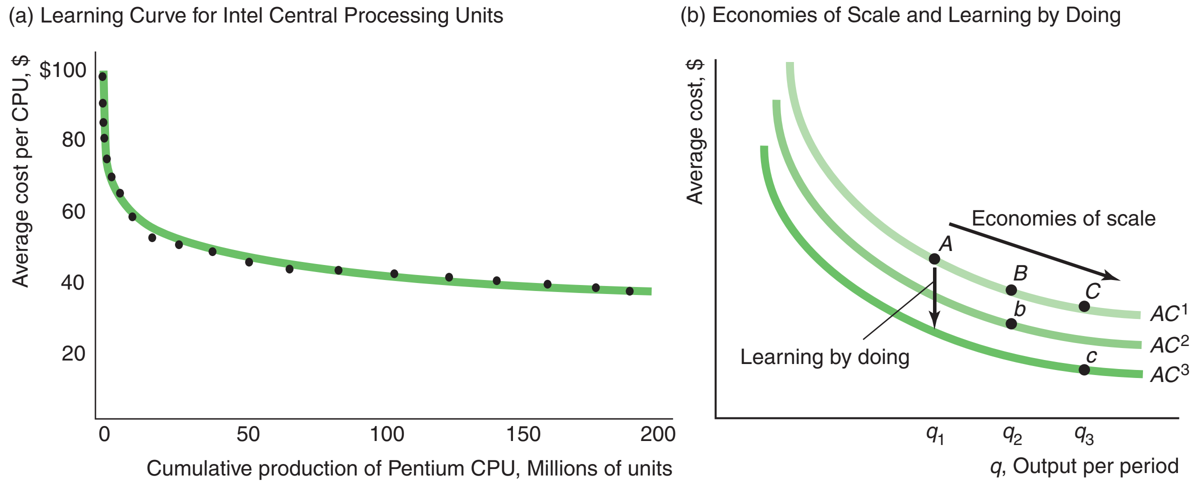
If a firm is operating in the economies of scale section of its average cost curve, expanding output lowers its cost for two reasons. Its average cost falls today because of economies of scale, and for any given level of output, its average cost is lower in the next period due to learning by doing.

In panel b of Figure 7.11, the firm is producing q_1 units of output at point A on average cost curve AC^1 in the first period. We assume that each period is long enough that the firm can vary all factors of production. If the firm expands its output to q_2 in Period 1, its average cost falls in that period to B because of economies of scale. The learning by doing in Period 1 results in a lower average cost, AC^2, in Period 2. If the firm continues to produce q_2 units of output in the next period, its average cost falls to b on AC^2.

If instead of expanding output to q_2 in Period 1, the firm expands to q_3, its average cost is even lower in Period 1 (C on AC^1) due to even more economies of scale. Moreover, its average cost in Period 2, AC^3, is even lower due to the extra experience producing more output in Period 1. If the firm continues to produce q_3 in Period 2, its average cost is c on AC^3. Thus, all else being the same, if learning by doing depends on cumulative output, firms have an incentive to produce more in any one period than they otherwise would to lower their costs in the future.

Figure 7.11 Learning by Doing MyLab Economics Video

(a) As Intel produced more cumulative CPUs, the average cost of production fell (Salgado, 2008). (b) In the short run, extra production reduces a firm's average cost owing to economies of scale: because $q_1 < q_2 < q_3$, A is higher than B, which is higher than C. In the long run, extra production reduces average cost because of learning by doing. To produce q_2 this period costs B on AC^1, but to produce that same output in the next period would cost only b on AC^2. If the firm produces q_3 instead of q_2 in this period, its average cost in the next period is AC^3 instead of AC^2 because of additional learning by doing. Thus, extra output in this period lowers the firm's cost in two ways: It lowers average cost in this period due to economies of scale and lowers average cost for any given output level in the next period due to learning by doing.



7.5 Cost of Producing Multiple Goods

Few firms produce only a single good, but we discuss single-output firms for simplicity. If a firm produces two or more goods, the cost of one good may depend on the output level of the other. For efficiency, a firm produces the two goods together if both require a single input. For example, sheep produce mutton and wool, cattle provide beef and hides, and oil supplies both heating fuel and gasoline. It is less expensive to produce beef and hides together than separately. If the firm produces the goods together, a single steer yields one unit of beef and one hide. If beef and hides are produced separately (throwing away the unused good), the same amount of output requires two steers and more labor.

economies of scope situation in which it is less expensive to produce goods jointly than separately

More generally, we refer to **economies of scope** if it is less expensive to produce goods jointly than separately (Panzar and Willig, 1977, 1981). A measure of the degree of economies of scope (SC) is

$$SC = \frac{C(q_1, 0) + C(0, q_2) - C(q_1, q_2)}{C(q_1, q_2)},$$

where $C(q_1, 0)$ is the cost of producing q_1 units of the first good by itself, $C(0, q_2)$ is the cost of producing q_2 units of the second good, and $C(q_1, q_2)$ is the cost of producing both goods together. If the cost of producing the two goods separately, $C(q_1, 0) + C(0, q_2)$, is the same as producing them together, $C(q_1, q_2)$, then SC is zero. If it is cheaper to produce the goods jointly, SC is positive. If SC is negative, the diseconomies of scope imply that it is cheaper to produce the goods separately.

production possibility frontier
the maximum amount of outputs that can be produced from a fixed amount of input

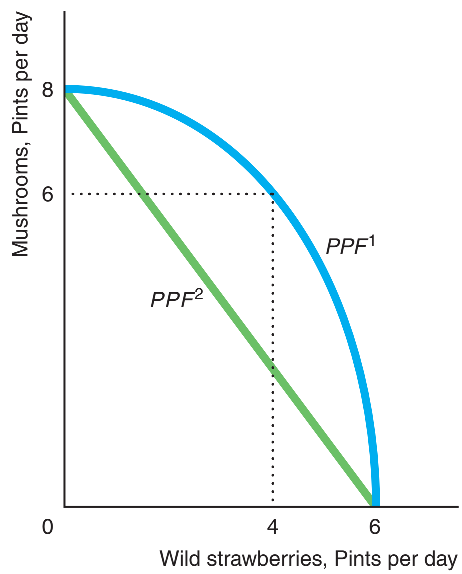
To illustrate this idea, suppose that Laura spends one day collecting mushrooms and wild strawberries in the woods. Her **production possibility frontier**—the maximum amounts of outputs (mushrooms and strawberries) that can be produced from a fixed amount of input (Laura's effort during one day)—is PPF^1 in Figure 7.12. The production possibility frontier summarizes the trade-off Laura faces: She picks fewer mushrooms if she collects more strawberries in a day.

If Laura spends all day collecting only mushrooms, she picks 8 pints; if she spends all day picking strawberries, she collects 6 pints. If she picks some of each, however, she can harvest more total pints: 6 pints of mushrooms and 4 pints of strawberries. The product possibility frontier is concave (the middle of the curve is farther from the origin than it would be if it were a straight line) because of the diminishing marginal returns from collecting only one of the two goods. If she collects only mushrooms, she must walk past wild strawberries without picking them. As a result, she has to walk farther if she collects only mushrooms than if she picks both. Thus, she benefits from economies of scope in jointly collecting mushrooms and strawberries.

If instead the production possibility frontier were a straight line, the cost of producing the two goods jointly would not be lower. Suppose, for example, that mushrooms grow in one section of the woods and strawberries in another section. In that case, Laura can collect only mushrooms without passing any strawberries. That production possibility frontier is a straight line, PPF^2 in Figure 7.12. By allocating her time between the two sections of the woods, Laura can collect any combination of mushrooms and strawberries by spending part of her day in one section of the woods and part in the other.

Figure 7.12 Joint Production

With economies of scope, the production possibility frontier is bowed away from the origin, PPF^1. If instead the production possibility frontier is a straight line, PPF^2, the cost of producing both goods does not fall if the firm produces them together.



Application

Medical Economies of Scope

Empirical studies show that some medical production processes have economies of scope, others have none, and some have diseconomies of scope. Is it cost-effective to separate outpatient and inpatient surgical procedures in a general hospital or should outpatient surgeries be provided separately? Carey et al. (2015) estimated small scope economies ($SC = 0.12$) at the median for-profit hospital.

Gonçalves and Barros (2013) examined whether providing auxiliary clinical services in Portuguese hospitals is cost-effective. They did not find economies of scope between clinical chemistry service and other medical services, so outsourcing that service would not raise costs. However, in medical imaging, computed tomography exhibited scope economies with most other services, suggesting that outsourcing computed tomography would raise the costs of producing those other outputs.

Cohen and Morrison Paul (2011) found large diseconomies of scope in Washington state hospitals to providing both inpatient and outpatient services for drug abuse treatment, so it is cheaper to provide these services separately.

Challenge Solution

Technology Choice at Home Versus Abroad

If a U.S. semiconductor manufacturing firm shifts production from the firm's home plant to a foreign plant, should it use the same mix of inputs as at home? The firm may choose to use a different technology because the firm's cost of labor relative to capital is lower abroad than in the United States.

If the firm's isoquant is smooth, the firm uses a different bundle of inputs abroad than at home, given that the relative factor prices differ as Figure 7.6 shows. However, semiconductor manufacturers may have kinked isoquants. The figure below shows the isoquant that we examined in Chapter 6 in the Application "A Semiconductor Integrated Circuit Isoquant."

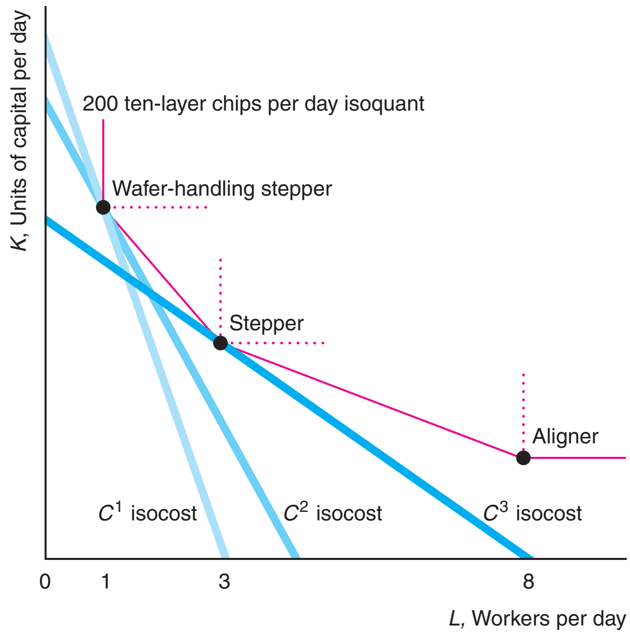
In its U.S. plant, the semiconductor manufacturing firm uses a wafer-handling stepper technology because the C^1 isocost line, which is the lowest isocost line that touches the isoquant, hits the isoquant at that technology.

The firm's cost of both inputs is less abroad than in the United States, and its cost of labor is relatively less than the cost of capital at its foreign plant than at its U.S. plant. The slope of its isocost line is $-w/r$, where w is the wage and r is the rental cost of the manufacturing equipment. The smaller w is relative to r, the less steeply sloped is its isocost curve. Thus, the firm's foreign isocost line is flatter than its domestic C^1 isocost line.

If the firm's isoquant were smooth, the firm would certainly use a different technology at its foreign plant than in its home plant. However, its isoquant has kinks, so a small change in relative input prices does not necessarily lead to a change in production technology. The firm could face either the C^2 or C^3 isocost curves, both of which are flatter than the C^1 isocost. If the firm faces the C^2 isocost line, which is only slightly flatter than the C^1 isocost, the firm still uses the capital-intensive wafer-handling stepper technology in its foreign plant. However, if the firm faces the much flatter C^3 isocost line, which hits the isoquant at the stepper technology, it switches technologies. (If the isocost line were even flatter, it could hit the isoquant at the aligner technology.)

Even if the wage change is small so that the firm's isocost is C^2 and the firm does not switch technologies abroad, the firm's cost will be lower abroad with the same technology because C^2 is less than C^1. However, if the wage is low enough that it can shift to a more labor-intensive technology, its costs will be even lower: C^3 is less than C^2.

Thus, the firm's decision about whether to use a different technology in its foreign plant than in its domestic plant depends on the relative factor prices in the two



locations and whether the firm's isoquant is smooth. If the isoquant is smooth, even a slight difference in relative factor prices will induce the firm to shift along the isoquant and use a different technology with a different capital-labor ratio. However, if the isoquant has kinks, the firm will use a different technology only if the relative factor prices differ substantially.

Summary

From all technologically efficient production processes, a firm chooses the one that is economically efficient. The economically efficient production process is the technologically efficient process for which the cost of producing a given quantity of output is lowest or for which the output at a given cost is highest.

1. **The Nature of Costs.** In making decisions about production, managers need to take into account the opportunity cost of an input, which is the value of the input's best alternative use. For example, if the owner of the company runs it and does not receive a salary, the amount that the owner could have earned elsewhere—the forgone earnings—is the opportunity cost of the owner's time and is relevant in deciding whether the firm should produce or not. A durable good's opportunity cost depends on its current alternative use. If the past expenditure for a durable input is sunk—that is, it cannot be recovered—then that input has no opportunity cost and hence should not influence current production decisions.
2. **Short-Run Costs.** In the short run, the firm can vary the costs of the factors that it can adjust, but the costs of other factors are fixed. The firm's average fixed cost falls as its output rises. If a firm has a short-run average cost curve that is U-shaped, its marginal cost curve is below the average cost curve when average cost is falling and above the average cost curve when average cost is rising, so the marginal cost curve crosses the average cost curve at its minimum.
3. **Long-Run Costs.** In the long run, all factors can be varied, so all costs are variable. As a result, average cost and average variable cost are identical. The firm chooses the combination of inputs that minimize its cost. To produce a given output level, it chooses the lowest isocost line that touches the relevant isoquant, which is tangent to the isoquant. Equivalently, to minimize cost, the firm adjusts inputs until the last dollar spent on any input increases output by as much as the last dollar spent on any other input. If the firm calculates the cost of producing every possible output level given current

input prices, it knows its cost function: Cost is a function of the input prices and the output level. If the firm's average cost falls as output increases, it has economies of scale. If its average cost rises as output increases, the firm suffers from diseconomies of scale.

4. **Lower Costs in the Long Run.** The firm can always do in the long run what it does in the short run, so its long-run cost can never be greater than its short-run cost. Because some factors are fixed in the short run, to produce more output, the firm must greatly increase its use of other factors, which is relatively costly. In the long run, the firm can adjust all factors, a process that keeps its cost down. Long-run cost may also be lower than short-run cost due to technological progress or learning by doing.

5. **Cost of Producing Multiple Goods.** For some goods, it is less expensive to produce two goods jointly rather than separately due to economies of scope. For other goods, diseconomies of scope make it is less expensive to produce the goods separately.

Questions

Select questions are available on MyLab Economics; ** = answer appears at the back of this book;* **A** *= algebra problem;* **C** *= calculus problem.*

1. The Nature of Costs

1.1 The Danish company Maersk Line is the largest shipping container company in the world, with 15–16% of the global container market. According to an article in *ShippingWatch* (Kristiansen, 2012), the trade imbalance between Europe and Asia causes Maersk Line to spend about $1 billion every year shipping 2 million empty containers to Asia to be filled with products for European consumers. Use the concept of opportunity cost to explain the pros and cons of shipping companies coordinating their container transport to use competitors' empty containers situated closer to markets, in comparison to transporting their own thousands of kilometers away.

1.2 You have a ticket to a concert by one of your favorite groups, the Hives, which you bought for $100 and that you cannot resell. However, you can buy a ticket for $30 to attend a talk by Steven Colbert, which is at the same time as the concert. You are willing to pay up to $90 to hear Colbert. Given that you incur no other costs from attending either event, what is your opportunity cost of attending the Hives concert? (*Hint*: See Solved Problem 7.1.)

*1.3 "'There are certain fixed costs when you own a plane,' [Andre] Agassi explained during a break in the action at the Volvo/San Francisco tennis tournament, 'so the more you fly it, the more economic sense it makes. . . . The first flight after I bought it, I took some friends to Palm Springs for lunch.'" (Scott Ostler, "Andre Even Flies Like a Champ," *San Francisco Chronicle*, February 8, 1993, C1.) Discuss whether Agassi's statement is reasonable.

*1.4 A firm purchased copper pipes a few years ago at $10 per pipe and stored them, using them only as the need arose. The firm could sell its remaining pipes in the market at the current price of $9. What is the opportunity cost of each pipe and what is the sunk cost?

1.5 Platinum Pipeline Inc. needs a Caterpillar D6T dozer to install water and sewer lines. How does its fixed cost change if it can rent a dozer rather than buy one? (*Hint*: See Application "The Sharing Economy and the Short Run.")

1.6 Erik is under contract with a textbook publisher to provide interesting and current examples to reinforce key theoretical concepts. He makes €50 per hour for doing this work. Erik also wants to repaint a number of rooms in his house. A professional painter estimated the cost of this work to be €500 (including the paint) and guarantees it will be completed in eight hours. If Erik instead paints the rooms himself, he believes it would take him at least twice that long. What is Erik's opportunity cost of painting the rooms himself?

2. Short-Run Costs

2.1 Many corporations allow CEOs to use the firm's corporate jet for personal travel. The Internal Revenue Service (IRS) requires that the firm report personal use of its corporate jet as taxable executive income, and the Securities and Exchange Commission (SEC) requires that publicly traded corporations report the value of this benefit to shareholders. An important issue is the determination of the value of this benefit. The IRS values a CEO's personal flight at or below the price of a first-class ticket. The SEC values the flight at the "incremental" cost of the flight: the additional costs to the corporation of the flight. The third alternative is the market value of chartering an aircraft. Of the three methods, the first-class ticket is least expensive and the chartered flight is most expensive.

a. What factors (such as fuel) determine the marginal explicit cost to a corporation of an executive's personal flight? Do any of the three valuation methods correctly determine the marginal explicit cost?

b. What is the marginal opportunity cost to the corporation of an executive's personal flight?

2.2 In the twentieth century, department stores and supermarkets largely replaced smaller specialty stores, as consumers found it more efficient to go to one store rather than many stores. Consumers incur a transaction or search cost to shop, primarily the opportunity cost of their time. This transaction cost consists of a fixed cost of traveling to and from the store and a variable cost that rises with the number of different types of items the consumer tries to find on the shelves. By going to a supermarket that carries meat, fruits and vegetables, and other items, consumers can avoid some of the fixed transaction costs of traveling to a separate butcher shop, produce mart, and so forth. Use math or figures to explain why a shopper's average costs are lower when buying at a single supermarket than from many stores. (*Hint*: Define the goods as the items purchased and brought home.)

2.3 Using the information in Table 7.1, construct another table showing how a lump-sum franchise tax of \$30 affects the various average cost curves of the firm.

2.4 In 1796, Gottfried Christoph Härtel, a German music publisher, calculated the cost of printing music using an engraved plate technology and used these estimated cost functions to make production decisions. Härtel figured that the fixed cost of printing a musical page—the cost of engraving the plates—was 900 pfennings. The marginal cost of each additional copy of the page was 5 pfennings (Scherer, 2001).

a. Graph the total cost, average total cost, average variable cost, and marginal cost functions.

b. Is the cost of only one music publisher printing a given composition lower than if several firms print it? Why?

c. Härtel used his data to do the following type of analysis. Suppose he expected to sell exactly 300 copies of a composition at 15 pfennings per page. What was the greatest amount the publisher would be willing to pay the composer per page of the composition?

2.5 The only variable input a janitorial service firm uses to clean offices is workers who are paid a wage, w, of \$8 an hour. Each worker can clean four offices in an hour. Use math to determine the variable cost, the average variable cost, and the marginal cost of cleaning one more office. Draw a diagram like Figure 7.1 to show the variable cost, average variable cost, and marginal cost curves. **A**

2.6 Give the formulas for and plot AFC, MC, AVC, and AC if the cost function is

a. $C = 10 + 10q$.

b. $C = 10 + q^2$.

c. $C = 10 + 10q - 4q^2 + q^3$. **A**

2.7 Gail works in a flower shop, where she produces ten floral arrangements per hour. She is paid \$10 an hour for the first eight hours she works and \$15 an hour for each additional hour she works. If Gail's wage is the firm's only variable cost, what is the firm's cost function? What are its AC, AVC, and MC functions? Draw the AC, AVC, and MC curves. **A**

2.8 A firm has the cost curve $C(q) = 100 + 150q - 46q^2 + 5q^3$. What are the equations of the firm's marginal cost, average variable cost, and average cost curves? What are the minimum values of marginal cost, average variable cost, and average cost, and at what levels of output do they occur? What is the value of marginal cost when average variable cost and average cost are at a minimum? **C**

2.9 A firm has two plants that produce identical output. The cost functions are $C_1 = 10q - 4q^2 + q^3$ and $C_2 = 10q - 2q^2 + q^3$.

a. At what output levels does the average cost curve of each plant reach its minimum?

b. If the firm wants to produce four units of output, how much should it produce in each plant? **C**

*2.10 A firm has the cost curve $C(q) = 25 + q^2$. Show how the firm's average cost varies with output. Is there a minimum average cost and, if so, at what level of output is average cost minimized?

2.11 A Chinese high technology manufacturing firm has a production function of $q = 10L^{0.28}K^{0.64}$ (based on Zhang et al., 2012). It faces factor prices of $w = 10$ and $r = 20$. What are its short-run marginal cost and average variable cost curves? (*Hint*: See Appendix 7B.) **A**

2.12 Suppose in Solved Problem 7.2 that the government charges the firm a franchise tax each year (instead of only once). Describe the effect of this tax on the marginal cost, average variable cost, short-run average cost, and long-run average cost curves.

2.13 In the Application "A Beer Manufacturer's Short-Run Cost Curves," the short-run variable cost function for Japanese beer is $VC = 0.55q^{1.67}$. If the fixed cost is 600 and the firm produces 550 units, determine the C, VC, MC, AFC, and AVC. What happens to these costs if the firm increases its output to 600? **A**

3. Long-Run Costs

3.1 A firm in Saudi Arabia uses capital and labor in its production process. The hourly cost of labor is SR30 and the initial rental rate of capital is SR60 per hour. What is the firm's isocost line? How does the slope of its isocost line change if the rental rate of capital falls to SR40 per hour? If the firm responds to the lower cost of capital by using three more units, how would it have to vary its labor input to keep its costs from changing?

*3.2 You have 60 minutes to take an exam with 2 questions. You want to maximize your score. Toward the end of the exam, the more time you spend on either question, the fewer extra points per minute you get for that question. How should you allocate your time between the two questions? (*Hint*: Think about producing an output of a score on the exam using inputs of time spent on each of the problems. Then use Equation 7.8.)

*3.3 A bottling company uses two inputs to produce bottles of the soft drink Sludge: bottling machines (K) and workers (L). The isoquants have the usual smooth shape. The machine costs \$1,000 per day to run. The workers earn \$200 per day. At the current level of production, the marginal product of the machine is an additional 200 bottles per day, and the marginal product of labor is 50 more bottles per day. Is this firm producing at minimum cost? If it is minimizing cost, explain why. If it is not minimizing cost, explain how the firm should change the ratio of inputs it uses to lower its cost. (*Hint*: Examine the conditions for minimizing cost: Equations 7.5 or 7.8. See Solved Problem 7.3.)

3.4 Use the tangency rule to determine the cost-minimizing bundles of labor and capital for a Japanese synthetic rubber firm's production function $q = L^{0.5}K^{0.5}$ (Flath, 2011) where $w = 10$ and $r = 10$. How does your answer change if $w = 20$ and $r = 10$? (*Hint*: See Solved Problem 7.3.)

3.5 Canada removed all duties and quotas on imports from Bangladesh in 2003. Since that time, Bangladesh has become the second largest source (after India) of Canadian merchandise imports from South Asia. Totalling over C\$1 billion, the vast majority of these imports are garments and textile products. Low labor costs are a key attraction for Canadian retailers. Suppose the production function for a textile firm is $q = K^{0.5}L^{0.5}$. What would be the minimum cost of producing 1,000 units of output if the cost of capital is the same in Country 1 and Country 2 (that is, $r =$ Tk20,000 per month in each), but the cost of labor in Country 1 is half of that in the other country (that is, $w_2 = 2w_1 =$ Tk10,000 per month)? (*Hint*: See Solved Problem 7.3.) **A**

*3.6 A firm is looking to locate its production facilities in one of two countries. The wage rate in Country 1 is 10 and the cost of capital is also 10. The wage rate in Country 2 is 20% higher, but the rental rate of capital is 20% lower. The firm's production function is $q = K^{0.2}L^{0.8}$. How much capital and labor would the firm use to produce 100 units of output in each country, and what would be the total cost of production in each? (*Hint*: See Solved Problem 7.4.) **A**

*3.7 While processes differ somewhat among the three main manufacturers of cricket balls in international matches, all top-quality cricket balls are handcrafted to exacting specifications and standards. Their cork cores are tightly wound with layers of yarn and covered with a leather case with a slightly raised seam that is hand stitched. Once finished, the balls are shaped, stamped, and polished. The labor required is highly skilled and a single craftsman can stitch up to eight cricket balls per day. The yarn used is high-quality linen. Alternatives to leather from cowhide have been tested but found to be of lesser quality. Given this information, what can you conclude about the production function for cricket balls? What is the cost function? (*Hint*: See Solved Problem 7.5.)

3.8 Governments may use, and have used, wage subsidies to reduce unemployment in times of economic recession. For example, from January to August 2009, Mexico provided wage subsidies to certain firms in eligible industries; they reduced work schedules rather than lay workers off in response to the fall in demand (Bruhn, 2016). The percentage of workers laid off had to be no more than one-third of the percentage drop in sales. The amount of the subsidy per worker was essentially Mex\$110 per day to a maximum of Mex\$5,100 every two months, adjusted by the firm's percentage decrease in sales and the percentage of workers retained. If a firm reduces its production in response to a fall in demand and the government covers 30% of the wage cost, what effect will this subsidy have on a cost-minimizing firm's choice of labor and capital to produce at the lower output level? Assume the firm has strictly convex isoquants and that labor and capital are normal inputs (that is, the use of each varies directly with the level of production).

3.9 A firm in the state of Karnataka in India can source one of its factors of production either within the state, F_K, or from the neighboring state of Maharashtra, F_M. Assume the quality and rental price of the input is identical in both states, $r = ₹1{,}000$ per hour, and there is no significant difference in the cost of transporting the input for use in the firm's production facilities. However, the rate of value-added tax in Karnataka is $t_K = 14.5\%$, while it is 2 percentage points lower in Maharashtra, $t_M = 12.5\%$. Use an isocost-isoquant diagram to explain how the cost-minimizing firm would likely respond to this tax rate differential.

3.10 The Bouncing Ball Ping Pong Co. sells table tennis sets that consist of two paddles and one net. What is the firm's long-run expansion path if it incurs no costs other than what it pays for paddles and nets, which it buys at market prices? How does its expansion path depend on the relative prices of paddles and nets? (*Hint*: See Solved Problem 7.5.)

3.11 A firm manufactures boxes of cereal using a fixed-proportion production function: One box and one unit (12 ounces) of cereal produce one box of cereal. What is the expansion path? What is the cost function? (*Hint*: See Solved Problem 7.5.)

*3.12 What is the long-run cost function if the production function is $q = L + K$?

3.13 Suppose that your firm's production function has constant returns to scale. What is the expansion path? (*Hint*: See Solved Problem 7.5.)

3.14 A firm's production function is $q = K^{0.4}L^{0.6}$. What is the firm's cost-minimizing input bundle when $w = 3$ and $r = 4$? Draw a diagram depicting the firm's long-run expansion path. What is the firm's long-run cost function, $C(q)$? **A**

3.16 See the Application "3D Printing." When fully incorporated by firms, how will 3D printing affect the shape of short-run and long-run cost curves?

4. Lower Costs in the Long Run

4.1 A U-shaped long-run average cost curve is the envelope of U-shaped short-run average cost curves. On what part of the short-run curve (downward sloping, flat, or upward sloping) does the short-run curve touch the long-run curve? (*Hint*: Your answer should depend on where on the long-run curve the two curves touch.)

*4.2 If a firm's long-run average cost function is $AC(q) = 2/q$, what happens to average cost as q increases? What is this relationship between average cost and returns to scale called? If the firm also learns by doing, what happens to average cost as q increases? How would learning by doing be reflected in the average cost function? **A**

4.3 In what types of industry would you expect to see substantial learning by doing? Why?

*4.4 Suppose that a firm's learning curve—that is, the relationship between its average cost, AC, and cumulative output, Q—is $AC = 10 + 2Q^a$, where $a \le 0$. If $a = 0$ and cumulative output increases from 1 to 2,000 units, what can you say about average cost and the firm's ability to learn by doing? Would your answer change if $a = -0.5$? **A**

5. Cost of Producing Multiple Goods

5.1 What can you say about Laura's economies of scope if her time is valued at \$10 an hour and her production possibility frontier is PPF^1 in Figure 7.12?

*5.2 A refiner produces heating fuel and gasoline from crude oil in virtually fixed proportions. What can you say about economies of scope for such a firm? What is the sign of its measure of economies of scope, SC?

5.3 According to Haskel and Sadun (2012), the United Kingdom started regulating the size of grocery stores in the early 1990s, and today the average size of a typical U.K. grocery store is roughly half the size of a typical U.S. store and two-thirds the size of a typical French store. What implications would such a restriction on size have on a store's average costs? Discuss in terms of economies of scale and scope.

6. Challenge

*6.1 In the figure in the Challenge Solution, show that there are wage and cost of capital services such that the firm is indifferent between using the wafer-handling stepper technology and the stepper technology. How does this wage/cost of capital ratio compare to those in the C^2 and C^3 isocost lines?

6.2 The harvesting of many crops can be done manually or by using a combination of machines and labor. Suppose that a crew of 15 workers using one machine produces as much output as a crew of 25 workers. If these two technologies are perfect substitutes, what is the equation of an isoquant? If the daily cost of the machine is 1,200 and the daily cost of one worker is 600, what is the equation of the isocost line for the unassisted crew and for the machine-assisted crew? Comparing the two isocost lines, which technology will a cost-minimizing firm use to harvest its crops? Draw a diagram showing the isoquants and isocost curves associated with 2.5 and 5 units of output per day.

Competitive Firms and Markets

8

Competition produces the best in markets and the worst in humans.

Challenge

The Rising Cost of Keeping On Truckin'

Businesses complain constantly about the costs and red tape that government regulations impose on them. U.S. truckers and trucking firms have a particular beef. In recent years, federal and state fees have increased substantially and truckers have had to adhere to many new regulations.

The Federal Motor Carrier Safety Administration (FMCSA), along with state transportation agencies in 41 states, administers interstate trucking licenses through the Unified Carrier Registration Agreement. According to FMCSA's website in 2013, it has 27 types of driver regulations, 16 types of vehicle regulations, 42 types of company regulations, 4 types of hazardous materials regulations, and 14 types of other regulatory guidance. (Of course, they may have added some additional rules while I wrote this last sentence.[1]) A trucker must also maintain minimum insurance coverage, pay registration fees, and follow policies that differ across states before the FMCSA will grant permission to operate. The registration process is so complex and time-consuming that firms pay substantial amounts to brokers who expedite the application process and take care of state licensing requirements.

For a large truck, the annual federal interstate registration fee can exceed $8,000. To operate, truckers and firms must pay for many additional fees and costly regulations. These largely lump-sum costs—which are not related to the number of miles driven—have increased substantially in recent years. During the 2007–2009 financial crisis, many states raised their annual fee from a few hundred to several thousand dollars per truck. Before going into the interstate trucking business, a firm must participate in the New Entrant Safety Assurance Process, which raised the standard of compliance for passing the new entrant safety audit starting in 2009. By 2017, each truck will have to add an electronic onboard recorder, which documents travel time and distance with an annualized cost of between $165 and $832 per truck.

What effect do these new fixed costs have on the trucking industry's market price and quantity? Are individual firms providing more or fewer trucking services? Does the number of firms in the market rise or fall? (As we'll discuss at the end of the chapter, the answer to one of these questions is surprising.)

[1]Indeed, the first time I checked after writing that sentence, I found that they had added a new rule forbidding truckers from texting while driving. (Of course, many of these rules and regulations help protect society and truckers in particular.)

One of the major questions a trucking or other firm faces is "How much should we produce?" To pick a level of output that maximizes its profit, a firm must consider its cost function and how much it can sell at a given price. The amount the firm thinks it can sell depends in turn on the market demand of consumers and its beliefs about how other firms in the market will behave. The behavior of firms depends on the **market structure**: the number of firms in the market, the ease with which firms can enter and leave the market, and the ability of firms to differentiate their products from those of their rivals.

market structure
the number of firms in the market, the ease with which firms can enter and leave the market, and the ability of firms to differentiate their products from those of their rivals

In this chapter, we look at a *competitive market structure*, one in which many firms produce identical products and firms can easily enter and exit the market. Because each firm produces a small share of the total market output and its output is identical to that of other firms, each firm is a *price taker* that cannot raise its price above the market price. If it were to try to do so, this firm would be unable to sell any of its output because consumers would buy the good at a lower price from the other firms in the market. The market price summarizes all a firm needs to know about the demand of consumers *and* the behavior of its rivals. Thus, a competitive firm can ignore the specific behavior of individual rivals in deciding how much to produce.[2]

In this chapter, we examine four main topics

1. **Perfect Competition.** A competitive firm is a price taker, and as such, it faces a horizontal demand curve.
2. **Profit Maximization.** To maximize profit, any firm must make two decisions: how much to produce and whether to produce at all.
3. **Competition in the Short Run.** Variable costs determine a profit-maximizing, competitive firm's supply curve and market supply curve, and with the market demand curve, the competitive equilibrium in the short run.
4. **Competition in the Long Run.** Firm supply, market supply, and competitive equilibrium are different in the long run than in the short run because firms can vary inputs that were fixed in the short run.

8.1 Perfect Competition

Competition is a common market structure that has very desirable properties, so it is useful to compare other market structures to competition. In this section, we describe the properties of competitive firms and markets.

Price Taking

When most people talk about "competitive firms," they mean firms that are rivals for the same customers. By this interpretation, any market with more than one firm is competitive. However, to an economist, only some of these multifirm markets are competitive.

[2]In contrast, each oligopolistic firm must consider the behavior of each of its small number of rivals, as we discuss in Chapter 13.

Economists say that a market is *competitive* if each firm in the market is a *price taker*: a firm that cannot significantly affect the market price for its output or the prices at which it buys inputs. Why would a competitive firm be a price taker? Because it has no choice. The firm is a price taker because it faces a demand curve that is horizontal at the market price. If the demand curve is horizontal at the market price, the firm can sell as much as it wants at that price, so it has no incentive to lower its price. Similarly, the firm cannot increase the price at which it sells by restricting its output because it faces an infinitely elastic demand (see Chapter 3): A small increase in price results in its demand falling to zero.

Why the Firm's Demand Curve Is Horizontal

Perfectly competitive markets have five characteristics that force firms to be price takers:

1. The market consists of many small buyers and sellers.
2. All firms produce identical products.
3. All market participants have full information about price and product characteristics.
4. Transaction costs are negligible.
5. Firms can freely enter and exit the market.

Many Small Buyers and Sellers If the sellers in a market are small and numerous, no single firm can raise or lower the market price. The more firms in a market, the less any one firm's output affects the market output and hence the market price.

For example, the 316,000 U.S. corn farmers are price takers. If a typical grower were to drop out of the market, market supply would fall by only $1/316{,}000 \approx 0.00032\%$, so the market price would not be noticeably affected. Each soybean farm can sell as much output as it can produce at the prevailing market equilibrium price, so each farm faces a demand curve that is a horizontal line at the market price.

Similarly, perfect competition requires that buyers be price takers as well. For example, if firms sell to only a single buyer—such as producers of weapons that are allowed to sell to only the government—then the buyer can set the price and the market is not perfectly competitive.

Identical Products Firms in a perfectly competitive market sell *identical* or *homogeneous* products. Consumers do not ask which farm grew a Granny Smith apple because they view all Granny Smith apples as essentially identical. If the products of all firms are identical, it is difficult for a single firm to raise its price above the going price charged by other firms.

In contrast, in the automobile market—which is not perfectly competitive—the characteristics of a BMW 5 Series and a Honda Civic differ substantially. These products are *differentiated* or *heterogeneous*. Competition from Civics would not be very effective in preventing BMW from raising its price.

Full Information If buyers know that different firms are producing identical products and they know the prices charged by all firms, no single firm can unilaterally raise its price above the market equilibrium price. If it tried to do so, consumers would buy the identical product from another firm. However, if consumers are unaware that products are identical or they don't know the prices charged by other firms, a single firm may be able to raise its price and still make sales.

Negligible Transaction Costs Perfectly competitive markets have very low transaction costs. Buyers and sellers can easily find each other and can trade without hiring lawyers to write contracts.[3] If transaction costs are low, it is easy for a customer to buy from a rival firm if the customer's usual supplier raises its price.

In contrast, if transaction costs are high, customers might absorb a price increase from a traditional supplier. For example, because some consumers prefer to buy milk at a local convenience store rather than travel several miles to a supermarket, the convenience store can charge slightly more than the supermarket without losing all its customers.

In some perfectly competitive markets, many buyers and sellers are brought together in a single room, so transaction costs are virtually zero. For example, transaction costs are very low at FloraHolland's daily flower auctions in the Netherlands, which attract 7,000 suppliers and 4,500 buyers from around the world. It has 125,000 auction transactions every day, with 12 billion cut flowers and 1.3 billion plants trading in a year.

Free Entry and Exit The ability of firms to enter and exit a market freely leads to a large number of firms in a market and promotes price taking. Suppose a firm can raise its price and increase its profit. If other firms can quickly and easily enter the market, the higher profit encourages entry by new firms until the price falls to the original level. Free exit is also important: If firms can freely enter a market but cannot exit easily if prices decline, they are reluctant to enter the market in response to a possibly temporary profit opportunity.[4] More generally, we assume perfect mobility of resources, which allows firms to alter their scale of production and to enter and exit an industry.

Perfect Competition in the Chicago Mercantile Exchange The Chicago Mercantile Exchange, where buyers and sellers can trade wheat and other commodities, exhibits the characteristics of perfect competition including thousands of buyers and sellers who are price takers. Anyone can be a buyer or a seller. Indeed, a trader might buy wheat in the morning and sell it in the afternoon. They trade virtually *identical products*. Buyers and sellers have *full information* about products and prices, which are posted for everyone to see. Market participants waste no time finding someone who wants to trade and they can easily place buy or sell orders in person, over the telephone, or electronically without paperwork, so *transaction costs are negligible*. Finally, *buyers and sellers can easily enter this market and trade wheat*. These characteristics lead to an abundance of buyers and sellers and to price-taking behavior by these market participants.

Deviations from Perfect Competition

Many markets possess some but not all the characteristics of perfect competition. Such markets are still highly competitive so that buyers and sellers are, for all practical purposes, price takers. For example, a government may limit entry into a market, but if the market has many buyers and sellers, they may still be price takers. Many cities use zoning laws to limit the number of certain types of stores or motels, yet such cities still have a large number of these firms. Other cities impose moderately large transaction costs on entrants by requiring them to buy licenses, post bonds, and deal with a slow-moving city bureaucracy, yet a significant number of firms

---

[3]Average number of hours per week that an American and a Chinese person, respectively, spend shopping: 4, 10.—*Harper's Index*, 2008.

[4]For example, many governments require that firms give workers six months' warning before they exit a market or pay them a severance fee.

enter the market. Similarly, even if only some customers have full information, that may be sufficient to prevent firms from deviating significantly from price taking. For example, tourists do not know the prices at various stores, but locals do and they use their knowledge to prevent one store from charging unusually high prices.

Economists use the terms *competition* and *competitive* more restrictively than do others. To an economist, a competitive firm is a price taker. In contrast, when most people talk about competitive firms, they mean that firms are rivals for the same customers. Even in a market with only a few firms, the firms compete for the same customers so they are competitive in this broader sense. From now on, we will use the terms *competition* and *competitive* to refer to all markets in which no single buyer or seller can significantly affect the market price—they are price takers—even if the market is not perfectly competitive.

Derivation of a Competitive Firm's Demand Curve

Are the demand curves faced by individual competitive firms actually flat? To answer this question, we use a modified supply-and-demand diagram to derive the demand curve for an individual firm.

residual demand curve
the market demand that is not met by other sellers at any given price

An individual firm faces a **residual demand curve**: the market demand that is not met by other sellers at any given price. The firm's residual demand function, $D^r(p)$, shows the quantity demanded from the firm at price p. A firm sells only to people who have not already purchased the good from another seller. We can determine how much demand is left for a particular firm at each possible price using the market demand curve and the supply curve for all *other* firms in the market. The quantity the market demands is a function of the price: $Q = D(p)$. The supply curve of the other firms is $S^o(p)$. The residual demand function equals the market demand function, $D(p)$, minus the supply function of all other firms:

$$D^r(p) = D(p) - S^o(p). \tag{8.1}$$

At prices so high that the amount supplied by other firms, $S^o(p)$, is greater than the quantity demanded by the market, $D(p)$, the residual quantity demanded, $D^r(p)$, is zero.

In Figure 8.1 we derive the residual demand for a Canadian manufacturing firm that produces metal chairs. Panel b shows the market demand curve, D, and the supply of all but one manufacturing firm, S^o.[5] At $p = \$66$ per chair, the supply of other firms, 500 units (where one unit is 1,000 metal chairs) per year, exactly equals the market demand (panel b), so the residual quantity demanded of the remaining firm (panel a) is zero.

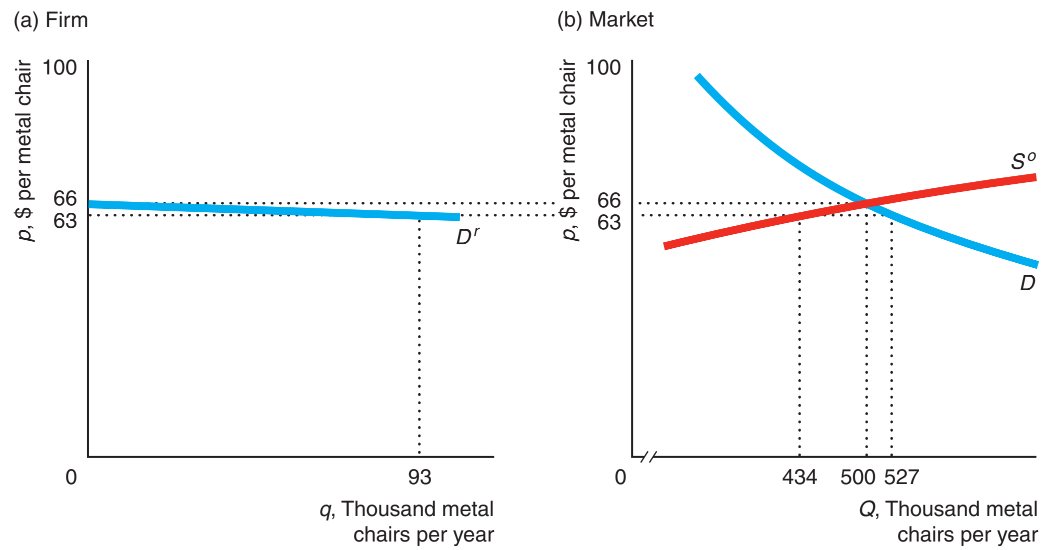
At prices below \$66, the other chair manufacturers are not willing to supply as much as the market demands. At $p = \$63$, for example, the market demand is 527 units, but other firms want to supply only 434 units. As a result, the residual quantity demanded from the individual firm at $p = \$63$ is 93 ($= 527 - 434$) units. Thus, the residual demand curve at any given price is the horizontal difference between the market demand curve and the supply curve of the other firms.

The residual demand curve the firm faces in panel a is much flatter than the market demand curve in panel b. As a result, the elasticity of the residual demand curve is much higher than the market elasticity.

[5]The figure uses constant elasticity demand and supply curves. The elasticity of supply, $\eta = 3.1$, is based on the estimated cost function from Robidoux and Lester (1988) for Canadian office furniture manufacturers. I estimate that the elasticity of demand is $= -1.1$ using data from Statistics Canada, *Office Furniture Manufacturers*.

Figure 8.1 Residual Demand Curve

The residual demand curve, $D^r(p)$, that a single office furniture manufacturing firm faces is the market demand, $D(p)$, minus the supply of the other firms in the market, $S^o(p)$. The residual demand curve is much flatter than the market demand curve.



If the market has n identical firms, the elasticity of demand, $_i$, facing Firm i is

$$_i = n \quad - (n-1)\eta_o, \tag{8.2}$$

where is the market elasticity of demand (a negative number), η_o is the elasticity of supply of the other firms (typically a positive number), and $n - 1$ is the number of other firms (see Appendix 8A for the derivation).

As Equation 8.2 shows, a firm's residual demand curve is more elastic with more firms, n, in the market, with more elastic market demand, , and with larger elasticity of supply of the other firms, η_o. The residual demand elasticity, $_i$, must be at least as elastic as n if the supply curve slopes up so that the second term makes the estimate more elastic. Thus, using n as an approximation is conservative. For example, even though the estimated market elasticity of demand for soybeans is very inelastic at about $= -0.2$, because $n = 107{,}000$, the residual demand facing a single soybean farm must be at least n $= 107{,}000 \times (-0.2) = -21{,}400$, which is extremely elastic.

Solved Problem 8.1

The Canadian metal chair manufacturing market has $n = 78$ firms. The estimated elasticity of supply is $\eta_o = 3.1$, and the estimated elasticity of demand is $= -1.1$. Assuming that the firms are identical, calculate the elasticity of demand facing a single firm. Is its residual demand curve highly elastic?

Answer

1. *Use Equation 8.2 and the estimated elasticities to calculate the residual demand elasticity facing a firm.* Substituting the elasticities into Equation 8.2, we find that

$$\begin{aligned} {}_i &= n \quad - (n - 1)\eta_o \\ &= [78 \times (-1.1)] - [77 \times 3.1] \\ &= -85.8 - 238.7 = -324.5. \end{aligned}$$

   That is, a typical firm faces a residual demand elasticity of −324.5.

2. *Discuss whether this elasticity is high.* The estimated $_i$ is nearly 300 times the market elasticity of −1.1. If a firm raises its price by one-tenth of a percent, the quantity it can sell falls by nearly one-third. Therefore, the competitive model assumption that this firm faces a horizontal demand curve with an infinite price elasticity is not much of an exaggeration.

Why We Study Perfect Competition

Perfectly competitive markets are important for two reasons:

First, many markets can be reasonably described as competitive. Many agricultural and other commodity markets, stock exchanges, retail and wholesale, building construction, and other types of markets have many or all of the properties of a perfectly competitive market. The competitive supply-and-demand model works well enough in these markets that it accurately predicts the effects of changes in taxes, costs, incomes, and other factors on market equilibrium.

Second, a perfectly competitive market has many desirable properties (see Chapter 9). Economists use the perfectly competitive model as the ideal against which real-world markets are compared. Throughout the rest of this book, we consider that society as a whole is worse off if the properties of the perfectly competitive market fail to hold. From this point on, for brevity, we use the phrase *competitive market* to mean a *perfectly competitive market* unless we explicitly note an imperfection.

8.2 Profit Maximization

"Too caustic?" To hell with the cost. If it's a good picture, we'll make it.
—Samuel Goldwyn

Economists usually assume that *all* firms—not just competitive firms—want to maximize their profits. One reason is that many businesspeople say that their objective is to maximize profits. A second reason is that a firm—especially a competitive firm—that does not maximize profit is likely to lose money and be driven out of business.

In this section, we examine how any type of firm—not just a competitive firm—maximizes its profit. We then examine how a competitive firm in particular maximizes profit.

Profit

A firm's *profit*, π, is the difference between a firm's revenues, R, and its cost, C:

$$\pi = R - C.$$

If profit is negative, $\pi < 0$, the firm makes a *loss*.

Measuring a firm's revenue is straightforward: revenue is price times quantity. Measuring cost is more challenging. For an economist, the correct measure of cost is the *opportunity cost* or *economic cost*: the value of the best alternative use of any input the firm employs. As discussed in Chapter 7, the full opportunity cost of inputs used might exceed the explicit or out-of-pocket costs recorded in financial accounting statements. This distinction is important because a firm may make a serious mistake if it incorrectly measures profit by ignoring some relevant opportunity costs.

economic profit revenue minus opportunity (economic) cost

We always refer to profit or **economic profit** as revenue minus opportunity (economic) cost. For tax or other reasons, *business profit* may differ. For example, if a firm uses only explicit cost, then its reported profit may be larger than its economic profit. A couple of examples illustrate the difference in the two profit measures and the importance of this distinction in dispelling a misconception:

Common Confusion: It pays to run your own firm if you are making a business profit.

That conclusion may not follow because business profit ignores opportunity cost (unlike economic profit).

Suppose that you start your own firm.[6] You have to pay explicit costs such as workers' wages and the price of materials. Like many owners, you do not pay yourself a salary. Instead, you take home a business profit of $20,000 per year. Economists (well-known spoilsports) argue that your profit is less than $20,000. Economic profit equals your business profit minus any additional opportunity cost. Suppose that instead of running your own business, you could have earned $25,000 a year working for someone else. The opportunity cost of your time working for your business is $25,000—your forgone salary. So even though your firm made a business profit of $20,000, your economic loss (negative economic profit) is $5,000. Put another way, the price of being your own boss is $5,000.

By looking at only the business profit and ignoring opportunity cost, you conclude that running your business is profitable. However, if you consider economic profit, you realize that working for others maximizes your income.

Similarly, when a firm decides whether to invest in a new venture, it must consider the next best alternative use of its funds. A firm that is considering setting up a new branch in Tucson must consider all the alternatives—placing the branch in Santa Fe, putting the money that the branch would cost in the bank and earning interest, and so on. If the best alternative use of the money is to put it in the bank and earn $10,000 per year in interest, the firm should build the new branch in Tucson only if it expects to make $10,000 or more per year in business profits. That is, the firm should create a Tucson branch only if the economic profit from the new branch is zero or positive. If the economic profit is zero, then the firm is earning the same return on its investment as it would from putting the money in its next best alternative, the bank. From this point on, when we use the term *profit*, we mean economic profit unless we specifically refer to business profit.

[6]Michael Dell started a mail-order computer company while he was in college. Today, his company is one of the world's largest personal computer companies. *Forbes* estimated Mr. Dell's wealth at $19.7 billion in 2015.

Two Decisions for Maximizing Profit

Any firm (not just a competitive firm) makes two successive decisions to maximize its profit. Because both revenue and cost vary with output, a firm's profit varies with its output level. The firm's profit function is

$$\pi(q) = R(q) - C(q),$$

where q is the number of units it produces, $R(q)$ is its revenue function, and $C(q)$ is its cost function. A firm chooses how much output to produce to maximize its profit by answering two questions:

- **Output decision.** If the firm produces a positive amount of output, what output level, q^*, maximizes its profit or minimizes its loss?
- **Shutdown decision.** Is it more profitable to produce q^* or to shut down and produce no output?

The profit curve in Figure 8.2 illustrates these two basic decisions. This firm makes losses at very low and very high output levels and positive profits at moderate output levels. The profit curve first rises and then falls, reaching a maximum profit of π^* when its output is q^*. Because the firm makes a positive profit at that output, it chooses to produce q^* units of output.

Output Decision Rules A firm can use one of three equivalent rules to choose how much output to produce. All types of firms maximize profit using the same rules. The first output rule is the most straightforward:

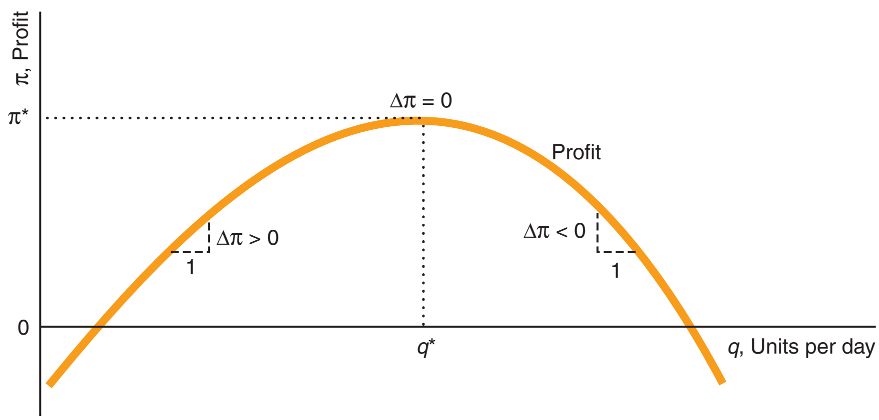
Output Rule 1: The firm sets its output where its profit is maximized.

The profit curve in Figure 8.2 is maximized at π^* when output is q^*. If the firm knows its entire profit curve, it can immediately set its output to maximize its profit.

Even if the firm does not know the exact shape of its profit curve, it may be able to find the maximum by experimenting. The firm slightly increases its output. If profit increases, the firm increases the output more. The firm keeps increasing output until profit does not change. At that output, the firm is at the peak of the profit curve. If profit falls when the firm first increases its output, the firm tries decreasing its output. It keeps decreasing its output until it reaches the peak of the profit curve.

Figure 8.2 Maximizing Profit MyLab Economics Video

By setting its output at q^*, the firm maximizes its profit at π^*.



marginal profit
the change in profit a firm gets from selling one more unit of output

What the firm is doing is experimentally determining the slope of the profit curve. The slope of the profit curve is the firm's **marginal profit**: the change in the profit the firm gets from selling one more unit of output, $\Delta\pi/\Delta q$.[7] In the figure, the marginal profit or slope is positive when output is less than q^*, zero when output is q^*, and negative when output is greater than q^*. Thus, a second, equivalent output rule follows:

Output Rule 2: A firm sets its output where its marginal profit is zero.

marginal revenue (*MR*)
the change in revenue a firm gets from selling one more unit of output

A third way to express a profit-maximizing output rule is in terms of cost and revenue. The marginal profit depends on a firm's *marginal cost* and *marginal revenue*. A firm's *marginal cost* (MC) is the amount by which a firm's cost changes if it produces one more unit of output (Chapter 7): $MC = \Delta C/\Delta q$, where ΔC is the change in cost when output changes by Δq. Similarly, a firm's **marginal revenue,** MR, is the change in revenue it gets from selling one more unit of output: $\Delta R/\Delta q$, where ΔR is the change in revenue.[8] If a firm that was selling q units of output sells one more unit of output, the extra revenue, $MR(q)$, raises its profit, but the extra cost, $MC(q)$, lowers its profit. The change in the firm's profit from producing one more unit is the difference between the marginal revenue and the marginal cost:[9]

$$\text{Marginal profit}(q) = MR(q) - MC(q).$$

Does it pay for a firm to produce one more unit of output? If the marginal revenue from this last unit of output exceeds its marginal cost, $MR(q) > MC(q)$, the firm's marginal profit is positive, $MR(q) - MC(q) > 0$, so it pays to increase output. The firm keeps increasing its output until its marginal profit $= MR(q) - MC(q) = 0$. At that quantity, its marginal revenue equals its marginal cost: $MR(q) = MC(q)$. If the firm produces more output where its marginal cost exceeds its marginal revenue, $MR(q) < MC(q)$, the extra output reduces the firm's profit. Thus we arrive at a third, equivalent output rule (Appendix 8B).

Output Rule 3: A firm sets its output where its marginal revenue equals its marginal cost,

$$MR(q) = MC(q).$$

Shutdown Decision Rule The firm chooses to produce q^* if it can make a profit. But even if the firm maximizes its profit at q^*, should it produce output if doing so makes a loss? "Common sense" suggest that it should not.

Common Confusion: A firm should shut down if it is making a loss.

---

[7]The calculus definition of the marginal profit is the derivative of the profit function, $\pi(q)$, with respect to quantity, $d\pi(q)/dq$.

[8]The calculus definition of the marginal revenue is the derivative of the revenue function with respect to quantity:

$$MR(q) = dR(q)/dq.$$

[9]Because profit is $\pi(q) = R(q) - C(q)$, marginal profit is the difference between marginal revenue and marginal cost:

$$\frac{d\pi(q)}{dq} = \frac{dR(q)}{dq} - \frac{dC(q)}{dq} = MR - MC.$$

This intuition holds if the firm is making a loss in the long run, but it may be wrong in the short run. The general rule, which holds for all types of firms in both the short and long run, is

Shutdown Rule 1: The firm shuts down only if it can reduce its loss by doing so.

In the short run, the firm has variable costs, such as from labor and materials, and fixed costs, such as plant and equipment. If the fixed cost is *sunk* (Chapter 7), the firm cannot avoid this expense by stopping operations. By shutting down, the firm stops receiving revenue and stops paying the avoidable costs, but it is still stuck with any its fixed cost. Consequently, it pays for the firm to shut down only if its revenue is less than its avoidable cost.

Suppose that the firm's weekly revenue is $R = \$2{,}000$, its variable cost is $VC = \$1{,}000$, and its sunk fixed cost is $F = \$3{,}000$, which is the price it paid for a machine that it cannot resell or use for any other purpose. This firm is making a short-run loss:

$$\pi = R - VC - F = \$2{,}000 - \$1{,}000 - \$3{,}000 = -\$2{,}000.$$

If the firm shuts down, it loses its sunk fixed cost, \$3,000, so it is better off operating. Its revenue more than covers its avoidable variable cost and offsets some of the fixed cost.

However, if its revenue is only \$500, its loss is \$3,500, which is greater than the loss from the sunk fixed cost alone of \$3,000. Because its revenue is less than its avoidable variable cost, the firm reduces its loss by shutting down.

In conclusion, the firm compares its revenue to its variable cost only when deciding whether to stop operating. Because the fixed cost is sunk, the firm pays this cost whether it shuts down or not. The sunk fixed cost is irrelevant to the shutdown decision.

We usually assume that fixed cost is *sunk* in the short run. However, if a firm can sell its capital for as much as it paid, its fixed cost is *avoidable* (Chapter 7) so that the firm should consider it when deciding whether to shut down. A firm with a fully avoidable fixed cost always shuts down if it makes a short-run loss. If a firm buys a specialized piece of machinery for \$1,000 that can be used only for its business but can be sold for scrap metal for \$100, then \$100 of the fixed cost is avoidable and \$900 is sunk. Only the avoidable portion of a fixed cost is relevant for the shutdown decision.

In the long run, all costs are avoidable because the firm can eliminate them all by shutting down. Thus, where the firm can avoid all losses by not operating, it pays to shut down if the firm faces any loss at all. As a result, we can restate the shutdown rule as:

Shutdown Rule 2: The firm shuts down only if its revenue is less than its avoidable cost.

Both versions of the shutdown rule hold for all types of firms in both the short run and the long run.

8.3 Competition in the Short Run

Having considered how firms maximize profit in general, we now examine the profit-maximizing behavior of competitive firms, derive their supply curves, and then determine the competitive equilibrium in the short run. In the next section, we examine competitive firms and competitive markets in the long run.

The *short run* is a period short enough that at least one input cannot be varied (Chapter 6). Because a firm cannot quickly build a new plant or make other large capital expenditures, a new firm cannot enter a market in the short run. Similarly, a firm cannot fully exit in the short run. It can choose not to produce—to shut down—but it is stuck with some fixed inputs such as a plant or other capital that it cannot quickly sell or assign to other uses. In the long run, all inputs can be varied so firms can enter and fully exit the industry.

We treat the short run and the long run separately for two reasons. First, profit-maximizing firms may choose to operate at a loss in the short run, whereas they do not do so in the long run. Second, a firm's long-run supply curve typically differs from its short-run supply curve.

In both the short run and the long run, a competitive firm, like other firms, determines the output at which it maximizes its profit (or minimizes its loss) and then decides whether to produce or to shut down.

Short-Run Output Decision

We've already seen that *any* firm maximizes its profit at the output where its marginal profit is zero or, equivalently, where its marginal cost equals its marginal revenue. Because it faces a horizontal demand curve, a competitive firm can sell as many units of output as it wants at the market price, p. Thus, a competitive firm's revenue, $R = pq$, increases by p if it sells one more unit of output, so its marginal revenue is p.[10] For example, if the firm faces a market price of \$2 per unit, its revenue is \$10 if it sells 5 units and \$12 if it sells 6 units, so its marginal revenue for the sixth unit is \$2 = \$12 − \$10 (the market price). Because a competitive firm's marginal revenue equals the market price, *a profit-maximizing competitive firm produces the amount of output at which its marginal cost equals the market price*:

$$MC(q) = p. \tag{8.3}$$

To illustrate how a competitive firm maximizes its profit, we examine a representative firm in the highly competitive Canadian lime manufacturing industry. Lime is a nonmetallic mineral used in mortars, plasters, cements, bleaching powders, steel, paper, glass, and other products. The lime plant's estimated average cost curve, AC, first falls and then rises in panel a of Figure 8.3.[11] As always, the marginal cost curve, MC, intersects the average cost curve at its minimum point.

If the market price of lime is $p = \$8$ per metric ton, the competitive firm faces a horizontal demand curve (marginal revenue curve) at \$8. The MC curve crosses the firm's demand curve (or price or marginal revenue curve) at point e, where the firm's output is 284 units (where a unit is a thousand metric tons).

At a market price of \$8, the competitive firm maximizes its profit by producing 284 units. If the firm produced fewer than 284 units, the market price would be above its marginal cost. The firm could increase its profit by expanding output because the firm earns more on the next ton, $p = \$8$, than it costs to produce it, $MC < \$8$. If the firm were to produce more than 284 units, the market price would be below its marginal cost, $MC > \$8$, and the firm could increase its profit by reducing its output. Thus, the competitive firm maximizes its profit by producing that output at which its marginal cost equals its marginal revenue, which is the market price.

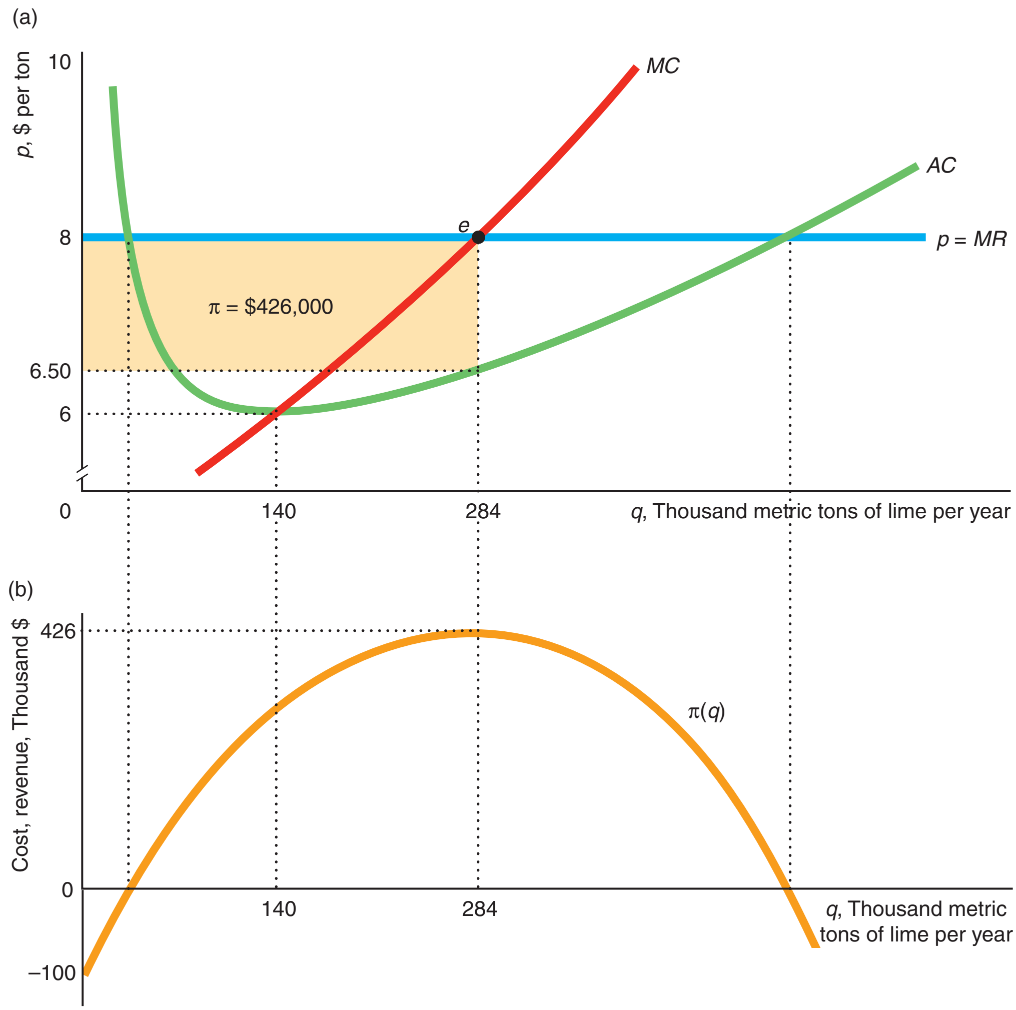
At that 284 units, the firm's profit is $\pi = \$426{,}000$, which is the shaded rectangle in panel a. The length of the rectangle is the number of units sold, $q = 284{,}000$ (or 284 units). The height of the rectangle is the firm's average profit per unit. Because the firm's profit is its revenue, $R(q) = pq$, minus its cost, $\pi(q) = R(q) - C(q)$,

[10]We can derive this result using calculus. Because $R(q) = pq$, $MR = \mathrm{d}R(q)/\mathrm{d}q = \mathrm{d}(pq)/\mathrm{d}q = p$.

[11]The figure is based on Robidoux and Lester's (1988) estimated variable cost function. In the figure, we assume that the minimum of the average variable cost curve is \$5 at 50,000 metric tons of output. Based on information from Statistics Canada, we set the fixed cost so that the average cost is \$6 at 140,000 tons.

Figure 8.3 How a Competitive Firm Maximizes Profit

(a) A competitive lime manufacturing firm maximizes its profit at $\pi^* = \$426{,}000$, where its marginal revenue, MR, which is the market price, $p = \$8$, equals its marginal cost, MC. (b) The corresponding profit curve reaches its peak at 284 units of lime. The estimated cost curves are based on Robidoux and Lester (1988).



its average profit per unit is the difference between the market price (or average revenue), $p = R(q)/q = pq/q$, and its average cost, $AC = C(q)/q$:

$$\frac{\pi(q)}{q} = \frac{R(q) - C(q)}{q} = \frac{R(q)}{q} - \frac{C(q)}{q} = p - AC. \qquad (8.4)$$

At 284 units, the lime firm's average profit per unit is $\$1.50 = p - AC(284) = \$8 - \$6.50$, and the firm's profit is $\pi = \$1.50 \times 284{,}000 = \$426{,}000$. Panel b shows that this profit is the maximum possible profit because it is the peak of the profit curve.

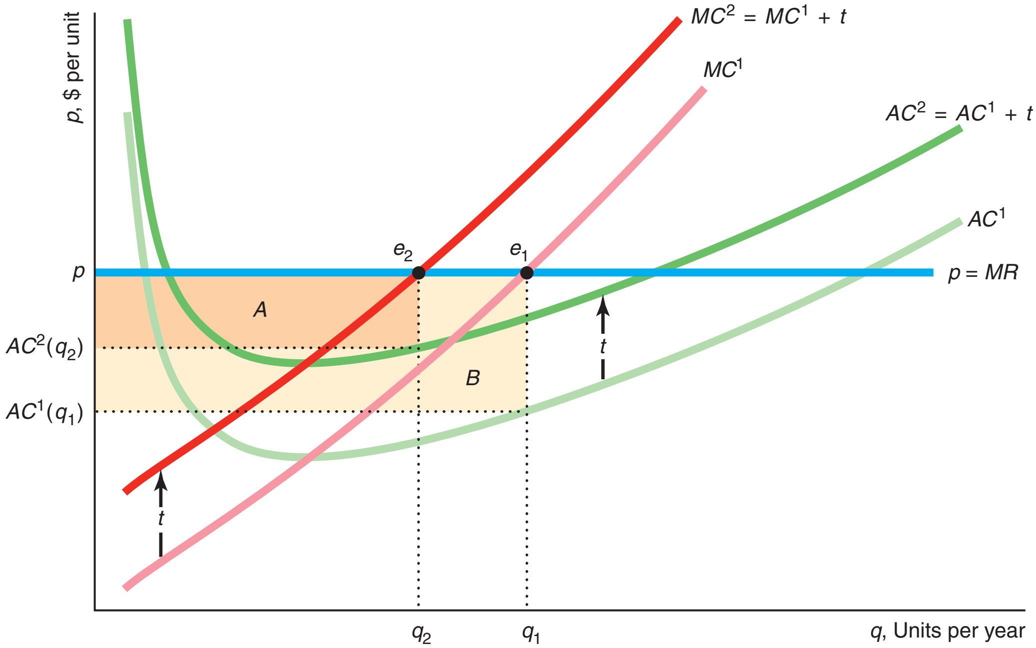
Solved Problem 8.2

MyLab Economics Solved Problem

If a competitive firm's cost increases due to an increase in the price of a factor of production or a tax, the firm's manager can quickly determine by how much to adjust output by calculating how the firm's marginal cost has changed and applying the profit-maximization rule. Suppose that the Canadian province of Manitoba imposes a specific (per-unit) tax of t per ton of lime produced in the province. No other provincial government imposes such a tax. Manitoba has only one lime-producing firm, so the tax affects only that firm and hence has virtually no effect on the market price. If the tax is imposed, how should the Manitoba firm change its output level to maximize its profit, and how does its maximum profit change?

Answer

1. *Show how the tax shifts the marginal cost and average cost curves*. The firm's before-tax marginal cost curve is MC^1 and its before-tax average cost curve is AC^1. Because the specific tax adds t to the per-unit cost, it shifts the after-tax marginal cost curve up to $MC^2 = MC^1 + t$ and the after-tax average cost curve to $AC^2 = AC^1 + t$ (see Chapter 7).



2. *Determine the before-tax and after-tax equilibria and the amount by which the firm adjusts its output*. Where the before-tax marginal cost curve, MC^1, hits the horizontal demand curve, p, at e_1, the profit-maximizing quantity is q_1. The after-tax marginal cost curve, MC^2, intersects the demand curve, p, at e_2 where the profit-maximizing quantity is q_2. Thus, in response to the tax, the firm produces $q_1 - q_2$ fewer units of output.

3. *Show how the profit changes after the tax*. Because the market price is constant but the firm's average cost curve shifts upward, the firm's profit at every output level falls. The firm sells fewer units (because of the increase in MC) and makes

less profit per unit (because of the increase in AC). The after-tax profit is area $A = \pi_2 = [p - AC^2(q_2)]q_2$, and the before-tax profit is area $A + B = \pi_1 = [p - AC^1(q_1)]q_1$, so profit falls by area B due to the tax.[12]

Short-Run Shutdown Decision

Once a firm determines the output level that maximizes its profit or minimizes its loss, it must decide whether to produce that output level or to shut down and produce nothing. This decision is easy for the lime firm in Figure 8.3 because, at the output that maximizes its profit, it makes a positive economic profit. However, the question remains whether a firm should shut down if it is making a loss in the short run.

All firms—not just competitive firms—use the same shutdown rule: The firm shuts down only if it can reduce its loss by doing so. The firm shuts down only if its revenue is less than its avoidable variable cost: $R(q) < VC(q)$. For a competitive firm, this rule is

$$pq < VC(q). \tag{8.5}$$

By dividing both sides of Equation 8.5 by output, we can write this condition as

$$p < \frac{VC(q)}{q} = AVC(q).$$

A competitive firm shuts down if the market price is less than its short-run average variable cost at the profit-maximizing quantity.

We illustrate the logic behind this rule for a competitive firm using our lime firm example. We look at three cases where the market price is (1) above the minimum average cost (AC), (2) less than the minimum average cost but at least equal to or above the minimum average variable cost, or (3) below the minimum average variable cost.

The Market Price Is Above the Minimum AC If the market price is above the firm's average cost at the quantity that it's producing, the firm makes a profit and so it operates. In panel a of Figure 8.3, the competitive lime firm's average cost curve reaches its minimum of \$6 per ton at 140 units. Thus, if the market price is above \$6, the firm makes a profit of $p - AC$ on each unit it sells and operates. In the figure, the market price is \$8, and the firm makes a profit of \$426,000.

The Market Price Is Between the Minimum AC and the Minimum AVC The tricky case is when the market price is less than the minimum average cost but at least as great as the minimum average variable cost. If the price is in this range, the firm makes a loss, but it reduces its loss by operating rather than shutting down.

Figure 8.4 (which reproduces the marginal and average cost curves for the lime firm from panel a of Figure 8.3 and adds the average variable cost curve) illustrates this case for the lime firm. The lime firm's average cost curve reaches a minimum of \$6 at 140 units, while its average variable cost curve hits its minimum of \$5 at 50 units. If the market price is between \$5 and \$6, the lime firm loses money (its profit is negative) because the price is less than its AC, but the firm does not shut down.

---

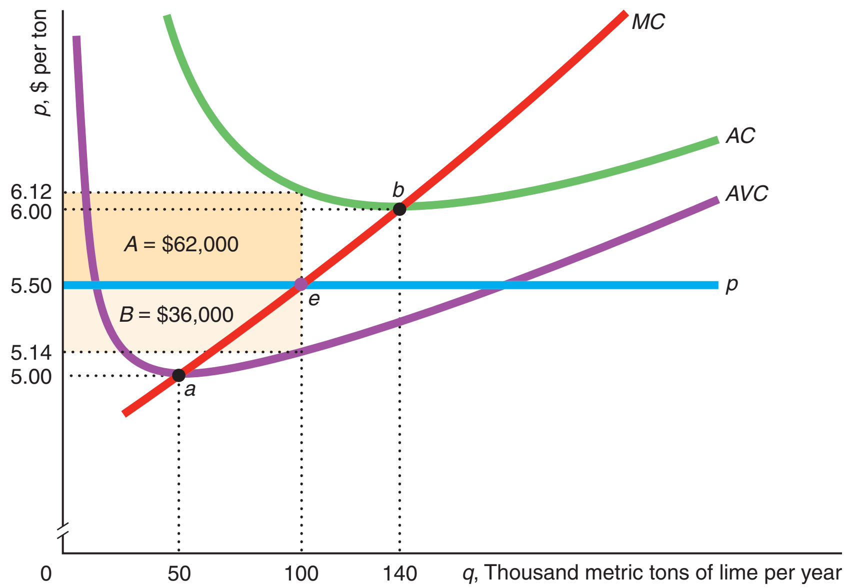
[12] We can solve this problem using calculus. The firm's profit is $\pi = pq - [C(q) + tq]$, where $C(q)$ is the firm's before-tax cost and $C(q) + tq$ is its after-tax cost. We obtain the necessary condition for the firm to maximize its after-tax profit by taking the first derivative of profit with respect to quantity and setting it equal to zero:

$$\frac{d\pi}{dq} = \frac{d(pq)}{dq} - \frac{d[C(q) + tq]}{dq} = p - \left[\frac{dC(q)}{dq} + t\right] = p - [MC + t] = 0.$$

Thus, the firm produces where $p = MC + t$.

Figure 8.4 The Short-Run Shutdown Decision MyLab Economics Video

The competitive lime manufacturing plant operates if price is above the minimum of the average variable cost curve, point *a*, at \$5. With a market price of \$5.50, the firm produces 100 units because that price is above $AVC(100) = \$5.14$, so the firm more than covers its out-of-pocket, variable costs. At that price, the firm makes a loss of area $A = \$62{,}000$ because the price is less than the average cost of \$6.12. If it shuts down, its loss is its fixed cost, area $A + B = \$98{,}000$. Thus, the firm does not shut down.



For example, if the market price is \$5.50, the firm minimizes its loss by producing 100 units where the marginal cost curve crosses the price line. At 100 units, the average cost is \$6.12, so the firm loses 62¢ [$= p - AC(100) = \$5.50 - \$6.12$] on each unit that it sells.

Why does the firm produce given that it is making a loss? The reason is that the firm reduces its loss by operating rather than shutting down because its revenue exceeds its variable cost—or equivalently, the market price exceeds its average variable cost.

If the firm shuts down in the short run it incurs a loss equal to its fixed cost of \$98,000, which is the sum of rectangles *A* and *B*.[13] If the firm operates and produces $q = 100$ units, its average variable cost is $AVC = \$5.14$, which is less than the market price of $p = \$5.50$ per ton. It makes $36¢ = p - AVC = \$5.50 - \$5.14$ more on each ton than its average variable cost. The difference between the firm's revenue and its variable cost, $R - VC$, is the rectangle $B = \$36{,}000$, which has a length of 100 thousand tons and a height of 36¢. Thus, if the firm operates, it loses only \$62,000 (rectangle *A*), which is less than its loss if it shuts down, \$98,000. The firm makes a smaller loss by operating than by shutting down because its revenue more than covers its variable cost and hence helps to reduce the loss from the fixed cost.

The Market Price Is Less than the Minimum AVC If the market price dips below the minimum of the average variable cost, \$5 in Figure 8.4, then the firm should shut down in the short run. Thus, the minimum of the average variable cost curve is the firm's *shutdown point*. At any price less than the minimum average variable cost, the firm's revenue is less than its variable cost, so it makes a greater loss by operating than by shutting down because it loses money on each unit sold in addition to the fixed cost that it loses if it shuts down.

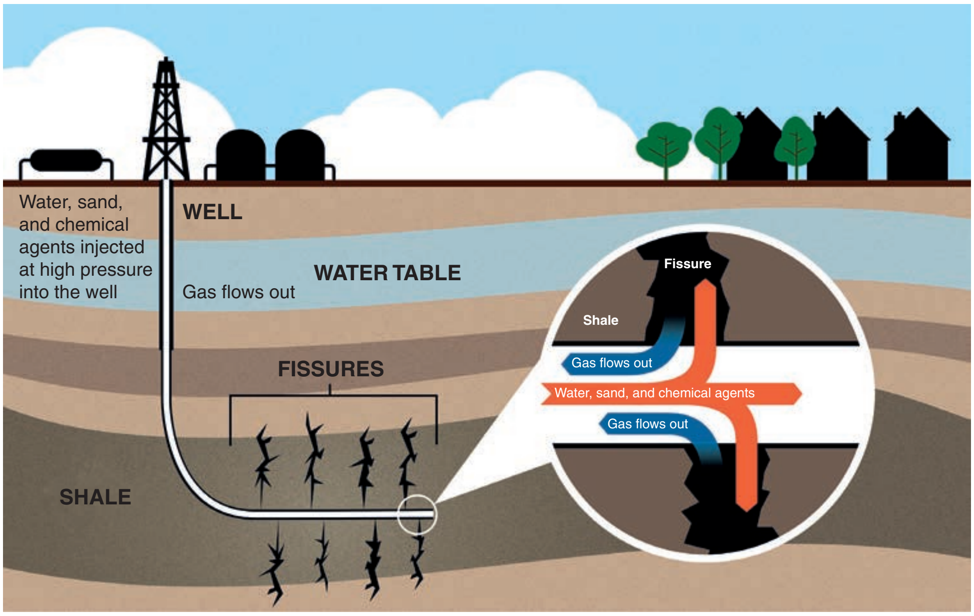
[13]The average cost is the sum of the average variable cost and the average fixed cost, $AC = AVC + F/q$ (Chapter 7). Thus, the gap between the average cost and the average variable cost curves at any given output is $AC - AVC = F/q$. Consequently, the height of the rectangle $A + B$ is $AC(100) - AVC(100) = F/100$, and the length of the rectangle is 100 units, so the area of the rectangle is F, or $\$98{,}000 = \$62{,}000 + \$36{,}000$.

In summary, a competitive firm uses two decisions to maximize its profit. First, the competitive firm determines the output that maximizes its profit or minimizes its loss when its marginal cost equals the market price (which is its marginal revenue): $p = MC$. Second, the firm chooses to produce that quantity unless it would lose more by operating than by shutting down. Thus, *a competitive firm shuts down in the short run only if the market price is less than the minimum of its average variable cost curve*.

Application

Fracking and Shutdowns

Oil production starts and stops in the short run as the market price fluctuates. In 1998–1999 when oil prices were historically low, U.S. oil-producing firms shut down or abandoned 74,000 of 136,000 oil wells. History repeats itself. From 2011 through the first half of 2014, oil prices were above $100 per barrel—nearly hitting $130 at one point—which was above the shutdown point for virtually all U.S. wells. However, when oil prices fell below $50 a barrel in 2015 and below $30 a barrel in 2016, many U.S. wells shut down.



Conventional oil wells—which essentially stick a pipe in the ground and pump oil—have low enough shutdown points that they operated in 2014 and 2015. Some Middle Eastern oil wells break even at a price as low as $10 a barrel. Older Texas wells often have a break-even point at $20 to $30 per barrel.

Most new U.S. oil wells use hydraulic fracturing (fracking). Fracking pumps pressurized liquid consisting of water, sand, and chemicals to fracture oil shale (rock containing oil), which releases natural gas and oil.[14] Current fracking operations have a shutdown point at between $50 and $77 per barrel with an average of about $65. Thus, fracking operations were more likely to shut down than were conventional wells during the recent period of low prices.

[14]The first fracking experiment was in 1947. At first, fracking wells had a minimum average variable cost that was too high to operate profitably. However, in recent years, technological innovation substantially lowered this cost, and the global price of oil was often high enough for fracking to be widely used. Due to fracking, U.S. oil production rose from 5.6 million barrels a day in 2010 to 9.3 million in 2015. Fracking is controversial because opponents fear it will create environmental problems and trigger earthquakes.

Solved Problem 8.3

A competitive firm's bookkeeper, upon reviewing the firm's books, finds that the firm spent twice as much on its plant, a fixed cost, as the firm's manager had previously thought. Should the manager change the output level because of this new information? How does this new information affect profit?

Answer

1. *Show that a change in fixed costs does not affect the firm's decisions*. How much the firm produces and whether it shuts down in the short run depend only on the firm's variable costs. (The firm picks its output level so that its marginal cost—which depends only on variable costs—equals the market price, and it shuts down only if market price is less than its minimum average variable cost.) Learning that the amount spent on the plant was greater than previously believed should not change the output level that the manager chooses.

2. *Show that the change in how the bookkeeper measures fixed costs does not affect economic profit*. The change in the bookkeeper's valuation of the historical amount spent on the plant may affect the firm's short-run business profit but does not affect the firm's true economic profit. The economic profit is based on opportunity costs—the amount for which the firm could rent the plant to someone else—and not on historical payments.

Short-Run Firm Supply Curve

We just demonstrated how a competitive firm chooses its output for a given market price in a way that maximizes its profit or minimizes its losses. By repeating this analysis at different possible market prices, we can show how the amount the competitive firm supplies varies with the market price.

As the market price increases from $p_1 = \$5$ to $p_2 = \$6$ to $p_3 = \$7$ to $p_4 = \$8$, the lime firm increases its output from 50 to 140 to 215 to 285 units per year, as Figure 8.5 shows. The profit-maximizing output at each market price is determined by the intersection of the relevant demand curve—market price line—and the firm's marginal cost curve, as equilibria e_1 through e_4 illustrate. That is, as the market price increases, the equilibria trace out the marginal cost curve. However, if the price falls below the firm's minimum average variable cost at \$5, the firm shuts down. Thus, the competitive firm's short-run supply curve is its marginal cost curve above its minimum average variable cost.

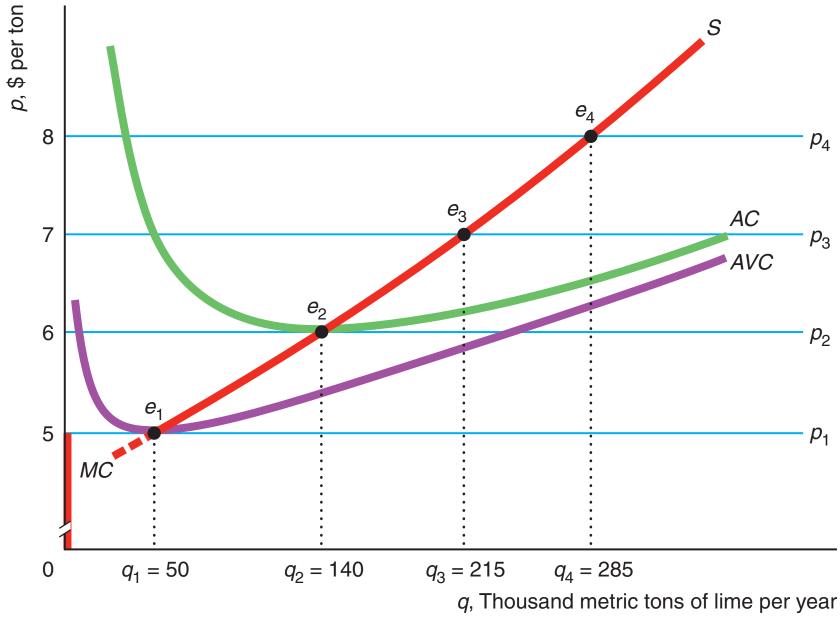
The firm's short-run supply curve, S, is a solid red line in the figure. At prices above \$5, the short-run supply curve is the same as the marginal cost curve. The supply is zero when price is less than the minimum of the AVC curve of \$5. (From now on, to keep the graphs as simple as possible, we will not show the supply curve at prices below the minimum AVC.)

Short-Run Market Supply Curve

The market supply curve is the horizontal sum of the supply curves of all the individual firms in the market (Chapter 2). In the short run, the maximum number of firms in a market, n, is fixed because new firms need time to enter the market. If all the firms in a competitive market are identical, each firm's supply curve is identical, so the market supply at any price is n times the supply of an individual firm. Where firms have different shutdown prices, the market supply reflects a different number of firms at various prices even in the short run. We examine competitive markets first with firms that have identical costs and then with firms that have different costs.

Figure 8.5 How the Profit-Maximizing Quantity Varies with Price MyLab Economics Video

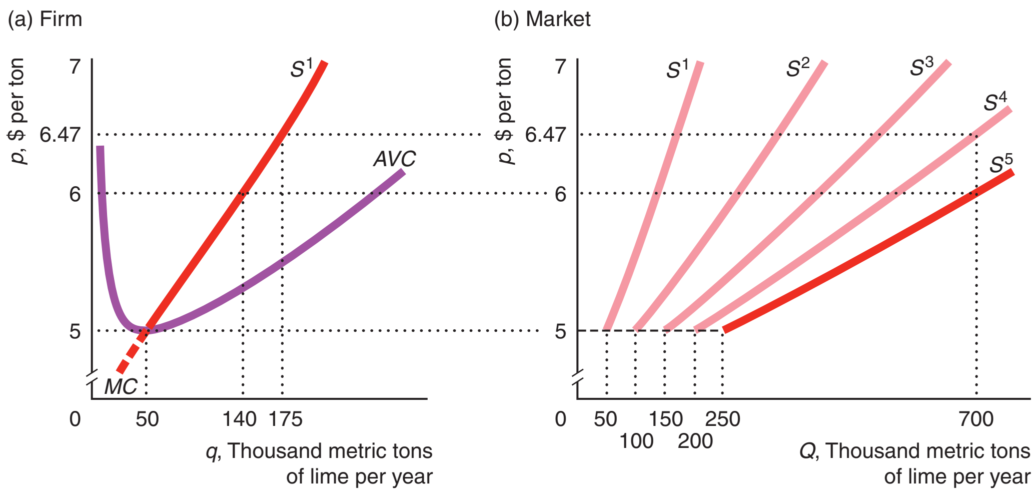
As the market price increases, the lime manufacturing firm produces more output. As the market price changes, it traces out the marginal cost (MC) curve of the firm. The firm's short-run supply (S) curve is the MC curve above the minimum of its AVC curve (at e_1).



Short-Run Market Supply with Identical Firms To illustrate how to construct a short-run market supply curve, we suppose that the lime manufacturing market has $n = 5$ competitive firms with identical cost curves. Panel a of Figure 8.6 plots the short-run supply curve, S^1, of a typical firm—the MC curve above the minimum AVC—where the horizontal axis shows the firm's output, q, per year. Panel b

Figure 8.6 Short-Run Market Supply with Five Identical Lime Firms MyLab Economics Video

(a) The short-run supply curve, S^1, for a typical lime manufacturing firm is its MC above the minimum of its AVC. (b) The market supply curve, S^5, is the horizontal sum of the supply curves of each of the five identical firms. Similarly, S^4 is the sum of the supply curves with only four firms in the market; S^3 is the sum with three firms; and so on.



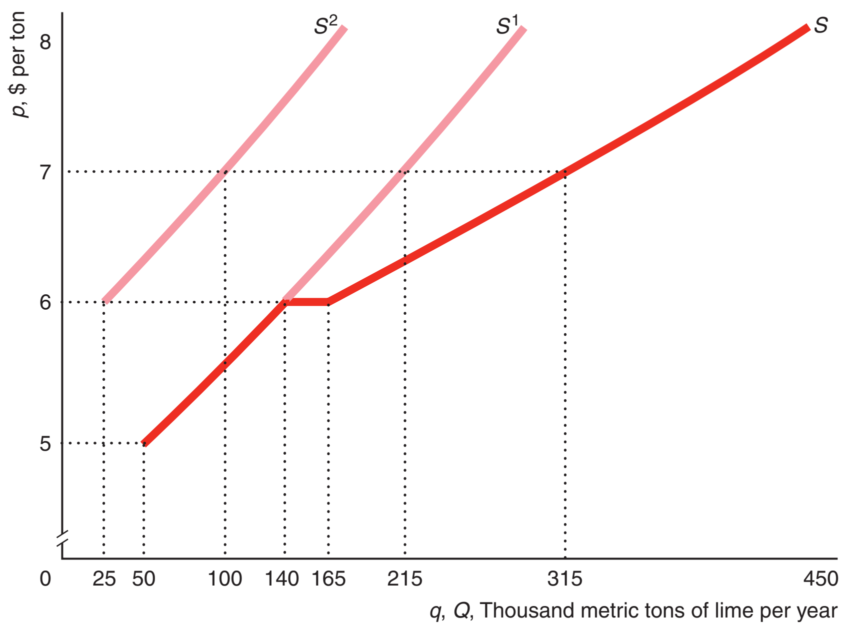
illustrates the competitive market supply curve, the dark line S^5, where the horizontal axis is market output, Q, per year. The price axis is the same in the two panels, but the scales differ for the quantity axes.

If the market price is less than \$5 per ton, no firm supplies any output, so the market supply is zero. At \$5, each firm is willing to supply $q = 50$ units, as in panel a. Consequently, the market supply is $Q = 5q = 250$ units in panel b. At \$6 per ton, each firm supplies 140 units, so the market supply is 700 ($= 5 \times 140$) units.

Suppose, however, that the market has fewer than five firms in the short run. The light-color lines in panel b show the market supply curves for various other numbers of firms. The market supply curve is S^1 with one price-taking firm, S^2 with two firms, S^3 with three firms, and S^4 with four firms. The market supply curve flattens as the number of firms in the market increases because the market supply curve is the horizontal sum of more and more upward-sloping firm supply curves. As the number of firms grows very large, the market supply curve approaches a horizontal line at \$5. Thus, *the more identical firms producing at a given price, the flatter (more elastic) the short-run market supply curve at that price*. As a result, the more firms in the market, the less the price needs to increase for the short-run market supply to increase substantially. Consumers pay \$6 per ton to obtain 700 units of lime with five firms but must pay \$6.47 per ton to obtain that much with only four firms in the market.

Short-Run Market Supply with Firms That Differ If the firms in a competitive market have different minimum average variable costs, not all firms produce at every price. These varying shutdown points affect the shape of the short-run market supply curve. Suppose that the only two firms in the lime market are our typical lime firm with a supply curve of S^1 and another firm with a higher marginal and minimum average cost with the supply curve of S^2 in Figure 8.7. The first firm produces if the market price is at least \$5, whereas the second firm does not produce unless the price is \$6 or more. At \$5, the first firm produces 50 units, so the quantity on the market supply curve, S, is 50 units. Between \$5 and \$6, only the first firm produces, so the

Figure 8.7 Short-Run Market Supply with Two Different Lime Firms



The supply curve S^1 is the same as for the typical lime firm in Figure 8.6. A second firm has an MC that lies to the left of the original firm's cost curve and a higher minimum of its AVC. Thus, its supply curve, S^2, lies above and to the left of the original firm's supply curve, S^1. The market supply curve, S, is the horizontal sum of the two supply curves. When the price is \$6 or higher, both firms produce, and the market supply curve is flatter than the supply curve of either individual firm.

market supply, S, is the same as the first firm's supply, S^1. At and above \$6, both firms produce, so the market supply curve is the horizontal summation of their two individual supply curves. For example, at \$7, the first firm produces 215 units, and the second firm supplies 100 units, so the market supply is 315 units.

As with the identical firms, where both firms are producing, the market supply curve is flatter than that of either firm. Because the second firm does not produce at as low a price as the first firm, the short-run market supply curve has a steeper slope (less elastic supply) at relatively low prices than it would if the firms were identical.

Where firms differ, only the low-cost firm supplies goods at relatively low prices. As the price rises, the other, higher-cost firm starts supplying, creating a stair-like market supply curve. The more suppliers with differing costs, the more steps in the market supply curve. As price rises and more firms are supplying goods, the market supply curve flattens, so it takes a smaller increase in price to increase supply by a given amount. Differences in cost curves across firms are one explanation for why some market supply curves are upward sloping.

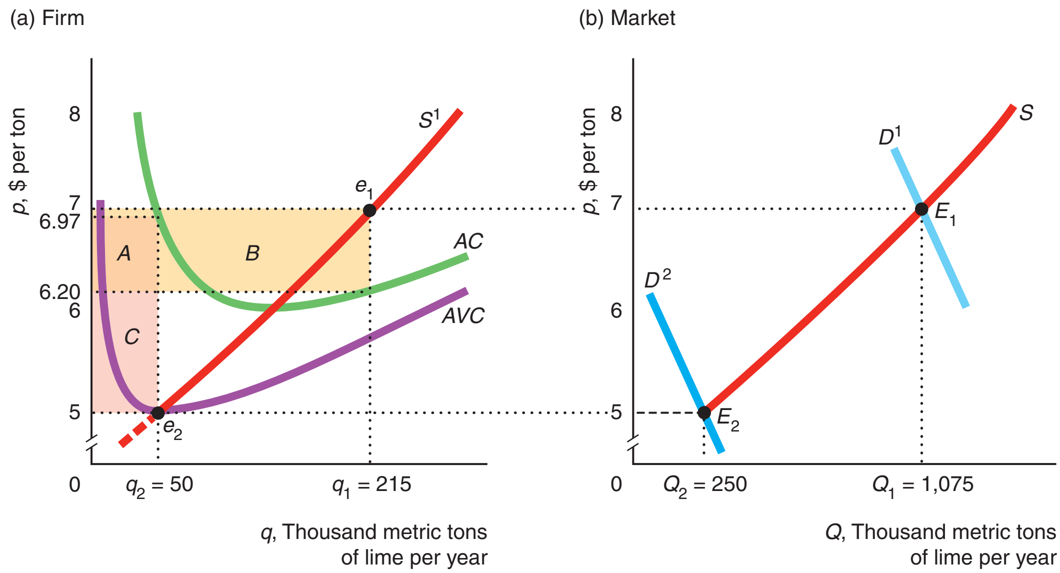
Short-Run Competitive Equilibrium

By combining the short-run market supply curve and the market demand curve, we can determine the short-run competitive equilibrium. We first show how to determine the equilibrium in the lime market, and we then examine how the equilibrium changes when firms are taxed.

Suppose that the lime market has five identical firms in the short-run equilibrium. Panel a of Figure 8.8 shows the short-run cost curves and the supply curve, S^1, for

Figure 8.8 Short-Run Competitive Equilibrium in the Lime Market MyLab Economics Video

(a) A lime firm's short-run supply curve is the marginal cost curve above minimum average variable cost of \$5. If the price is \$5, each firm makes a short-run loss of \$98,500 [= $(p - AC)q$ = (\$5 − \$6.97) × 50,000], area $A + C$. At a price of \$7, the short-run profit of a typical lime firm is \$172,000 [= $(p - AC)q$ = (\$7 − \$6.20) × 215,000], area $A + B$. (b) If the lime market has only five firms in the short run, the market supply curve is S. The market demand curve, D^1, intersects S at the short-run equilibrium, E_1, where the market price is \$7 and market output is Q_1 = 1,075 units. If the demand curve shifts to D^2, the market equilibrium is p = \$5 and Q_2 = 250 units.



a typical firm, and panel b shows the corresponding short-run competitive market supply curve, S.

In panel b, the initial demand curve, D^1, intersects the market supply curve at E_1, the market equilibrium. The equilibrium quantity is $Q_1 = 1{,}075$ units of lime per year, and the equilibrium market price is \$7.

Each competitive firm faces a horizontal demand curve at the equilibrium price of \$7. Each price-taking firm chooses its output where its marginal cost curve intersects the horizontal demand curve at e_1 in panel a. Because each firm is maximizing its profit at e_1, no firm wants to change its behavior, so e_1 is the firm's equilibrium. Each firm makes a short-run profit of area $A + B$ = \$172,000, which is the average profit per ton, $p - AC$ = \$7 − \$6.20 = 80¢, times the firm's output, $q_1 = 215$ units. The equilibrium market output, Q_1, is the number of firms, n, times the equilibrium output of each firm: $Q_1 = nq_1 = 5 \times 215$ units $= 1{,}075$ units (panel b).

Now suppose that the demand curve shifts to D^2. The new market equilibrium is E_2, where the price is only \$5. At that price, each firm produces $q = 50$ units, and market output is $Q = 250$ units. In panel a, each firm loses \$98,500, area $A + C$, because it makes an average per ton of ($p - AC$) = (\$5 − \$6.97) = −\$1.97 and it sells $q_2 = 50$ units. However, such a firm does not necessarily shut down because price equals the firm's average variable cost, so the firm is covering its out-of-pocket expenses.

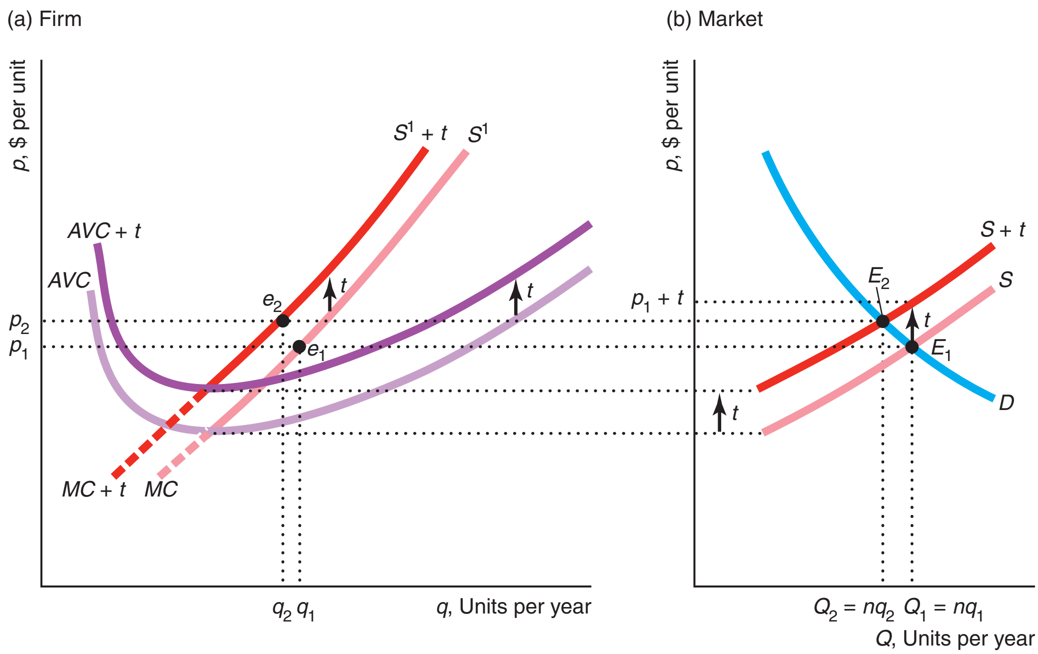
Solved Problem 8.4

MyLab Economics Solved Problem

What is the effect on the short-run equilibrium of a specific tax of t per unit that is collected from all n firms in a market? What is the incidence of the tax?

Answer

1. *Show how the tax shifts a typical firm's marginal cost and average cost curves and hence its supply curve.* In Solved Problem 8.2, we showed that such a tax



causes the marginal cost curve, the average cost curve, and (hence) the minimum average cost of the firm to shift up by t, as illustrated in panel a of the figure. As a result, the short-run supply curve of the firm, labeled $S^1 + t$, shifts up by t from the pre-tax supply curve, S^1.

2. *Show how the market supply curve shifts.* The market supply curve is the sum of all the individual firm supply curves, so it too shifts up by t, from S to $S + t$ in panel b of the figure.

3. *Determine how the short-run market equilibrium changes.* The pre-tax, short-run market equilibrium is E_1, where the downward-sloping market demand curve D intersects S in panel b. In that equilibrium, price is p_1 and quantity is Q_1, which equals n (the number of firms) times the quantity q_1 that a typical firm produces at p_1. The after-tax, short-run market equilibrium, E_2, determined by the intersection of D and the after-tax supply curve, $S + t$, occurs at p_2 and Q_2. Because the after-tax price p_2 is above the after-tax minimum average variable cost, all the firms continue to produce, but they produce less than before: $q_2 < q_1$. Consequently, the equilibrium quantity falls from $Q_1 = nq_1$ to $Q_2 = nq_2$.

4. *Discuss the incidence of the tax.* The equilibrium price increases, but by less than the full amount of the tax: $p_2 < p_1 + t$. The incidence of the tax is shared between consumers and producers because both the supply and the demand curves are sloped (Chapter 3).

8.4 Competition in the Long Run

I think there is a world market for about five computers.
—Thomas J. Watson, IBM chairman, 1943

In the long run, competitive firms can vary inputs that were fixed in the short run, so the long-run firm and market supply curves differ from the short-run curves. After briefly looking at how a firm determines its long-run supply curve so as to maximize its profit, we examine the relationship between short-run and long-run market supply curves and competitive equilibria.

Long-Run Competitive Profit Maximization

The firm's two profit-maximizing decisions—how much to produce and whether to produce at all—are simpler in the long run than in the short run. In the long run, typically all costs are variable, so the firm does not have to consider whether fixed costs are sunk or avoidable.

The firm chooses the quantity that maximizes its profit using the same rules as in the short run. The firm picks the quantity that maximizes long-run profit, which is the difference between revenue and long-run cost. Equivalently, it operates where long-run marginal profit is zero and where marginal revenue equals long-run marginal cost.

After determining the output level, q^*, that maximizes its profit or minimizes its loss, the firm decides whether to produce or shut down. As always, the firm shuts down if its revenue is less than its avoidable cost. Because all costs are variable in the long run, the firm shuts down in the long run if it would make an economic loss by operating.

Long-Run Firm Supply Curve

A firm's long-run supply curve is its long-run marginal cost curve above the minimum of its long-run average cost curve (because all costs are variable in the long run). The firm is free to choose its capital in the long run, so the firm's long-run supply curve may differ substantially from its short-run supply curve.

The firm chooses a plant size to maximize its long-run economic profit in light of its beliefs about the future. If its forecast is wrong, it may be stuck with a plant that is too small or too large for its level of production in the short run. The firm acts to correct this mistake in plant size in the long run.

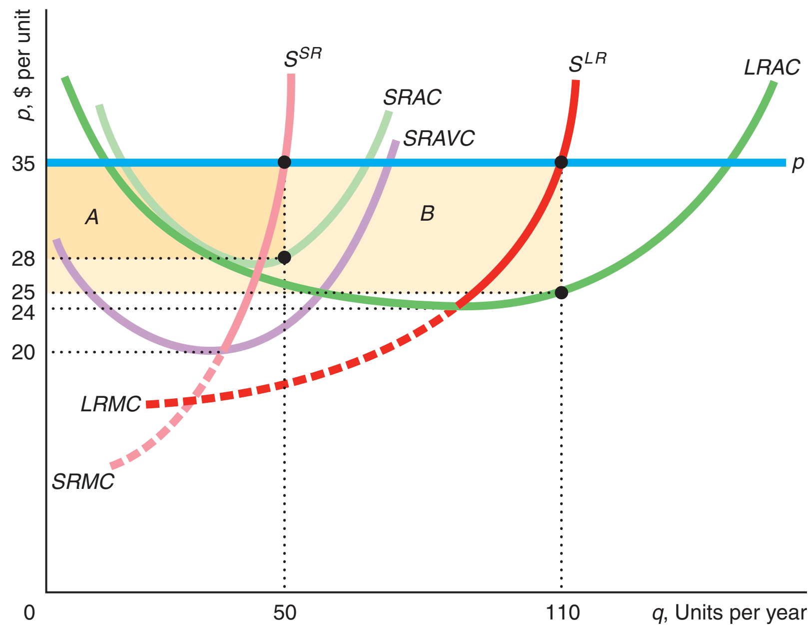
The firm in Figure 8.9 has different short- and long-run cost curves. In the short run, the firm uses a plant that is smaller than the optimal long-run size if the price is \$35. (Having a short-run plant size that is too large is also possible.) The firm produces 50 units of output per year in the short run, where its short-run marginal cost, *SRMC*, equals the price, and makes a short-run profit equal to area *A*. The firm's short-run supply curve, S^{SR}, is its short-run marginal cost above the minimum, \$20, of its short-run average variable cost, *SRAVC*.

If the firm expects the price to remain at \$35, it builds a larger plant in the long run. Using the larger plant, the firm produces 110 units per year, where its long-run marginal cost, *LRMC*, equals the market price. It expects to make a long-run profit, area $A + B$, which is greater than its short-run profit by area *B* because it sells 60 more units and its equilibrium long-run average cost, $LRAC = \$25$, is lower than its short-run average cost in equilibrium, \$28.

The firm does not operate at a loss in the long run when all inputs are variable. It shuts down if the market price falls below the firm's minimum long-run average cost of \$24. Thus, the competitive firm's long-run supply curve is its long-run marginal cost curve above \$24.

Figure 8.9 The Short-Run and Long-Run Supply Curves

The firm's long-run supply curve, S^{LR}, is zero below its minimum average cost of \$24 and equals the long-run marginal cost, *LRMC*, at higher prices. The firm produces more in the long run than in the short run, 110 units instead of 50 units, and earns a higher profit, area $A + B$ instead of just area *A*.



Application

The Size of Ethanol Processing Plants

When a large number of firms initially built ethanol processing plants, they built relatively small ones. When the ethanol market took off in the first few years of the twenty-first century, with the price reaching a peak of $4.23 a gallon in June 2006, many firms built larger plants or greatly increased their plant size. From 1999 to 2006, the number of plants nearly doubled and the average plant capacity nearly tripled (36 to 106 million gallons per year).

However, since then, the ethanol market price has collapsed. The price was generally below $3 and often below $1.50 from 2007 through 2016, hitting a low of $1.26 in January 2016. As a result, many firms closed plants or reduced their size. The average plant capacity fell by a third from 2006 to 2016 (106 to 73 million gallons per year).

Long-Run Market Supply Curve

The competitive market supply curve is the horizontal sum of the supply curves of the individual firms in both the short run and the long run. Because the maximum number of firms in the market is fixed in the short run (it takes time to build a new plant, buy equipment, and hire workers), we add the supply curves of a known number of firms to obtain the short-run market supply curve. The only way for the market to supply more output in the short run is for existing firms to produce more.

However, in the long run, firms can enter the market. Thus, before we can add all the relevant firm supply curves to obtain the long-run market supply curve, we need to determine how many firms are in the market at each possible market price. We now look in detail at how market entry and exit affect the long-run market supply curve.

Entry and Exit *Entry* and *exit* by firms determines the long-run number of firms in a market. In the long run, each firm decides whether to enter or exit depending on whether it can make a long-run profit:

- A firm enters the market if it can make a long-run profit, $\pi > 0$.
- A firm exits the market to avoid a long-run loss, $\pi < 0$.

If firms in a market are making zero long-run profit, they are indifferent between staying in the market and exiting. We presume that if they are already in the market, they stay in the market when they are making zero long-run profit.

In the United States in the fourth quarter of 2014, 238 thousand firms entered the market and 197 thousand firms exited.[15] The annual rates of entry and exit of such firms are both about 10% of the total number of firms per year.

Even in the long run, entry is limited in many markets, such as manufacturing, because firms face significant costs to enter, such as large start-up costs. In other markets, government restrictions create a barrier to entry. For instance, many city governments limit the number of liquor stores, creating an insurmountable barrier that prevents new ones from entering. Similarly, patent protection prevents new firms from producing the patented product until the patent expires.

However, in unregulated, perfectly competitive markets, firms can enter and exit freely in the long run. For example, many construction firms that provide only labor services enter and exit a market several times a year.

[15] **http://www.bls.gov/web/cewbd/table9_1.txt** (viewed June 16, 2016).

In markets with free entry, when the demand curve shifts to the right so that the market price and profit rise, entry occurs until the last firm to enter—the *marginal firm*—makes zero long-run profit. Similarly in markets with free exit, if the demand curve shifts to the left so that the market price drops, firms with minimum average costs above the new, lower market price exit the market. Firms continue to exit until the next firm considering leaving—the marginal firm—is again making a zero long-run profit.

Application

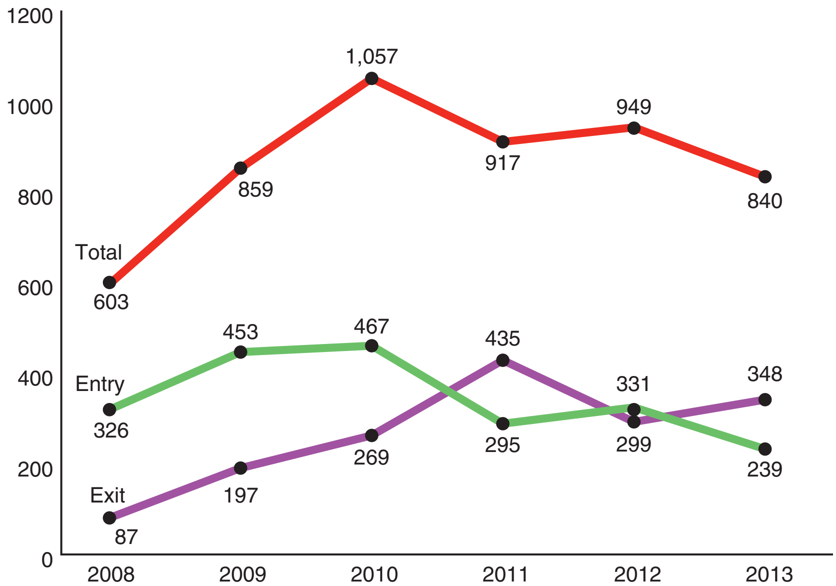
Entry and Exit of Solar Power Firms

About one-half of all U.S. residential solar energy systems are installed in California. The incentives of firms to enter or exit the California solar market have changed considerably since 2007.

The manufacturing cost of a solar energy system has plummeted since 2007. California offered a large residential solar subsidy, which fell significantly from 2007 to 2013, when it ended. Homeowners could also get a federal subsidy of 30% on the cost of a system up to $2,000. In 2009, the $2,000 cap was eliminated. The federal program is scheduled to end in 2016.

Solar firms differ substantially. A new firm with relatively low costs or that expects demand to rise may enter the market in the same year that another firm with higher costs or less rosy expectations exits the market.

The figure (based on Davis and Chun, 2015) shows the number of firms that entered the California solar market (green), the number that exited (purple), and the total number of firms (red) annually from 2008–2013. For example in 2009, 453 firms entered and 197 firms exited so the total number of firms increased by 256, from 603 in 2008 to 859 in 2009.



From 2008–2010, the number of firms that entered the market was much greater than the number of firms that exited. The total number of firms rose by 75% from 603 in 2008 to 1,057 in 2010. In each of these three rapid growth years, nearly half the firms in the market were new to the market that year.

Since 2010, the number of firms fell, dropping from the 2010 peak of 1,057 to 840 in 2013. Even during this period, about 30% of the firms were new in each year. However, entry was more than offset by exits in 2011 and 2013 (the number of firms entering and exiting were roughly equal in 2012).

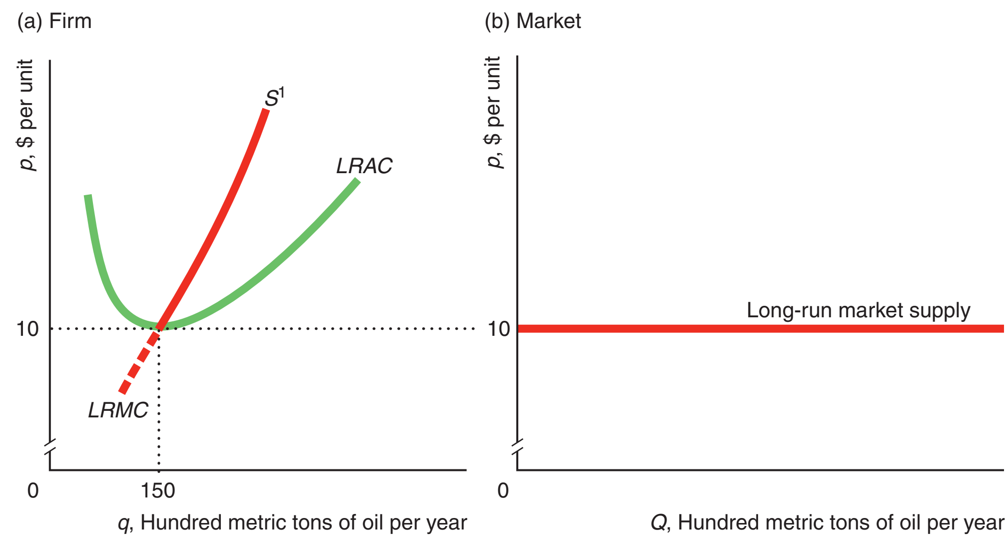
Given the large number of firms entering and exiting the market each year, the costs of entry and exit must be minimal. The number of firms in this market is enormous. Most of these firms are small. The median company installs only about five systems per year, although SolarCity and a few other large firms install many systems annually.

Long-Run Market Supply with Identical Firms and Free Entry The *long-run market supply curve is flat* at the minimum long-run average cost *if firms can freely enter and exit* the market, an unlimited number of *firms have identical costs*, and *input prices are constant*. This result follows from our reasoning about the short-run supply curve, in which we showed that the more firms in the market, the flatter the market supply curve. With many firms in the market in the long run, the market supply curve is effectively flat. ("Many" is 10 firms in the vegetable oil market.)

The long-run supply curve of a typical vegetable oil mill, S^1 in panel a of Figure 8.10, is the long-run marginal cost curve above a minimum long-run average cost of $10. Because each firm shuts down if the market price is below $10, the long-run market supply curve is zero at a price below $10. If the price rises above $10, firms are making positive profits, so new firms enter, expanding market output until profits

Figure 8.10 Long-Run Firm and Market Supply with Identical Vegetable Oil Firms

(a) The long-run supply curve of a typical vegetable oil mill, S^1, is the long-run marginal cost curve above the minimum average cost of $10. (b) The long-run market supply curve is horizontal at the minimum of the long-run minimum average cost of a typical firm. Each firm produces 150 units, so market output is $150n$, where n is the number of firms.



are driven to zero, where price is again \$10. The long-run market supply curve in panel b is a horizontal line at the minimum long-run average cost of the typical firm, \$10. At a price of \$10, each firm produces $q = 150$ units (where one unit equals 100 metric tons). Thus, the total output produced by n firms in the market is $Q = nq = n \times 150$ units. Extra market output is obtained by new firms entering the market.

In summary, the long-run market supply curve is horizontal if the market has free entry and exit, an unlimited number of firms have identical costs, and input prices are constant.

We next examine four reasons why a long-run market supply curve is not flat:

1. limited entry,
2. firms' cost functions differ,
3. input prices vary with output,
4. a buyer is large (demands a large share of a good from the market).

Long-Run Market Supply with Limited Entry First, if the number of firms in a market is limited in the long run, the long-run market supply curve slopes upward. The number of firms is limited if the government restricts that number, if firms need a scarce resource, or if entry is costly. An example of a scarce resource is the limited number of lots on which a luxury beachfront hotel can be built in Miami Beach. High entry costs restrict the number of firms in a market because firms enter only if the long-run economic profit is greater than the cost of entering.

The only way to get more output if the number of firms is limited is for existing firms to produce more. Because individual firms' supply curves slope upward, the long-run market supply curve is also upward sloping. The reasoning is the same as in the short run, as panel b of Figure 8.6 illustrates, given that no more than five firms can enter. The market supply curve is the upward-sloping S^5 curve, which is the horizontal sum of the five firms' upward-sloping marginal cost curves above minimum average cost.

Long-Run Market Supply When Firms' Cost Functions Differ A second reason why some long-run market supply curves slope upward is that firms differ. Firms with relatively low minimum long-run average costs are willing to enter the market at lower prices than others, resulting in an upward-sloping long-run market supply curve (similar to the short-run example in Figure 8.7).

Many markets have a number of low-cost firms and other higher-cost firms.[16] If lower-cost firms can produce as much output as the market wants, only low-cost firms produce, and the long-run market supply curve is horizontal at the minimum of the low-cost firm's average cost curve. The long-run supply curve is upward sloping *only* if lower-cost firms cannot produce as much output as the market demands because each of these firms has a limited capacity and the number of these firms is limited.

---

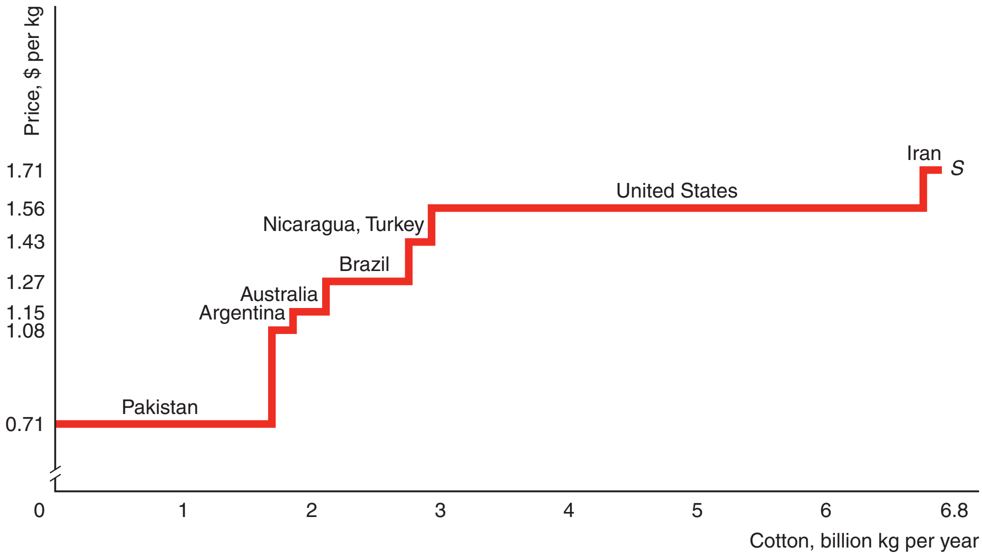
[16]Syverson (2004) estimated that, in the typical narrowly defined (4-digit) U.S. manufacturing industry, the 90th percentile plant produces 90% more output from the same input as the 10th percentile plant. Atalay (2012) found that 7% of the standard deviation (a measure of variation) of these plant-level productivities is due to differences in the price of materials that plants face. Thus, these markets have substantial variations in costs.

Application

Upward-Sloping Long-Run Supply Curve for Cotton

Many countries produce cotton. Production costs differ among countries because of differences in the quality of land, rainfall, costs of irrigation, costs of labor, and other factors.

The length of each step-like segment of the long-run supply curve of cotton in the graph is the quantity produced by the labeled country. The amount that the low-cost countries can produce must be limited, or we would not observe production by the higher-cost countries.



The height of each segment of the supply curve is the typical minimum average cost of production in that country. The average cost of production in Pakistan is less than half that in Iran. The supply curve has a step-like appearance because we are using an average of the estimated average cost in each country, which is a single number. If we knew the individual firms' supply curves in each of these countries, the market supply curve would have a smoother shape.

As the market price rises, the number of countries producing rises. At market prices below $1.08 per kilogram, only Pakistan produces. If the market price is below $1.50, the United States and Iran do not produce. If the price increases to $1.56, the United States supplies a large amount of cotton. In this range of the supply curve, supply is very elastic. For Iran to produce, the price has to rise to $1.71. Price increases in that range result in only a relatively small increase in supply. Thus, the supply curve is relatively inelastic at prices above $1.56.

Long-Run Market Supply When Input Prices Vary with Output A third reason why long-run market supply curves may slope upward is non-constant input prices. In markets where factor prices rise when output increases, the long-run supply curve slopes upward even if firms have identical costs and can freely enter and exit. (Similarly, the long-run supply slopes downward if factor prices fall when output increases.)

If a market's product uses a relatively small share of the total quantity of a factor of production, as that market's output expands, the price of the factor is not likely to change. For example, dentists do not hire enough receptionists to affect the market wage for receptionists.

In contrast, if the market's product uses a very large share of a factor, the price of that input is more likely to vary with that market's output. As jet aircraft manufacturers expand and buy more jet engines, the price of these engines rises because the aircraft manufacturers are the sole purchasers of these engines.

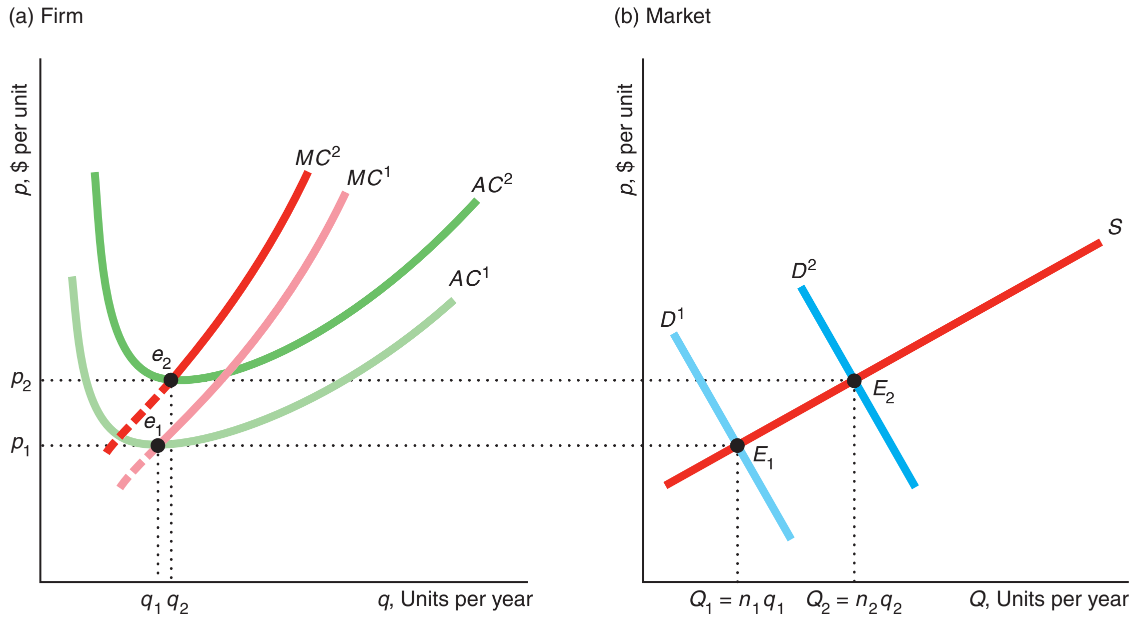
To produce a larger quantity in a market, firms must use more inputs. If as the firms use more of some or all inputs, the prices of those inputs may rise, so that the cost of producing the final good also rises. We call a market in which input prices rise with output an *increasing-cost market*. Because most ironworkers are afraid of heights, few of them are willing to construct tall buildings, so their supply curve is steeply upward sloping. As the number of skyscrapers under construction skyrockets, the demand curve for these workers shifts to the right, the equilibrium moves up along the supply curve, and their wage rises.

We assume that all firms in a market have the same cost curves and that input prices rise as market output expands. We use the cost curves of a representative firm in panel a of Figure 8.11 to derive the upward-sloping market supply curve in panel b.

The initial demand curve is D^1 in panel b. The market price is p_1 and the market produces relatively little output, Q_1, so input prices are relatively low. In panel a, each firm has the same long-run marginal cost curve, MC^1, and average cost curve, AC^1. Each firm produces at minimum average cost, e_1, and sells q_1 units of output. The n_1 firms collectively sell $Q_1 = n_1q_1$ units of output, which is point E_1 on the market supply curve in panel b.

Figure 8.11 Long-Run Firm and Market Supply in an Increasing-Cost Market

With the initial demand curve, D_1, and the relatively low market output, Q_1, in panel b, each firm's long-run marginal and average cost curves are MC^1 and AC^1 in panel a. When the demand shifts rightward to D^2, market quantity increases to Q_2, which increases input prices, so the cost curves shift upward to MC^2 and AC^2. Each firm produces at minimum average cost, such as points e_1 and e_2. The long-run market supply, S, in panel b is upward sloping.



If the market demand curve shifts outward to D^2 (panel b), the market price rises to p_2, new firms enter, and market output rises to Q_2, causing input prices to rise. As a result, the marginal cost curve shifts from MC^1 to MC^2, and the average cost curve rises from AC^1 to AC^2. The typical firm produces at a higher minimum average cost, e_2. At this higher price, the market has n_2 firms, so market output is $Q_2 = n_2 q_2$ at point E_2 on the market supply curve.

Thus, in both an increasing-cost market and a *constant-cost market*—where input prices remain constant as output increases—firms produce at minimum average cost in the long run. The difference is that the minimum average cost rises as market output increases in an increasing-cost market, whereas minimum average cost remains constant in a constant-cost market. In conclusion, *the long-run supply curve is upward sloping in an increasing-cost market and flat in a constant-cost market*.

In a decreasing-cost market, as market output rises, at least some factor prices fall. As a result, in a decreasing-cost market, the long-run market supply curve is downward sloping.

Increasing returns to scale may cause factor prices to fall. For example, when firms introduced Blu-ray drives, they manufactured and sold relatively few drives, and the cost of manufacturing was relatively high. Due to the high price of Blu-ray drives and the lack of Blu-ray disks, consumers demanded few Blu-ray drives. As demand for Blu-ray drives increased, it became practical to automate more of the production process so firms could produce drives at a lower average cost. The resulting decrease in the price of these drives lowered the cost of personal computers with these drives. Thus, theory tells us that competitive long-run market supply curves may be flat, upward sloping, or downward sloping.

Long-Run Market Supply Curve with a Large Buyer A fourth reason why a market supply curve may slope is that a buyer, such as a country, demands a large share of a good sold in the market. Many goods such as cotton and oil are traded on world markets. The world equilibrium price and quantity for a good are determined by the intersection of the world supply curve—the horizontal sum of the supply curves of each producing country—and the world demand curve—the horizontal sum of the demand curves of each consuming country.

A country that imports a good—a buyer—has a supply curve that is the horizontal sum of its domestic industry's supply curve and the import supply curve. The domestic long-run supply curve is the sum of the supply curves of all the domestic firms that we have just derived. However, we need to determine the import supply curve.

residual supply curve the quantity the market supplies that is not consumed by other demanders at any given price

The country imports the world's residual supply, where the **residual supply curve** is the quantity the market supplies that is not consumed by other demanders at any given price.[17] The country's import supply function is its residual supply function, $S^r(p)$, which is the quantity supplied to this country at price p. Because the country buys only that part of the world supply, $S(p)$, that is not consumed by any *other* demander elsewhere in the world, $D^o(p)$, its residual supply function is

$$S^r(p) = S(p) - D^o(p). \tag{8.6}$$

At prices so low that $D^o(p)$ is greater than $S(p)$, the residual supply, $S^r(p)$, is zero.

[17] *Jargon alert*: It is traditional to use the expression *excess supply* when discussing international trade and *residual supply* otherwise, though the terms are equivalent.

In Figure 8.12, we derive Japan's residual supply curve for cotton in panel a using the world supply curve, S, and the demand curve of the rest of the world, D^o, in panel b. The scales differ for the quantity axes in the two panels. At a price of \$850 per metric ton, the demand in other countries exhausts world supply (D^o intersects S at 32 million metric tons per year), so Japan has no residual supply. At a much higher price, \$935, Japan's excess supply, 4 million metric tons, is the difference between the world supply, 34 million tons, and the quantity demanded elsewhere, 30 million tons. As the figure illustrates, the residual supply curve facing Japan is much closer to horizontal than is the world supply curve.

The elasticity of residual supply, η_r, facing a given country is (by a similar argument to that in Appendix 8A)

$$\eta_r = \frac{\eta}{\theta} - \frac{1-\theta}{\theta}\;{}_o, \tag{8.7}$$

where η is the market supply elasticity, ${}_o$ is the demand elasticity of the other countries, and $\theta = Q_r/Q$ is the importing country's share of the world's output.

If a country imports a small fraction of the world's supply, we expect it to face a nearly perfectly elastic, horizontal residual supply curve. On the other hand, a relatively large consumer of the good might face an upward-sloping residual supply curve.

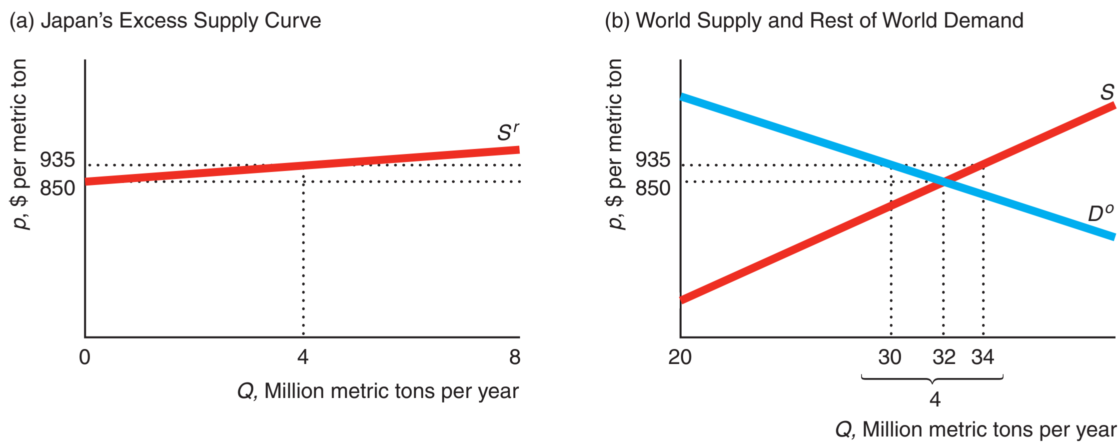
We can illustrate this difference for cotton, where $\eta = 0.5$ and ${}_o = -0.7$ (Green et al., 2005), which is virtually equal to ${}_o$. The United States imports $\theta = 0.1\%$ of the world's cotton, so its residual supply elasticity is

$$\begin{aligned} \eta_r &= \frac{\eta}{0.001} - \frac{0.999}{0.001}\;{}_\mathrm{o} \\ &= 1{,}000\eta - 999\;{}_\mathrm{o} \\ &= (1{,}000 \times 0.5) - (999 \times [-0.7]) = 1{,}199.3, \end{aligned}$$

which is 2,398.6 times more elastic than the world's supply elasticity. Canada's import share is 10 times larger, $\theta = 1\%$, so its residual supply elasticity is "only"

Figure 8.12 Excess or Residual Supply Curve

Japan's excess supply curve, S^r, for cotton is the horizontal difference between the world's supply curve, S, and the demand curve of the other countries in the world, D^o.



119.3. Nonetheless, its residual supply curve is nearly horizontal: A 1% increase in its price would induce imports to more than double, rising by 119.3%. Even Japan's $\theta = 2.5\%$ leads to a relatively elastic $\eta_r = 46.4$. In contrast, China imports 18.5% of the world's cotton, so its residual supply elasticity is 5.8. Even though its residual supply elasticity is more than 11 times larger than the world's elasticity, it is still small enough that its excess supply curve is upward sloping.

Thus, if a country imports a small share of the world's output, then it faces a horizontal import supply curve at the world equilibrium price. If its domestic supply curve is everywhere above the world price, then it only imports and faces a horizontal supply curve. If some portion of its upward-sloping domestic supply curve is below the world price, then its total supply curve is the upward-sloping domestic supply curve up to the world price, and then is horizontal at the world price (Chapter 9 shows such a supply curve for oil).

This analysis of trade applies to trade within a country too. The following Application shows that it can be used to look at trade across geographic areas or jurisdictions such as states.

Application

Reformulated Gasoline Supply Curves

You can't buy the gasoline sold in Milwaukee in other parts of Wisconsin. Houston gas isn't the same as western Texas gas. California, Minnesota, Nevada, and most of America's biggest cities use one or more of at least 46 specialized blends (sometimes called *boutique fuels*), while much of the rest of the country uses regular gas. The U.S. Clean Air Act Amendments, state laws, and local ordinances in areas with serious pollution problems require special, more highly refined blends that cut air pollution. For example, the objective of the federal Reformulated Fuels Program (RFG) is to reduce ground-level ozone-forming pollutants. It specifies both content criteria (such as benzene content limits) and emissions-based performance standards for refiners.

In states in which regular gasoline is used, wholesalers in one state ship gasoline to neighboring states with even slightly higher prices. Consequently, the residual supply curve for regular gasoline for a given state is close to horizontal.

In contrast, jurisdictions that require special blend rarely import gasoline. Few refiners produce any given special blend. Only one Wisconsin refinery produces Milwaukee's special low-polluting blend of gasoline. Because refineries require expensive upgrades to produce a new kind of gas, they generally do not switch from producing one type of gas to another type. Thus, even if the price of gasoline rises in Milwaukee, wholesalers in other states do not send gasoline to Milwaukee, because they cannot legally sell regular gasoline there and it would cost too much to start producing the reformulated gasoline.

Consequently, unlike the nearly horizontal residual supply curve for regular gasoline, the reformulated gasoline residual supply curve is eventually upward sloping. At relatively small quantities, refineries can produce more gasoline without incurring higher costs, so the supply curve in this region is relatively flat. However, to produce much larger quantities of gasoline, when the quantity demanded increases during warm months, refiners must run their plants around the clock and convert a larger fraction of oil into gasoline, incurring higher costs of production. Because of this higher cost, they are willing to sell larger quantities in this range only at a higher price, so the supply curve slopes upward. When the refineries reach capacity, no matter how high the price gets, firms cannot produce more gasoline (at least until new refineries go online), so the supply curve becomes vertical.

Milwaukee and five other counties in southeastern Wisconsin use reformulated gasoline during warm months, while the rest of Wisconsin uses regular gasoline. At the beginning of spring, when the refinery starts switching to cleaner-burning reformulated gasoline and consumers drive more, Milwaukee operates in the steeply upward-sloping section of its supply curve. During March 2015, while reformulated gasoline was in particularly short supply, motorists in Milwaukee, Wisconsin, were paying 45¢ or one-fifth more for a gallon of regular than were motorists in Madison, Wisconsin, which uses regular gasoline. Nationally, motorist who use reformulated gasoline paid 17¢ more than those who use regular gasoline in June 2016.

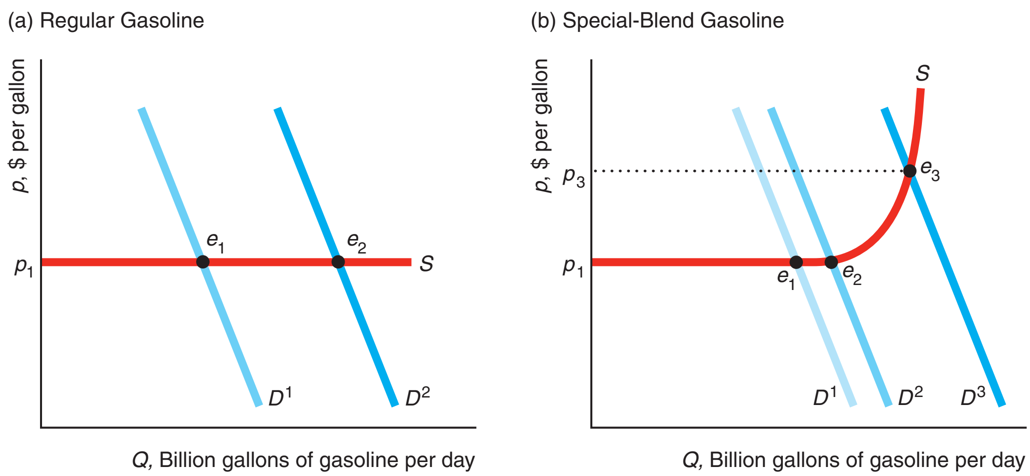
Solved Problem 8.5

MyLab Economics Solved Problem

In the short run, what happens to the competitive market price of gasoline when a state's demand curve shifts to the right during warm months? In your answer, distinguish between areas in which regular gasoline is sold and jurisdictions that require special blends.

Answer

1. *Show the effect of a shift of the demand curve in areas that use regular gasoline.* In an area using regular gasoline, the supply curve is horizontal, as panel a of the figure shows. Thus, as the demand curve shifts to the right from D^1 to D^2, the equilibrium shifts along the supply curve from e_1 to e_2, and the price remains at p_1.



2. *Show the effects of both a small and large shift of the demand curve in a jurisdiction that uses a special blend.* The supply curve in panel b is drawn as described in the Application "Reformulated Gasoline Supply Curves." If the demand curve shifts to the right from D^1 to D^2, the price remains unchanged at p_1 because the demand curve continues to intersect the supply curve in the flat region. However, if the demand curve shifts farther to the right to D^3, then the new intersection is in the upward-sloping section of the supply curve and the price increases to p_3. Consequently, unforeseen "jumps" in demand are more likely to cause a *price spike*—a large increase in price—in jurisdictions that use special blends.

Long-Run Competitive Equilibrium

The intersection of the long-run market supply and demand curves determines the long-run competitive equilibrium. With identical firms, free entry and exit, and constant input prices, the long-run competitive market supply is horizontal at minimum long-run average cost, so the equilibrium price equals long-run average cost. A shift in the demand curve affects only the equilibrium quantity and not the equilibrium price, which remains constant at minimum long-run average cost.

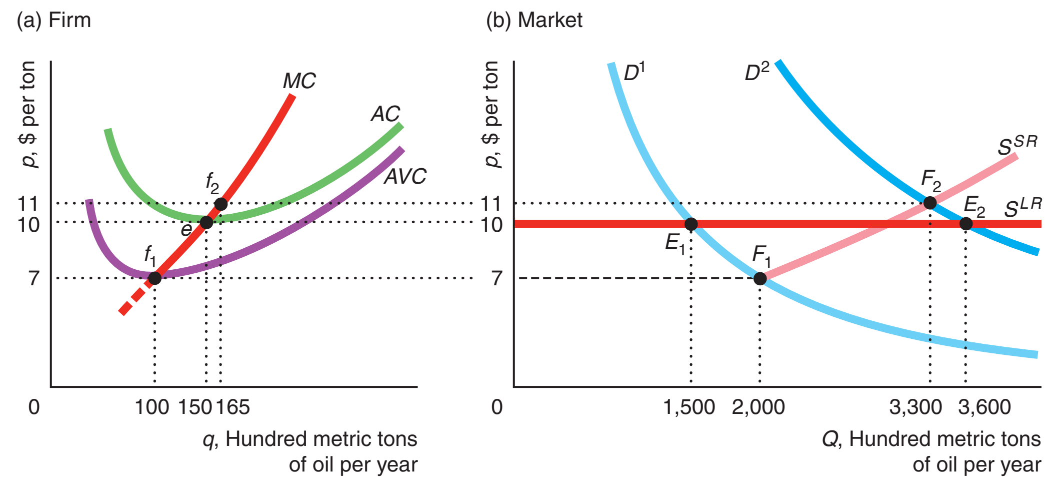
The market supply curve is different in the short run than in the long run, so the long-run competitive equilibrium differs from the short-run equilibrium. The relationship between the short- and long-run equilibria depends on where the market demand curve crosses the short- and long-run market supply curves. Figure 8.13 illustrates this point using the short- and long-run supply curves for the vegetable oil mill market.

The short-run supply curve for a typical firm in panel a is the marginal cost above the minimum of the average variable cost, \$7. At a price of \$7, each firm produces 100 units, so the 20 firms in the market in the short run collectively supply 2,000 ($= 20 \times 100$) units of oil in panel b. At higher prices, the short-run market supply curve slopes upward because it is the horizontal summation of the firm's upward-sloping marginal cost curves.

We assume that the firms use the same size plant in the short and long run so that the minimum average cost is \$10 in both the short and long run. Because all firms have the same costs and can enter freely, the long-run market supply curve is flat at the minimum average cost, \$10, in panel b. At prices between \$7 and \$10, firms supply goods at a loss in the short run but not in the long run.

Figure 8.13 The Short-Run and Long-Run Equilibria for Vegetable Oil MyLab Economics Video

(a) A typical vegetable oil mill produces where price equals its MC, so it is willing to produce 150 units of oil at a price of \$10 or 165 units at \$11. (b) The short-run market supply curve, S^{SR}, is the horizontal sum of 20 individual firms' short-run marginal cost curves above minimum average variable cost, \$7. The long-run market supply curve, S^{LR}, is horizontal at the minimum average cost, \$10. If the demand curve is D^1, in the short-run equilibrium, F_1, 20 firms sell 2,000 units of oil at \$7. In the long-run equilibrium, E_1, 10 firms sell 1,500 units at \$10. If demand is D^2, the short-run equilibrium is F_2 (\$11; 3,300 units; 20 firms) and the long-run equilibrium is E_2 (\$10; 3,600 units; 24 firms).



If the market demand curve is D^1, the short-run market equilibrium, F_1, is below and to the right of the long-run market equilibrium, E_1. This relationship is reversed if the market demand curve is D^2.[18]

In the short run, if the demand is as low as D^1, the market price in the short-run equilibrium, F_1, is \$7. At that price, each of the 20 firms produces 100 units, at f_1 in panel a. The firms lose money because the price of \$7 is below average cost at 100 units. These losses drive some of the firms out of the market in the long run, so market output falls and the market price rises. In the long-run equilibrium, E_1, price is \$10, and each firm produces 150 units, *e*, and breaks even. As the market demands only 1,500 units, only 10 (= 1,500/150) firms produce, so half the firms that produced in the short run exit the market.[19] Thus, with the D^1 demand curve, price rises and output falls in the long run.

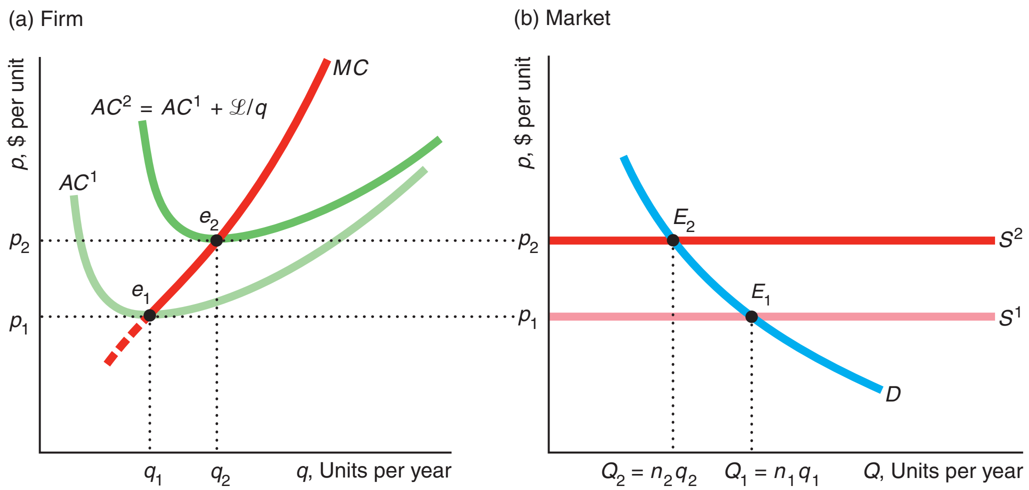
If demand expands to D^2, in the short run, each of the 20 firms expands its output to 165 units, f_2, and the price rises to \$11, where the firms make profits: The price of \$11 is above the average cost at 165 units. These profits attract entry in the long run, and the price falls. In the long-run equilibrium, each firm produces 150 units, *e*, and 3,600 units are sold by the market, E_2, by 24 (= 3,600/150) firms. Thus, with the D^2 demand curve, price falls and output rises in the long run.

Challenge Solution

The Rising Cost of Keeping On Truckin'

We return to the Challenge questions about the effects of higher annual fees and other lump-sum costs on the trucking market price and quantity, the output of individual firms, and the number of trucking firms (assuming that the demand curve remains constant). Because firms may enter and exit this industry in the long run, such higher lump-sum costs can have a counterintuitive effect on the competitive equilibrium.

All trucks of a certain size are essentially identical, and trucks can easily enter and exit the industry (government regulations aside). A typical firm's cost curves are shown in panel a and the market equilibrium is shown in panel b of the figure.



[18]Using data from Statistics Canada, I estimate that the elasticity of demand for vegetable oil is = −0.8. Both D^1 and D^2 are constant elasticity demand curves, but the quantity demanded at any price on D^2 is 2.4 times that on D^1.

[19]How do we know which firms leave? If the firms are identical, the theory says nothing about which ones leave and which ones stay. The firms that leave make zero economic profit, and those that stay also make zero economic profit, so firms are indifferent as to whether they stay or exit.

The new, higher fees and other lump-sum costs raise the fixed cost of operating by $\mathcal{L}$. In panel a, a lump-sum franchise tax shifts the typical firm's average cost curve upward from AC^1 to $AC^2 = AC^1 + \mathcal{L}/q$ but does not affect the marginal cost (see the answer to Solved Problem 7.2). As a result, the minimum average cost rises from e_1 to e_2.

Given that an unlimited number of identical truckers are willing to operate in this market, the long-run market supply curve is horizontal at the minimum average cost. Thus, the market supply curve shifts upward in panel b by the same amount as the minimum average cost increases. Given a downward-sloping market demand curve D, the new equilibrium, E_2, has a lower quantity, $Q_2 < Q_1$, and higher price, $p_2 > p_1$, than the original equilibrium, E_1.

As the market price rises, the quantity that a firm produces rises from q_1 to q_2 in panel a. Because the marginal cost curve is upward sloping at the original equilibrium, when the average cost curve shifts up due to the higher fixed cost, the new minimum point on the average cost curve corresponds to a larger output than in the original equilibrium. Thus, any trucking firm still operating in the market produces at a larger volume.

Because the market quantity falls but each firm remaining in the market produces more, the number of firms in the market must fall. At the initial equilibrium, the number of firms was $n_1 = Q_1/q_1$. The new equilibrium number of firms, $n_2 = Q_2/q_2$, must be smaller than n_1 because $Q_2 < Q_1$ and $q_2 > q_1$. Therefore, an increase in fixed cost causes the market price to rise and the market quantity and the number of trucking firms to fall, as most people would have expected, but it has the surprising effect that it causes producing firms to increase the amount of services that they provide.

Summary

1. **Perfect Competition.** Perfect competition is a market structure in which buyers and sellers are price takers. Each firm faces a horizontal demand curve. A firm's demand curve is horizontal because perfectly competitive markets have five characteristics: a large number of small buyers and sellers, firms produce identical (homogeneous) products, buyers have full information about product prices and characteristics, transaction costs are negligible, and firms can enter and exit freely in the long run. Many markets are highly competitive—firms are very close to being price takers—even if they do not strictly possess all five of the characteristics associated with perfect competition.
2. **Profit Maximization.** Most firms maximize economic profit, which is revenue minus economic cost (explicit and implicit cost). Because business profit, which is revenue minus only explicit cost, does not include implicit cost, economic profit tends to be less than business profit. A firm earning zero economic profit is making as much as it could if its resources were devoted to their best alternative uses. To maximize profit, all firms (not just competitive firms) must make two decisions. First, the firm determines the quantity at which its profit is highest. The firm maximizes its profit when marginal profit is zero or, equivalently, when marginal revenue equals marginal cost. Second, the firm decides whether to produce at all.
3. **Competition in the Short Run.** To maximize its profit, a competitive firm (like a firm in any other market structure) chooses its output level where marginal revenue equals marginal cost. Because a competitive firm is a price taker, its marginal revenue equals the market price, so it sets its output so that price equals marginal cost. New firms cannot enter in the short run. In addition, firms have some sunk fixed inputs. In this sense, firms cannot exit the industry in the short run. However, a profit-maximizing firm shuts down and produces no output if the market price is less than its minimum average variable cost. Thus, a competitive firm's short-run supply curve is its marginal cost

curve above its minimum average variable cost. The short-run market supply curve is the sum of the supply curves of the fixed number of firms producing in the short run. The intersection of the market demand curve and the short-run market supply curve determines the short-run competitive equilibrium.

4. **Competition in the Long Run.** In the long run, a competitive firm sets its output where the market price equals its long-run marginal cost. It shuts down if the market price is less than the minimum of its long-run average cost, because all costs are variable in the long run. Consequently, the competitive firm's supply curve is its long-run marginal cost above its minimum long-run average cost. The long-run supply curve of a firm may have a different slope than the short-run curve because the firm can vary its fixed factors in the long run. The long-run market supply curve is the horizontal sum of the supply curves of all the firms in the market. If all firms are identical, entry and exit are easy, and input prices are constant, the long-run market supply curve is flat at minimum average cost. If firms differ, entry is difficult or costly, or input prices vary with output, the long-run market supply curve has an upward slope. The long-run market supply curve slopes upward if input prices increase with output and slopes downward if input prices decrease with output. The long-run market equilibrium price and quantity are different from the short-run price and quantity.

Questions

Select questions are available on MyLab Economics*; * = answer appears at the back of this book;* **A** *= algebra problem;* **C** *= calculus problem.*

1. Perfect Competition

1.1 A large city has nearly 500 restaurants, with new ones entering regularly as the population grows. The city decides to limit the number of restaurant licenses to 500. Which characteristics of this market are consistent with perfect competition and which are not? Is this restaurant market likely to be nearly perfectly competitive? Explain your answer.

1.2 Why would high transaction costs or imperfect information tend to prevent price-taking behavior?

1.3 Based on Roberts and Schlenker (2013), the corn demand elasticity is $= -0.3$, and the supply elasticity is $\eta = 0.15$. According to the 2007 Census of Agriculture, the United States has 347,760 corn farms. Assuming that the farms are of roughly equal size, what is the elasticity of demand facing a single farm? (*Hint*: See Solved Problem 8.1.)

1.4 Based on Equation 8.2, by how much does the residual elasticity of demand facing a firm increase as the number of firms increases by one firm? (*Hint*: See Solved Problem 8.1.) **A** or **C**

2. Profit Maximization

2.1 Should a competitive firm ever produce when it is losing money (making a negative economic profit)? Why or why not?

2.2 Should a firm shut down (and why) if its revenue is $R = \$1{,}000$ per week,

a. its variable cost is $VC = \$500$, and its sunk fixed cost is $F = \$600$?

b. its variable cost is $VC = \$1{,}001$, and its sunk fixed cost $F = \$500$?

2.3 Should a firm shut down if its weekly revenue is \$1,000, its variable cost is \$500, and its fixed cost is \$800, of which \$600 is avoidable if it shuts down? Why?

*2.4 A firm's profit function is $\pi(q) = R(q) - C(q) = 120q - (200 + 40q + 10q^2)$. What is the positive output level that maximizes the firm's profit (or minimizes its loss)? What are the firm's revenue, variable cost, and profit? Should it operate or shut down in the short run? **C**

3. Competition in the Short Run

*3.1 A marginal cost curve may be U-shaped. As a result, the MC curve may hit the firm's demand curve or price line at two output levels. Which is the profit-maximizing output? Why?

3.2 If the cost function for John's Shoe Repair is $C(q) = 100 + 10q - q^2 + \frac{1}{3}q^3$, and its marginal cost function is $MC = 10 - 2q + q^2$, what is its profit-maximizing condition given that the market price is p? **A**

*3.3 If a competitive firm's cost function is $C(q) = a + bq + cq^2 + dq^3$, where a, b, c, and d are constants, what is the firm's marginal cost function? What is the firm's profit-maximizing condition? **C**

3.4 The cost function for Acme Laundry is $C(q) = 10 + 10q + q^2$, so its marginal cost function is $MC = 10 + 2q$, where q is tons of laundry cleaned. Derive the firm's average cost and average variable cost curves. What q should the firm choose so as to maximize its profit if the market price is p? How much does it produce if the competitive market price is $p = 50$? **A**

3.5 If a specific subsidy (negative tax) of s is given to only one competitive firm, how should that firm change its output level to maximize its profit, and how does its maximum profit change? (*Hint*: See Solved Problem 8.2.)

3.6 How would the answer to Solved Problem 8.2 change if, instead of a specific tax, Manitoba imposes an ad valorem tax (see Chapter 3) of v percent on lime produced in that province? (*Hints*: What role does the market price play in the analyses of the two types of tax?)

3.7 Beta Laundry's pre-tax cost function is $C(q) = 30 + 20q + q^2$, so its marginal cost function is $MC = 20 + 2q$.

a. What quantity maximizes the firm's profit if the market price is p? How much does it produce if $p = 60$?

b. If the government imposes a specific tax of $t = 2$, what quantity maximizes its after-tax profit? Does it operate or shut down? (*Hint*: See Solved Problem 8.2.) **A**

3.8 If the pre-tax cost function for John's Shoe Repair is $C(q) = 100 + 10q - q^2 + \frac{1}{3}q^3$, and it faces a specific tax of $t = 10$, what condition determines the profit-maximizing output if the market price is p? Can you solve for a single, profit-maximizing q in terms of p? (*Hint*: See Solved Problem 8.2.) **C**

3.9 Initially, the market price was $p = 20$, and the competitive firm's minimum average variable cost was 18, while its minimum average cost was 21. Should it shut down? Why? Now this firm's average variable cost increases by 3 at every quantity, while other firms in the market are unaffected. What happens to its average cost? Should this firm shut down? Why?

3.10 It is generally accepted that one advantage of an online store is lower overhead costs (expenses not directly related to production such as rent, utilities, and insurance). How would a move to e-commerce affect a competitive firm's pricing and production decisions and its profits in the short run? (*Hint*: See Solved Problem 8.3.)

3.11 According to the Application "Fracking and Shutdowns," conventional oil wells have lower shutdown points than those that use fracking. Use figures to compare the supply curves of firms with conventional wells and those that use fracking. On your figures, show the shutdown points and label the relevant costs on the vertical axes.

3.12 In June 2014, the international price of crude oil was about \$114 per barrel (Brent benchmark). Three years later, it was about \$47 per barrel. This dramatic price reduction caused petroleum companies to write down the value of their high-cost Canadian oil sands reserves and some companies to sell off their oil sands assets. Regardless, oil sands production has continued to grow. If the price at which a competitive oil sands mine becomes uneconomic is \$60 per barrel but the mine continues to produce, use a diagram to show what this implies about the firm's short-run cost curves. Also use your diagram to indicate what the firm's profit was in June 2014 and June 2017.

*3.13 Norway is the world's second largest exporter of fish and fish products by value, with key markets in Europe, Russia, Japan, China, and the United States. Haddock is one of the main capture species. However, the price of haddock fluctuates widely. For example, it was about 60% lower in March 2016 than its peak in January 2015, and exports of fresh haddock products declined by 7% between 2015 and 2016. If some Norwegian fishermen stay in port when prices fall dramatically while others continue to harvest haddock, what might explain their differing behavior?

3.14 According to an article in the *Financial Times* on March 8, 2016, the capacity of the biggest container ships afloat rose sharply between 2011 and 2016 and has more than doubled since 2000. What is the economic explanation for this behavior? Use a diagram to show what this implies about the firm's cost curves in 2000 and 2016.

3.15 The Internet is affecting holiday shipping. In years past, the busiest shipping period was Thanksgiving week. Now as people have become comfortable with e-commerce, they purchase later in the year and are more likely to have gifts shipped (rather than purchasing locally). FedEx, along with Amazon and other e-commerce firms, hires extra workers during this period, and many regular workers log substantial overtime hours.

a. Are a firm's marginal and average costs likely to rise or fall with this extra business? (Discuss economies of scale and the slopes of marginal and average cost curves.)

b. Use side-by-side firm-market diagrams to show the effects on the number of firms, equilibrium price and output, and profits of such a seasonal shift in demand for e-retailers in both the short run and the long run. Explain your reasoning.

3.16 About two-third of the world's wine is produced in Europe. France, Italy, and Spain account for almost half of all the wine produced in the world. Except for France, none of the top 10 wine-producing countries in Europe levy an excise (per-unit) tax on still wine, and France's excise tax is very small (about €0.038 per liter). If the wine industry were perfectly competitive, explain what effect a significant excise tax in France would have on the long-run supply curve for wine from those two countries when the entry of firms is limited. (*Hint*: See Solved Problem 8.4 and the Application "Upward-Sloping Long-Run Supply Curve for Cotton.")

3.17 What is the effect on the short-run equilibrium of a specific subsidy of s per unit that is given to all n firms in a market? What is the incidence of the subsidy? (*Hint*: See Solved Problem 8.4.)

3.18 As of October 15, 2015, large stores (with 250 or more full-time employees) in England are required to charge shoppers 5 pence per new plastic bag. Plastic bags at airport shops or aboard trains, planes, and ships are not included, and neither are paper bags. An article in *The Guardian* indicated that the number of single-use plastic bags handed out by the seven largest supermarkets fell by more than 85% in the first six months of the charge, more than £29 million was raised for charities and community groups, and more people were bringing their own bags when shopping (Barkham, 2016). Does such a charge affect the marginal cost of any particular good? If so, by how much? Is this charge likely to affect the overall amount that consumers pay for the goods they buy?

3.19 Each of the 10 firms in a competitive market has a cost function of $C = 25 + q^2$, so its marginal cost is $MC = 2q$. The market demand function is $Q = 120 - p$. Determine the equilibrium price, quantity per firm, and market quantity. **A**

3.20 A shock causes the demand curve to shift to the right. What properties of the market are likely to lead to a large increase in the equilibrium price? (*Hint*: See the discussion of the shape of the market supply curve and Solved Problem 8.5.)

4. Competition in the Long Run

4.1 In June 2005, Eastman Kodak announced that it no longer would produce black-and-white photographic paper—the type used to develop photographs by a traditional darkroom process. Kodak based its decision on the substitution of digital photography for traditional photography. In making its exit decision, did Kodak compare the price of its paper and average variable cost (at its optimal output)? Alternatively, did Kodak compare the price of its paper and average total cost (again at its optimal output)?

4.2 Redraw Figure 8.9 showing a situation in which the short-run plant size is too large relative to the optimal long-run plant size.

*4.3 In Brazil, layoffs are generally treated as terminations without cause, for which firms are required to provide 30 days' notice plus an additional three days' notice for every year the employee has worked for the company. In addition, employees are entitled to their wages, holiday pay, and access to severance funds. In addition, Brazilian Labor Courts have held that the employer should also inform and negotiate the conditions of any collective redundancy with the employees' union beforehand so that additional severance payments are possible. What is the effect on the firm and market equilibrium of these labor laws?

4.4 Each firm in a competitive market has a cost function of $C = q^2$, so its marginal cost function is $MC = 2q$. The market demand function is $Q = 24 - p$. Determine the long-run equilibrium price, quantity per firm, market quantity, and number of firms. **A**

4.5 The iPhone debuted in January 2007. In the 10 years since then, smartphones have revolutionized and enriched our lives in many ways, and they are continuing to do so. According to the International Telecommunications Union, a specialized agency of the United Nations for information and communication technologies, there were more than 7 billion mobile cellular subscriptions at the end of 2015, a little less than the world's population at that time (ITU, n.d.). Among many other things, smartphones have dramatically increased access to information at a relatively low cost. Use side-by-side diagrams to illustrate how this technological innovation affected the long-run supply curve and the equilibrium in the market for information.

4.6 Draw a figure to illustrate why the size of ethanol processing plants has fallen in recent years (see "The Size of Ethanol Processing Plants" Application).

4.7 The "Upward-Sloping Long-Run Supply Curve for Cotton" Application shows a supply curve for cotton. Discuss the equilibrium if the world demand curve crosses this supply curve in either (a) the flat section labeled Brazil or (b) the vertical section to the right of that. What do farms in the United States do?

4.8 Redraw Figure 8.11 to show what happens if factor costs fall as the industry's quantity increases.

*4.9 Derive Equation 8.7. (*Hint*: Use a method similar to that used in Appendix 8A.) **C**

*4.10 According to the European Commission's Weekly Oil Bulletin for June 12, 2017, taxes made up 66.2% of the €1.51 per liter consumer price of gasoline in Italy. Included in the tax total is an excise tax of €0.728 per liter levied by the national government, plus excise taxes applied in certain regions; for example, the additional tax in Molise is €0.031 per liter. Suppose that the national and regional governments each increased their excise taxes on gasoline by 1 cent and that the former increases the retail price of gasoline by 0.5 cents, while the latter increases it by 1 cent. What is the incidence of each of these taxes on consumers and producers? Explain why the incidence on consumers in a competitive market would be lower for an increase in the national tax than for an increase in a regional tax. How much more elastic is the supply elasticity for any one of the 20 regions in Italy compared to the national supply elasticity? (*Hint*: Use Equations 3.12 and 8.7 in answering this question.) **A**

4.11 On January 28, 2017, an empty LPG tanker leaving the Kamarajar Port in the State of Tamil Nadu, India, collided with an inbound tanker carrying 45,000 metric tons of oil, causing a spill of between 20 and 60 metric tons. A physical barrier to keep the oil contained was not effective, and the spill spread about 34 kilometers to the south, affecting marine life and beaches in the area. Over 1,000 personnel from several government agencies and many volunteers helped in the cleanup efforts. Use diagrams of a competitive market for cleanup services and of a representative firm providing those services to show the effects of an increase in demand on price, output, profit, and the number of firms in both the short and long run, assuming a horizontal long-run supply curve. Explain how your answer depends on whether the shift in demand is of a short or long duration.

4.12 A 2016 survey by OpinionWay for kayak.co.uk found that the use of travel agents has changed dramatically over the past 10 years in the United Kingdom (Sullivan, 2016). While 57% of respondents indicated that they booked through a travel agency a decade ago, only 19% do so today. Instead, 79% of respondents are booking their trips online today (up from 35% a decade ago). Technology and new business models have significantly reduced the cost of doing it oneself, and many travel agencies are going out of business. Use diagrams of a competitive market for travel services and of a representative travel agency providing those services to explain what happened in the market for travel services, assuming there is no change in the demand for travel services.

5. Challenge

5.1 In the Challenge Solution, would it make a difference to the analysis whether lump-sum costs such as registration fees are collected annually or only once when the firm starts operation? How would each of these franchise taxes affect the firm's long-run supply curve? Explain your answer.

5.2 Change the answer given in the Challenge Solution for the short run rather than for the long run. (*Hint*: The answer depends on where the demand curve intersects the original short-run supply curve.)

5.3 Fixed costs can significantly affect the profitability of a small business. One way to reduce fixed costs might be to outsource the company's sales and marketing functions to third parties. For example, a third party could schedule appointments with prospective clients and track progress, and web development might be undertaken by qualified students hired as the need arises. How would the market price and quantity be affected if all firms face the same fixed costs as opposed to a situation where one group has higher fixed costs than the other? Assume a competitive market with a horizontal long-run supply curve.

*5.4 Uber, a car-for-hire service that connects passengers with private drivers, presently operates in 613 cities worldwide. Lower overheads allow Uber to charge significantly less than a taxi company would charge for an equivalent ride, due in part to lower or nonexistent city licence fees. Taxi companies argue that this advantage for Uber and other ride-booking companies has adversely affected their businesses. If the market for car-for-hire services was perfectly competitive, what effect did the entry of Uber have? What would be the effect on taxi companies of reducing the annual licencing fee levied by the local authority?

5.5 Question 3.16 asked about the effect of imposing a specific tax on a perfectly competitive industry when entry into the industry is limited. If entry were instead unlimited, how would the market and firm equilibria change in the long run when a specific tax is imposed?

Applying the Competitive Model

9

No more good must be attempted than the public can bear. —Thomas Jefferson

Challenge

Liquor Licenses

After you graduate, do you want to open a restaurant that serves drinks? If so, you'll need a liquor license, which, in some states, costs a lot of money.

Seventeen states and Washington, D.C., limit the number of liquor licenses.[1] Massachusetts issues one license per 2,000 residents; New Jersey, one per 3,000; and Utah, one per 4,925. In these limit or quota states, buying a liquor license from the state only costs a few hundred dollars. However, you probably can't get one from the state. You'll have to buy one from someone who already has one. Buying a license will cost you as much as $200,000 in Philadelphia suburbs; $450,000 in Massachusetts; $1 million in Montana, New Mexico, or Utah; and $1.6 million in parts of New Jersey.

What effect does setting a limit on the number of liquor licenses have on the price of meals (including liquor)? What determines the value of a license? How much profit beyond the cost of the license can a restaurant earn? Who benefits and who loses from limiting the number of liquor licenses?

In this chapter, we illustrate how to use the competitive market model to answer these types of questions. One of the major strengths of the competitive model is that it can predict how changes in government policies such as those concerning licensing, trade tariffs (taxes on imports), import quotas, global warming, and major cost-saving discoveries affect consumers and producers. We start by examining the properties of a competitive market and then consider how government actions and other shocks affect the market and its properties.

We concentrate on two main properties of a competitive market. First, firms in a competitive equilibrium generally make zero (economic) profit in the long run. Second, competition maximizes a measure of societal welfare.

To most people, the term *welfare* refers to the government's payments to poor people. However, economists imply no such meaning when they use the term. Economists use *welfare* to refer to the well-being of society. This welfare measure is the sum of

[1]Alaska, Arizona, California, Florida, Idaho, Kentucky, Massachusetts, Michigan, Minnesota, Montana, New Jersey, New Mexico, Ohio, Pennsylvania, South Dakota, Utah, and Washington.

a measure of consumer welfare called *consumer surplus* and producer welfare called *producer surplus*. Economists call an analysis of the impact of a change on various groups' well-being a study of *welfare economics*.

By predicting the effects of a proposed policy on consumer surplus, producer surplus, and welfare, economists can advise policymakers as to who will benefit, who will lose, and what the net effect of this policy will likely be. To decide whether to adopt a particular policy, policymakers may combine these predictions with their normative views (values), such as whether they are more interested in helping the group that gains or the group that loses.

In this chapter, we examine seven main topics

1. **Zero Profit for Competitive Firms in the Long Run.** In long-run competitive market equilibrium, profit-maximizing firms break even, so firms that do not try to maximize profit lose money and leave the market.
2. **Consumer Welfare.** How much consumers are helped or harmed by a change in the equilibrium price can be measured by using information from demand curves or utility functions.
3. **Producer Welfare.** How much producers gain or lose from a change in the equilibrium price can be measured by using information from the marginal cost curve or by measuring the change in profits.
4. **Competition Maximizes Welfare.** Competition maximizes a measure of social welfare based on consumer and producer welfare.
5. **Policies That Shift Supply and Demand Curves.** Government policies that shift the supply or demand curves in perfectly competitive markets harm consumers and lower welfare.
6. **Policies That Create a Wedge Between Supply and Demand.** Government policies such as taxes, price ceilings, price floors, and tariffs that create a wedge between the supply and demand curves reduce the equilibrium quantity, raise the equilibrium price to consumers, and lower welfare.
7. **Comparing Both Types of Policies: Imports.** Policies that limit supply (such as quotas or bans on imports) or create a wedge between supply and demand (such as *tariffs*) have different welfare effects when both policies reduce imports by equal amounts.

9.1 Zero Profit for Competitive Firms in the Long Run

Competitive firms earn zero profit in the long run whether or not entry is completely free. Consequently, competitive firms must maximize profit.

Zero Long-Run Profit with Free Entry

The long-run supply curve is horizontal if firms are free to enter the market, if firms have identical cost, and if input prices are constant. All firms in the market are operating at minimum long-run average cost. That is, they are indifferent between shutting down or not because they are earning zero profit.

One implication of the shutdown rule (Chapter 8) is that the firm is willing to operate in the long run even if it is making zero *economic* profit, which is revenue minus opportunity cost. Because opportunity cost includes the value of the next best investment, at a zero long-run economic profit, the firm is earning the normal business profit that it could earn by investing elsewhere in the economy.

For example, if a firm's owner had not built the plant the firm uses to produce, the owner could have spent that money on another business or put the money in a bank. The opportunity cost of the current plant, then, is the forgone profit from what the owner could have earned by investing the money elsewhere.

The five-year after-tax accounting return on capital across all firms was 10.5%, indicating that the typical firm earned a business profit of 10.5¢ for every dollar it invested in capital (*Forbes*). These firms were earning roughly zero economic profit but positive business profit. Because business cost does not include all opportunity costs, business profit is larger than economic profit. Thus, a profit-maximizing firm may stay in business if it earns zero long-run economic profit, but it shuts down if it earns zero long-run business profit.

Zero Long-Run Profit When Entry Is Limited

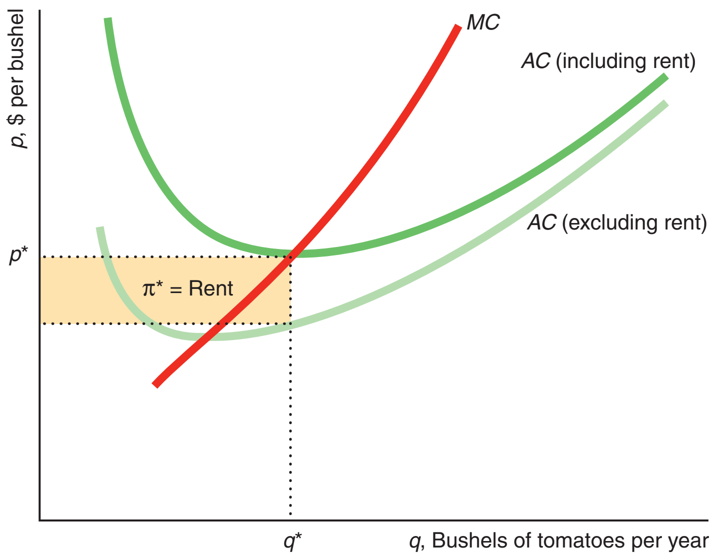
In some markets, firms cannot enter in response to long-run profit opportunities. One reason for the limited number of firms is that the supply of an input is limited: only so much land is suitable for mining uranium; only a few people have the superior skills needed to play professional basketball.

One might think that firms could make positive long-run economic profits in such markets; however, that's not true. The reason why firms earn zero economic profits is that firms bidding for the scarce input drive its price up until the firms' profits are zero.

Suppose that the number of acres suitable for growing tomatoes is limited. Figure 9.1 shows a typical farm's average cost curve if the rental cost of land is zero (the average cost curve includes only the farm's costs of labor, capital, materials, and energy—not land). At the market price p^*, the firm produces q^* bushels of tomatoes and makes a profit of π^*, the shaded rectangle in the figure.

Figure 9.1 Rent

If it did not have to pay rent for its land, a farm with high-quality land would earn a positive long-run profit of π^*. Due to competitive bidding for this land, however, the rent equals π^*, so the landlord reaps all the benefits of the superior land and the farmer earns a zero long-run economic profit.



Thus, if the owner of the land does not charge rent, the farmer makes a profit. Unfortunately for the farmer, the landowner rents the land for π^*, so the farmer actually earns zero profit. Why does the landowner charge that much? The reason is that π^* is the opportunity cost of the land: The land is worth π^* to other potential farmers. These farmers will bid against each other to rent this land until the rent is driven up to π^*.

This rent is a fixed cost to the farmer because it doesn't vary with the amount of output. Thus, the rent affects the farm's average cost curve but not its marginal cost curve.

As a result, if the farm produces at all, it produces q^*, where its marginal cost equals the market price, no matter what rent is charged. The higher average cost curve in the figure includes a rent equal to π^*. The minimum point of this average cost curve is p^* at q^* bushels of tomatoes, so the farmer earns zero economic profit.

If the demand curve shifts to the left so that the market price falls, the farmer suffers short-run losses. In the long run, the rental price of the land will fall enough that once again each farm earns zero economic profit.

Does it make a difference whether farmers own or rent the land? Not really. The opportunity cost to a farmer who owns superior land is the amount for which that land could be rented in a competitive land market. Thus, the economic profit of both owned and rented land is zero at the long-run equilibrium.

Good-quality land is not the only scarce resource. The price of any fixed factor will be bid up in a similar fashion until economic profit for the firm is zero in the long run. Similarly, the government may require that a firm have a license to operate and then limit the number of licenses available. The price of the license gets bid up by potential entrants, driving profit to zero. For example, the license fee was \$643,000 a year for a hot dog stand next to the steps of the Metropolitan Museum of Art in New York City.[2]

rent
a payment to the owner of an input beyond the minimum necessary for the factor to be supplied

A scarce input, such as a person with high ability or land, earns an extra opportunity value. This extra value is called a **rent**: a payment to the owner of an input beyond the minimum necessary for the factor to be supplied.

Bonnie manages a store for an annual salary of \$40,000, the amount paid to a typical manager. In this highly competitive retail market, firms typically earn zero economic profit. However, because Bonnie is an outstanding manager, her firm earns an economic profit of \$50,000 a year. Other firms, seeing what a good job Bonnie is doing, offer her a higher salary. The bidding for her services drives her salary up to \$90,000: her \$40,000 base salary plus the \$50,000 rent. After paying this rent to Bonnie, the store that employs her makes zero economic profit, just like the other firms in the market.

To summarize, if some firms in a market make short-run economic profits due to a scarce input, the other firms in the market bid for that input. This bidding drives up the price of the factor until all firms earn zero long-run profits. In such a market, the supply curve is flat because all firms have the same minimum long-run average cost.

[2]The auction value for this license hit \$643,000 in 2009, but has fallen since then. (In the hot dog stand photo, I'm the fellow in the blue shirt with the dopey expression.) As of 2013, the highest fee in New York City is \$1.39 million a year to operate a hot dog cart outside the former Tavern on the Green restaurant in Central Park.

Application

What's a Name Worth?

People with unusual abilities can earn staggering incomes, which are rents for their abilities. Though no law stops anyone from trying to become a professional entertainer or athlete, most of us do not have so much talent that others will pay to watch us perform.

According to **Forbes.com**, Cristiano Ronaldo earned $88 million in 2015, Taylor Swift $80 million, LeBron James $77 million, Diddy $60 million, Lady Gaga $59 million.

Indeed, the estates of major celebrities continue to collect rents even after they die. People will still pay to listen to their music, view their cartoons, or use their image. In 2015, Michael Jackson's estate earned $115 million, Elvis Presley $55 million, Charles Schulz $40 million, Bob Marley $21 million, and Albert Einstein $11 million.

To put these receipts in perspective, these amounts exceed some small nations' gross domestic product (the value of the country's total output) such as the $37 million for Tuvalu with 10,869 people.

The Need to Maximize Profit

The worst crime against working people is a company which fails to operate at a profit. —Samuel Gompers, first president of the American Federation of Labor

In a competitive market with identical firms and free entry, if most firms are profit-maximizing, profits are driven to zero at the long-run equilibrium. Any firm that does not maximize profit—that is, any firm that sets its output so that its marginal cost exceeds the market price or that fails to use the most cost-efficient methods of production—will lose money. Thus, *to survive in a competitive market, a firm must maximize its profit*.

9.2 Consumer Welfare

Economists and policymakers want to know how much consumers benefit from or are harmed by shocks that affect the equilibrium price and quantity. To what extent are consumers harmed if a local government imposes a sales tax to raise additional revenues? To answer such a question, we need some way to measure consumers' welfare. Economists use measures of welfare based on consumer theory (Chapters 4 and 5).

If we knew a consumer's utility function, we could directly answer the question of how an event affects a consumer's welfare. If the price of beef increases, the budget line facing someone who eats beef rotates inward, so the consumer is on a lower indifference curve at the new equilibrium. If we knew the levels of utility associated with the original indifference curve and the new one, we could measure the impact of a new tax in terms of the change in the utility level.

This approach is not practical for a couple of reasons. First, we rarely, if ever, know individuals' utility functions. Second, even if we had utility measures for various consumers, we would have no obvious way to compare them. One person might say that he got 1,000 utils (units of utility) from the same bundle that another consumer says gives her 872 utils of pleasure. The first person is not necessarily happier—he may just be using a different scale.

As a result, *we measure consumer welfare in terms of dollars*. Instead of asking the rather silly question "How many utils would you lose if your daily commute increased by 15 minutes?" we could ask "How much would you pay to avoid having your daily commute grow a quarter of an hour longer?" or "How much would it cost you in forgone earnings if your daily commute were 15 minutes longer?" It is easier to compare dollars across people than utils.

We first present the most widely used method of measuring consumer welfare. Then we show how it can be used to measure the effect of a change in price on consumer welfare.

Measuring Consumer Welfare Using a Demand Curve

Consumer welfare from a good is the benefit a consumer gets from consuming that good minus what the consumer paid to buy the good. How much pleasure do you get from a good above and beyond its price? If you buy a good for exactly what it's worth to you, you are indifferent between making that transaction and not. Frequently, however, you buy things that are worth more to you than what they cost. Imagine that you've played tennis in the hot sun and are very thirsty. You can buy a soft drink from a vending machine for \$1, but you'd be willing to pay much more because you are so thirsty. As a result, you're much better off making this purchase than not.

If we can measure how much more you'd be willing to pay than you did pay, we'd know how much you gained from this transaction. Luckily for us, the demand curve contains the information we need to make this measurement.

Marginal Willingness to Pay To develop a welfare measure based on the demand curve, we need to know what information is contained in a demand curve. The demand curve reflects a consumer's *marginal willingness to pay*: the maximum amount a consumer will spend for an extra unit of a good. The consumer's marginal willingness to pay is the *marginal value* the consumer places on the last unit of output.

David's demand curve for magazines per week, panel a of Figure 9.2, indicates his marginal willingness to pay for various numbers of magazines. David places a marginal value of \$5 on the first magazine, \$4 on a second magazine, and \$3 on a third one. As a result, if the price of a magazine is \$5, David buys one magazine, point *a* on the demand curve. If the price is \$4, he buys two magazines, *b*. If the price of magazines is \$3, he buys three magazines, *c*.

consumer surplus (*CS*) the monetary difference between what a consumer is willing to pay for the quantity of the good purchased and what the good actually costs

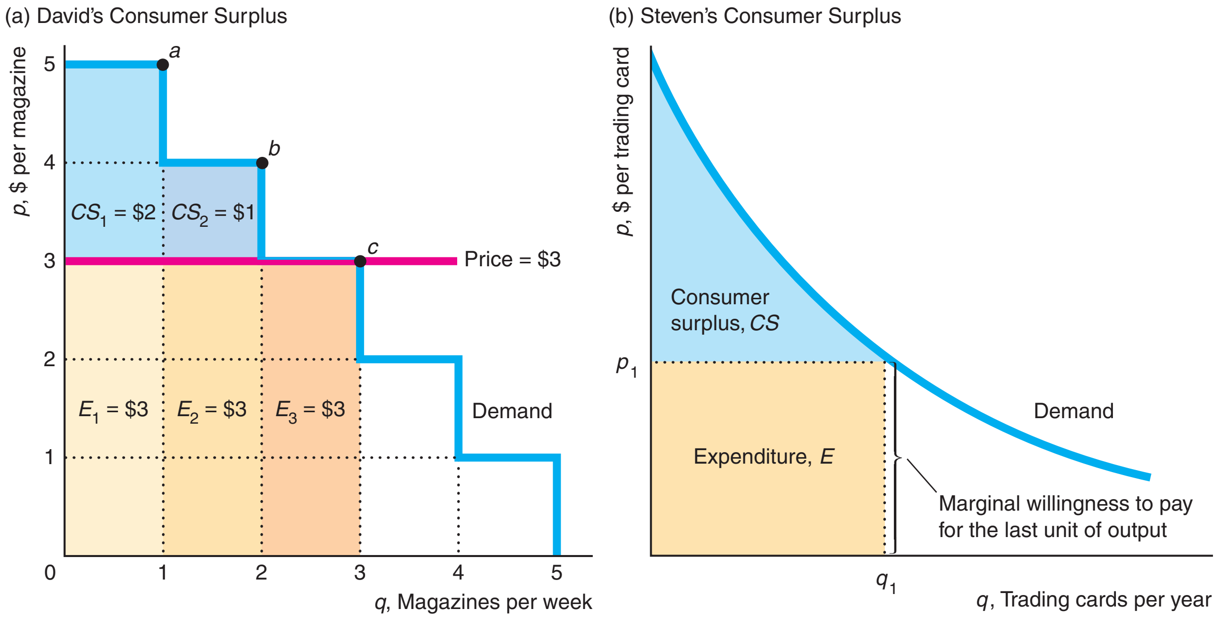
Consumer Surplus The monetary difference between what a consumer is willing to pay for the quantity of the good purchased and what the good actually costs is called **consumer surplus** (*CS*). Consumer surplus is a dollar-value measure of the extra pleasure the consumer receives from the transaction beyond its price.

David's consumer surplus from each additional magazine is his marginal willingness to pay minus what he pays to obtain the magazine.

His marginal willingness to pay for the first magazine, \$5, is area $CS_1 + E_1$ in panel a of Figure 9.2. If the price is \$3, his expenditure on the first magazine is area $E_1 = \$3 \times 1 = \$3$. Thus, his consumer surplus on the first magazine is his marginal willingness to pay for that magazine, CS_1, minus his expenditure, E_1, which is area $CS_1 = (CS_1 + E_1) - E_1 = \$5 - \$3 = \$2$. Because his marginal willingness to pay for the second magazine is \$4, his consumer surplus for the second magazine is the smaller area $CS_2 = \$1$. His marginal willingness to pay for the third magazine is \$3, which equals what he must pay to obtain it, so his consumer surplus is zero, $CS_3 = \$0$ (and hence not shown as an area in the figure). He is indifferent between buying and not buying the third magazine.

Figure 9.2 Consumer Surplus

(a) David's demand curve for magazines has a step-like shape. When the price is \$3, he buys three magazines, point *c*. David's marginal value for the first magazine is \$5, area $CS_1 + E_1$, and his expenditure is \$3, area E_1, so his consumer surplus is $CS_1 = \$2$. His consumer surplus is \$1 for the second magazine, area CS_2, and is \$0 (and hence not labeled) for the third, CS_3—he is indifferent between buying and not buying it. Thus, his total consumer surplus is the shaded area $CS_1 + CS_2 + CS_3 = \$3$. (b) Steven's willingness to pay for trading cards is the height of his smooth demand curve. At price p_1, Steven's expenditure is $E\ (= p_1 q_1)$, his consumer surplus is CS, and the total value he places on consuming q_1 trading cards per year is $CS + E$.



At a price of \$3, David buys three magazines. His total consumer surplus from the three magazines he buys is the sum of the consumer surplus he gets from each of these magazines: $CS_1 + CS_2 + CS_3 = \$2 + \$1 + \$0 = \$3$. This total consumer surplus of \$3 is the extra amount that David is willing to spend for the right to buy three magazines at \$3 each. Thus, *an individual's consumer surplus is the area under the demand curve and above the market price up to the quantity the consumer buys.*

David is unwilling to buy a fourth magazine unless the price drops to \$2 or less. If David's mother gives him a fourth magazine as a gift for which she pays \$3, the marginal value that David puts on that fourth magazine, \$2, is less than what it cost his mother.

We can determine consumer surplus for smooth demand curves in the same way as with David's unusual stair-like demand curve. Steven has a smooth demand curve for baseball trading cards, panel b of Figure 9.2. The height of this demand curve measures his willingness to pay for one more card. This willingness varies with the number of cards he buys in a year. The total value he places on obtaining q_1 cards per year is the area under the demand curve up to q_1, the areas CS and E. Area E is his actual expenditure on q_1 cards. Because the price is p_1, his expenditure is $p_1 q_1$. Steven's consumer surplus from consuming q_1 trading cards is the value of consuming those cards, areas CS and E, minus his actual expenditures, E, to obtain them, or CS.

Thus, his consumer surplus, *CS*, is the area under the demand curve and above the horizontal line at the price p_1 up to the quantity he buys, q_1.

Just as we measure the consumer surplus for an individual using that individual's demand curve, we measure the consumer surplus of all consumers in a market using the market demand curve. *Market consumer surplus is the area under the market demand curve above the market price up to the quantity consumers buy.*

To summarize, consumer surplus is a practical and convenient measure of consumer welfare. Using consumer surplus has two advantages over using utility to discuss consumer welfare. First, the dollar-denominated consumer surplus of several individuals can be easily compared or combined, whereas the utility of various individuals cannot be easily compared or combined. Second, it is relatively easy to measure consumer surplus, whereas it is difficult to get a meaningful measure of utility directly. To calculate consumer surplus, all we have to do is measure the area under a demand curve.

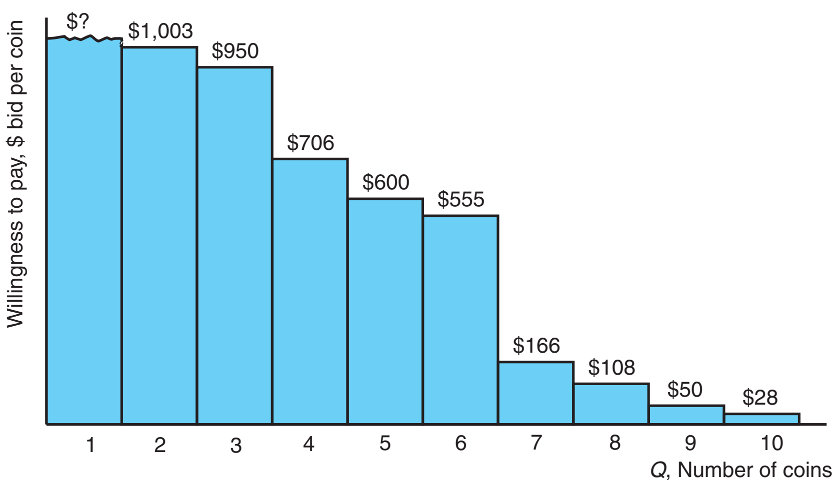
Application

Willingness to Pay on eBay

People differ in their willingness to pay for a given item. We can determine willingness to pay of individuals for a 238 AD Roman coin—a sesterce (originally equivalent in value to four asses) of the image of Emperor Balbinus—by how much they bid in an eBay auction. On its website, eBay correctly argues (as we show in Chapter 14) that an individual's best strategy is to bid his or her *willingness to pay*: the maximum value that the bidder places on the item. From what eBay reports, we know the maximum bid of each person except the winner, who paid the second-highest amount bid plus an increment.[3]

In the figure, the bids for the coin are arranged from highest to lowest. Because each bar on the graph indicates the bid for one coin, the figure shows how many units this group of bidders would have been willing to buy at various prices. That is, it is the market inverse demand curve.

Bapna et al. (2008) set up a website, **www.Cniper.com** (no longer active), that automatically bid on eBay at the last moment (a process called sniping). To use the site, an individual had to specify a maximum willingness to pay, so that the



[3]The increment depends on the size of the bid. It is $1 for the bids between $25 and $199.99 and $25 for bids between $1,000 and $2,499.99.

authors knew the top bidder's willingness to pay. The study found that the median consumer surplus was \$4 on goods that cost \$14 on average. They estimated the *CS* and the expenditures, *E*, for all eBay buyers and calculated that *CS/E* = 30%. That is, bidders' consumer surplus gain is 30% of their expenditures. Hasker et al. (2014) estimated that the median *CS* from buying a computer monitor on eBay was \$28 and that *CS/E* was about 19%.

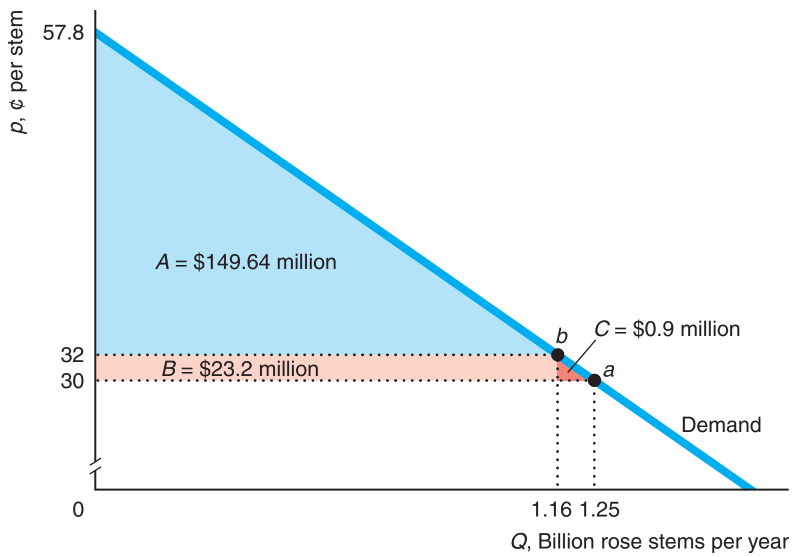
Effect of a Price Change on Consumer Surplus

If the supply curve shifts upward or a government imposes a new sales tax, the equilibrium price rises, reducing consumer surplus. We illustrate the effect of a price increase on market consumer surplus using estimated supply and demand curves for sweetheart and hybrid tea roses sold in the United States.[4] We then discuss which markets are likely to have the greatest loss of consumer surplus due to a price increase.

Suppose that the introduction of a new tax causes the (wholesale) price of roses to rise from the original equilibrium price of 30¢ to 32¢ per rose stem, a shift along the demand curve in Figure 9.3. The consumer surplus is area $A + B + C = \$173.74$ million stems per year at a price of 30¢, and it is only area $A = \$149.64$ million at a price of 32¢.[5] Thus, the loss in consumer surplus from the increase in the price is $B + C = \$24.1$ million per year.

Figure 9.3 Fall in Consumer Surplus from Roses as Price Rises MyLab Economics Video

As the price of roses rises 2¢ per stem from 30¢ per stem, the quantity demanded decreases from 1.25 to 1.16 billion stems per year. The loss in consumer surplus from the higher price, areas *B* and *C*, is \$24.1 million per year.



[4]I estimated this model using data from the *Statistical Abstract of the United States*, *Floriculture Crops*, *Floriculture and Environmental Horticulture Products*, and **http://www.usda.mannlib.cornell.edu**. The prices are in real 1991 dollars.

[5]The height of triangle *A* is $25.8¢ = 57.8¢ - 32¢$ per stem and the base is 1.16 billion stems per year, so its area is $\frac{1}{2} \times \$0.258 \times 1.16$ billion $= \$149.64$ million per year. Rectangle *B* is $\$0.02 \times 1.16$ billion $= \$23.2$ million. Triangle C is $\frac{1}{2} \times \$0.02 \times 0.09$ billion $= \$0.9$ million.

In general, as the price increases, consumer surplus falls more (1) the greater the initial revenues spent on the good and (2) the less elastic the demand curve (Appendix 9A). More is spent on a good when its demand curve is farther to the right so that areas like A, B, and C in Figure 9.3 are larger. The larger $B + C$ is, the greater is the drop in consumer surplus from a given percentage increase in price. Similarly, the less elastic a demand curve is (the closer it is to being vertical), the less willing consumers are to give up the good, so consumers do not cut their consumption much as the price increases, with the result of greater consumer surplus losses.

Higher prices cause greater consumer surplus loss in some markets than in others. Consumers would benefit if policymakers, before imposing a tax, considered in which market the tax is likely to harm consumers the most.

Application

Goods with a Large Consumer Surplus Loss from Price Increases

We can use estimates of demand curves to predict for which good a price increase causes the greatest loss of consumer surplus. The table shows the change in consumer surplus (ΔCS) in billions of 2016 dollars from a 10% increase in the price of various goods. As we would expect, the table shows that the larger the revenue (price times quantity) that is spent on a good, the larger the loss in consumer surplus.[6] A 10% increase in price causes a much greater loss of consumer surplus if it is imposed on medical services, \$184 billion, than if it is imposed on alcohol and tobacco, \$22 billion, because much more is spent on medical services.

	2012 Revenue	Estimated Elasticity of Demand,	Change in Consumer Surplus, ΔCS
Medical	1,902	−0.604	−184
Housing	1,698	−0.633	−164
Food	867	−0.245	−86
Clothing	383	−0.405	−38
Transportation	326	−0.461	−31
Utilities	321	−0.448	−31
Alcohol and tobacco	218	−0.162	−22

At first glance, the relationship between elasticities of demand and the loss in consumer surplus in the table looks backward: A given percent change in prices has a larger effect on consumer surplus for the relatively elastic demand curves. However, this relationship is coincidental: The large revenue goods happen to have relatively elastic demand curves. The effect of a price change depends on both revenue and the demand elasticity. In this table, the relative size of the revenues is more important than the relative elasticities.

If we could hold revenue constant and vary the elasticity, we would find that consumer surplus loss from a price increase is larger as the demand curve becomes less elastic. If the demand curve for alcohol and tobacco were 10 times more elastic, −1.62, while the revenue stayed the same—the demand curve became flatter at the initial price and quantity—the consumer surplus loss would be about \$2 billion less.

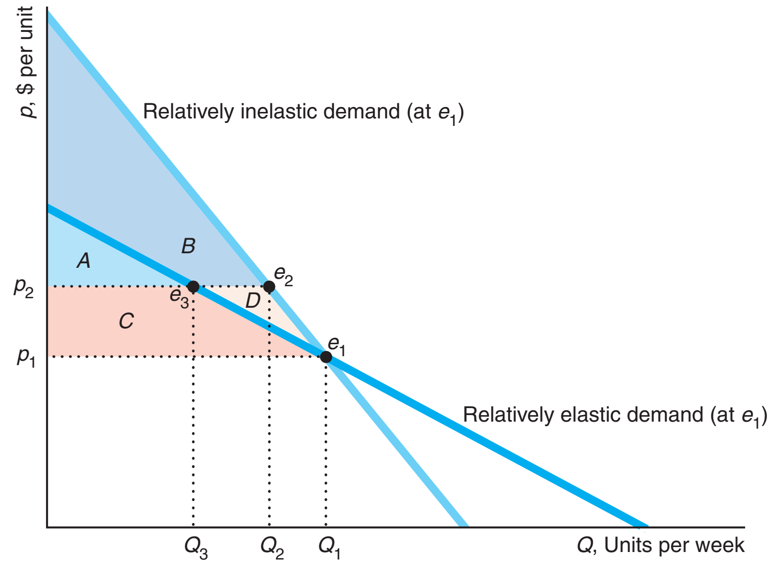
[6]Appendix 9A shows how to calculate ΔCS.

Solved Problem 9.1

Suppose that two linear demand curves go through the initial equilibrium, e_1. One demand curve is less elastic than the other at e_1. For which demand curve will a price increase cause the larger consumer surplus loss?

Answer

1. *Draw the two demand curves, and indicate which one is less elastic at the initial equilibrium.* Two demand curves cross at e_1 in the diagram. The steeper demand curve is less elastic at e_1.[7]



	Relatively Elastic Demand Curve	Relatively Inelastic Demand Curve
Consumer Surplus at p_1	$A + C$	$A + B + C + D$
Consumer Surplus at p_2	A	$A + B$
Consumer Surplus Loss	$-C$	$-C - D$

2. *Illustrate that a price increase causes a larger consumer surplus loss with the less elastic demand curve.* If the price rises from p_1 to p_2, the consumer surplus falls by only C with the relatively elastic demand curve and by $C + D$ with the relatively inelastic demand curve.

9.3 Producer Welfare

A supplier's gain from participating in the market is measured by its **producer surplus** (***PS***), which is the difference between the amount for which a good sells and the minimum amount necessary for the seller to be willing to produce the good. The

[7]As we discussed in Chapter 3, the price elasticity of demand, $= (\Delta Q/\Delta p)(p/Q)$, is 1 over the slope of the demand curve, $\Delta p/\Delta Q$, times the ratio of the price to the quantity. At the point of intersection where both demand curves have the same price, p_1, and quantity, Q_1, the steeper the demand curve, the lower the elasticity of demand.

producer surplus (*PS*) the difference between the amount for which a good sells and the minimum amount necessary for the seller to be willing to produce the good

minimum amount a seller must receive to be willing to produce is the firm's avoidable production cost (the shutdown rule in Chapter 8). Thus, for a particular quantity, a firm's producer surplus is the difference between its revenue from selling that quantity and its variable cost of producing that quantity.

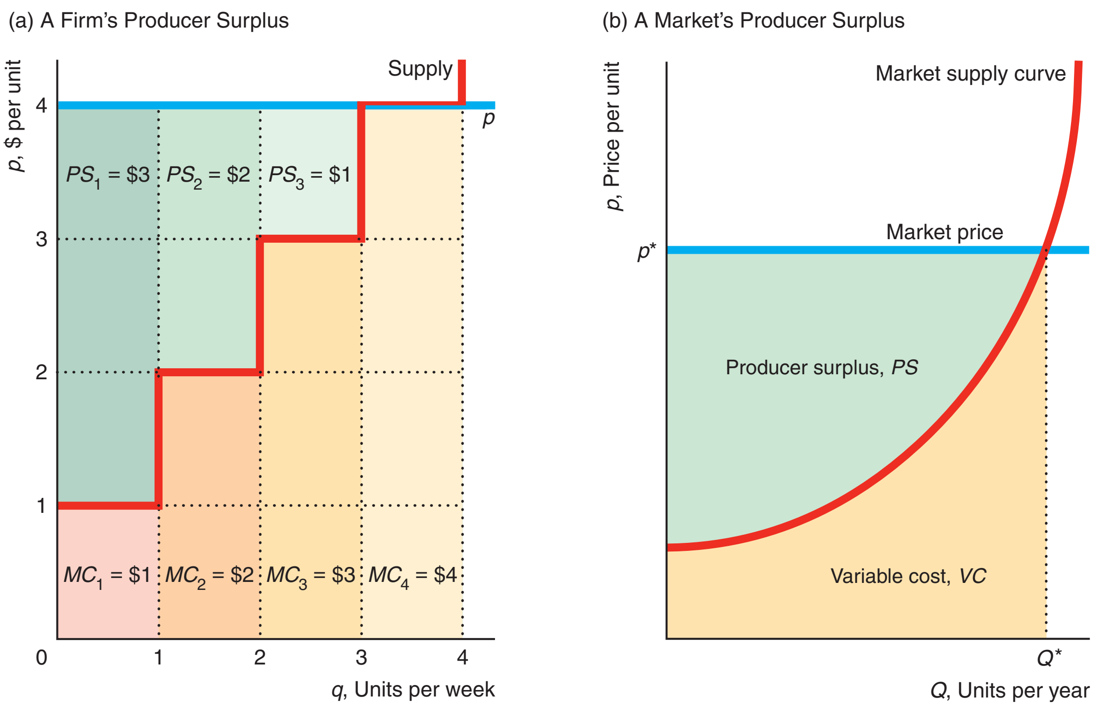
Measuring Producer Surplus Using a Supply Curve

To determine a competitive firm's producer surplus, we use its supply curve: its marginal cost curve above its minimum average variable cost (Chapter 8). The firm's supply curve in panel a of Figure 9.4 looks like a staircase. The marginal cost of producing the first unit is $MC_1 = \$1$, which is the area under the marginal cost curve between 0 and 1. The marginal cost of producing the second unit is $MC_2 = \$2$, and so on. The variable cost, VC, of producing four units is the sum of the marginal costs for the first four units:

$$VC = MC_1 + MC_2 + MC_3 + MC_4 = \$1 + \$2 + \$3 + \$4 = \$10.$$

Figure 9.4 Producer Surplus

(a) The firm's producer surplus, \$6, is the area below the market price, \$4, and above the marginal cost (supply curve) up to the quantity sold, 4. The area under the marginal cost curve up to the number of units actually produced is the variable cost of production. (b) The market producer surplus is the area above the supply curve and below the line at the market price, p^*, up to the quantity produced, Q^*. The area below the supply curve and to the left of the quantity produced by the market, Q^*, is the variable cost of producing that level of output.



If the market price, p, is \$4, the firm's revenue from the sale of the first unit exceeds its cost by $PS_1 = p - MC_1 = \$4 - \$1 = \$3$, which is its producer surplus on the first unit. The firm's producer surplus is \$2 on the second unit and \$1 on the third unit. On the fourth unit, the price equals marginal cost, so the firm just breaks even. As a result, the firm's total producer surplus, PS, from selling four units at \$4 each is the sum of its producer surplus on these four units:

$$PS = PS_1 + PS_2 + PS_3 + PS_4 = \$3 + \$2 + \$1 + \$0 = \$6.$$

Graphically, the total producer surplus is the area above the supply curve and below the market price up to the quantity actually produced. This same reasoning holds when the firm's supply curve is smooth.

The producer surplus is closely related to profit. Producer surplus is revenue, R, minus variable cost, VC:

$$PS = R - VC.$$

In panel a of Figure 9.4, revenue is $\$4 \times 4 = \$16$ and variable cost is \$10, so producer surplus is \$6. Profit is revenue minus total cost, C, which equals variable cost plus fixed cost, F: $\pi = R - C = R - VC - F$. Thus, the difference between producer surplus and profit,

$$PS - \pi = (R - VC) - (R - VC - F) = F,$$

is the fixed cost. If the fixed cost is zero (as often occurs in the long run), producer surplus equals profit.[8]

Another interpretation of producer surplus is as a gain to trade. In the short run, if the firm produces and sells its good—trades—it earns a profit of $R - VC - F$. If the firm shuts down—does not trade—it loses its fixed cost of $-F$. Thus, producer surplus equals the profit from trade minus the profit (loss) from not trading of

$$PS = R - VC = \pi - (-F) = (R - VC - F) + F.$$

Using Producer Surplus

Even in the short run, we can use producer surplus to study the effects of any shock that does not affect the fixed cost of firms, such as a change in the price of a substitute or an input. Such shocks change profit by exactly the same amount as they change producer surplus, $\Delta\pi = \Delta PS$, because fixed costs do not change.

A major advantage of producer surplus is that we can use it to measure the effect of a shock on *all* the firms in a market without having to measure the profit of each firm in the market separately. We can calculate market producer surplus using the market supply curve in the same way as we calculate a firm's producer surplus using its supply curve. The market producer surplus in panel b of Figure 9.4 is the area above the supply curve and below the market price, p^*, up to the quantity sold, Q^*. The market supply curve is the horizontal sum of the marginal cost curves of each of the firms (Chapter 8). As a result, the variable cost for all the firms in the market of producing Q is the area under the supply curve between 0 and the market output, Q.

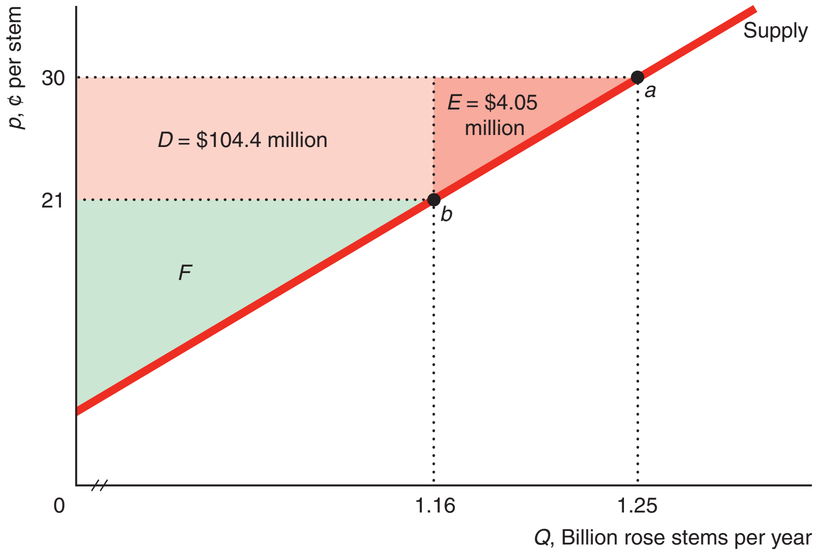
[8]Even though each competitive firm makes zero profit in the long run, owners of scarce resources used in that market may earn rents. Thus, owners of scarce resources may receive positive producer surplus in the long run.

Solved Problem 9.2

If the estimated supply curve for roses is linear, how much producer surplus is lost when the price of roses falls from 30¢ to 21¢ per stem (so that the quantity sold falls from 1.25 billion to 1.16 billion rose stems per year)?

Answer

1. *Draw the supply curve, and show the change in producer surplus caused by the price change.* The figure shows the estimated supply curve for roses. Point *a* indicates the quantity supplied at the original price, 30¢, and point *b* reflects the quantity supplied at the lower price, 21¢. The loss in producer surplus is the sum of rectangle *D* and triangle *E*.



	Original Price, 30¢	Lower Price, 21¢	Change ($ millions)
Producer Surplus	$D + E + F$	F	$-(D + E) = -108.45$

2. *Calculate the lost producer surplus by adding the areas of rectangle D and triangle E.* The height of rectangle *D* is the difference between the original and the new price, 9¢, and its base is 1.16 billion stems per year, so the area of *D* (not all of which is shown in the figure because of the break in the quantity axis) is \$0.09 per stem $\times$ 1.16 billion stems per year $=$ \$104.4 million per year. The height of triangle *E* is also 9¢, and its length is 90 million stems per year, so its area is $\frac{1}{2}$ $\times$ \$0.09 per stem $\times$ 90 million stems per year $=$ \$4.05 million per year. Thus, the loss in producer surplus from the drop in price is \$108.45 million per year.

9.4 Competition Maximizes Welfare

How should we measure society's welfare? People have proposed many reasonable welfare measures. One commonly used measure of the welfare of society, *W*, is the sum of consumer surplus plus producer surplus:

$$W = CS + PS.$$

This measure implicitly weights the well-being of consumers and producers equally. By using this measure, we are making a value judgment that the well-being of consumers and that of producers are equally important.

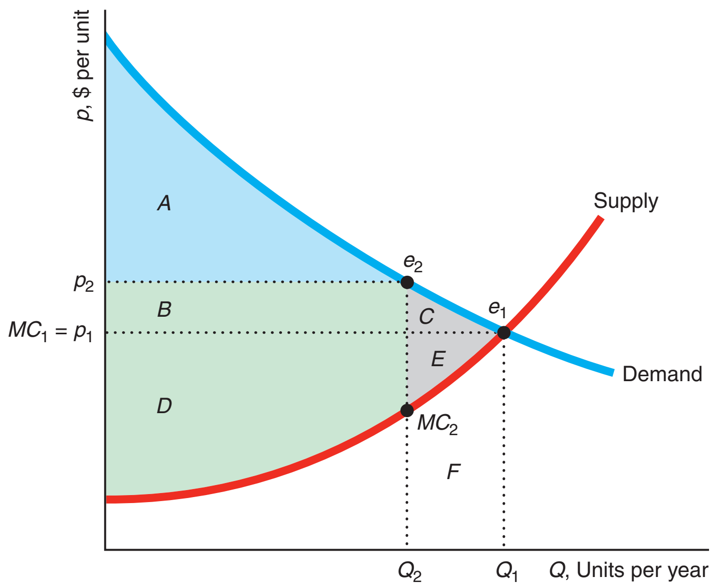
Not everyone agrees that society should try to maximize this measure of welfare. Groups of producers argue for legislation that helps them even if it hurts consumers by more than the producers gain—as though only producer surplus matters. Similarly, some consumer advocates argue that we should care only about consumers, so social welfare should include only consumer surplus.

We use the consumer surplus plus producer surplus measure of welfare in this chapter (and postpone a further discussion of other welfare concepts until the next chapter). One of the most striking results in economics is that competitive markets maximize this measure of welfare. If either less or more output than the competitive level is produced, welfare falls.

Producing less than the competitive output lowers welfare. At the competitive equilibrium in Figure 9.5, e_1, where output is Q_1 and price is p_1, consumer surplus is $CS_1 = A + B + C$, producer surplus is $PS_1 = D + E$, and total welfare is $W_1 = A + B + C + D + E$. If output is reduced to Q_2 so that price rises to p_2 at

Figure 9.5 Why Reducing Output from the Competitive Level Lowers Welfare MyLab Economics Video

Reducing output from the competitive level, Q_1, to Q_2 causes price to increase from p_1 to p_2. Consumers suffer: Consumer surplus is now A, a fall of $\Delta CS = -B - C$. Producers may gain or lose: Producer surplus is now $B + D$, a change of $\Delta PS = B - E$. Overall, the change in welfare is $\Delta W = -C - E$, which is a deadweight loss (DWL) to society.



	Competitive Output, Q_1 (1)	Smaller Output, Q_2 (2)	Change (2) – (1)
Consumer Surplus, CS	$A + B + C$	A	$-B - C = \Delta CS$
Producer Surplus, PS	$D + E$	$B + D$	$B - E = \Delta PS$
Welfare, $W = CS + PS$	$A + B + C + D + E$	$A + B + D$	$-C - E = \Delta W = DWL$

e_2, consumer surplus is $CS_2 = A$, producer surplus is $PS_2 = B + D$, and welfare is $W_2 = A + B + D$.

The change in consumer surplus is

$$\Delta CS = CS_2 - CS_1 = A - (A + B + C) = -B - C.$$

Consumers lose B because they have to pay $p_2 - p_1$ more than the competitive price for the Q_2 units they buy. Consumers lose C because they buy only Q_2 rather than Q_1 at the higher price.

The change in producer surplus is

$$\Delta PS = PS_2 - PS_1 = (B + D) - (D + E) = B - E.$$

Producers gain B because they now sell Q_2 units at p_2 rather than at p_1. They lose E because they sell $Q_2 - Q_1$ fewer units.

The change in welfare, $\Delta W = W_2 - W_1$, is[9]

$$\Delta W = \Delta CS + \Delta PS = (-B - C) + (B - E) = -C - E.$$

The area B is a transfer from consumers to producers—the extra amount consumers pay for the Q_2 units goes to the sellers—so it does not affect welfare. Welfare drops because the consumer loss of C and the producer loss of E benefit no one. This loss of welfare, $\Delta W = -C - E$, is a **deadweight loss** (***DWL***): the net reduction in welfare from a loss of surplus by one group that is not offset by a gain to another group from an action that alters a market equilibrium.

deadweight loss (*DWL*) the net reduction in welfare from a loss of surplus by one group that is not offset by a gain to another group from an action that alters a market equilibrium

The deadweight loss results because consumers value extra output by more than the marginal cost of producing it. At each output between Q_2 and Q_1, consumers' marginal willingness to pay for another unit—the height of the demand curve—is greater than the marginal cost of producing the next unit—the height of the supply curve. For example, at e_2, consumers value the next unit of output at p_2, which is much greater than the marginal cost, MC_2, of producing it. Increasing output from Q_2 to Q_1 raises firms' variable cost by area F, the area under the marginal cost (supply) curve between Q_2 and Q_1. Consumers value this extra output by the area under the demand curve between Q_2 and Q_1, area $C + E + F$. Thus, consumers value the extra output by $C + E$ more than it costs to produce it.

Society would be better off producing and consuming extra units of this good than spending this amount on other goods. In short, *the deadweight loss is the opportunity cost of giving up some of this good to buy more of another good*. Deadweight loss reflects a **market failure**—inefficient production or consumption—and is often due to the price not equaling the marginal cost.

market failure inefficient production or consumption due to the price not equaling the marginal cost

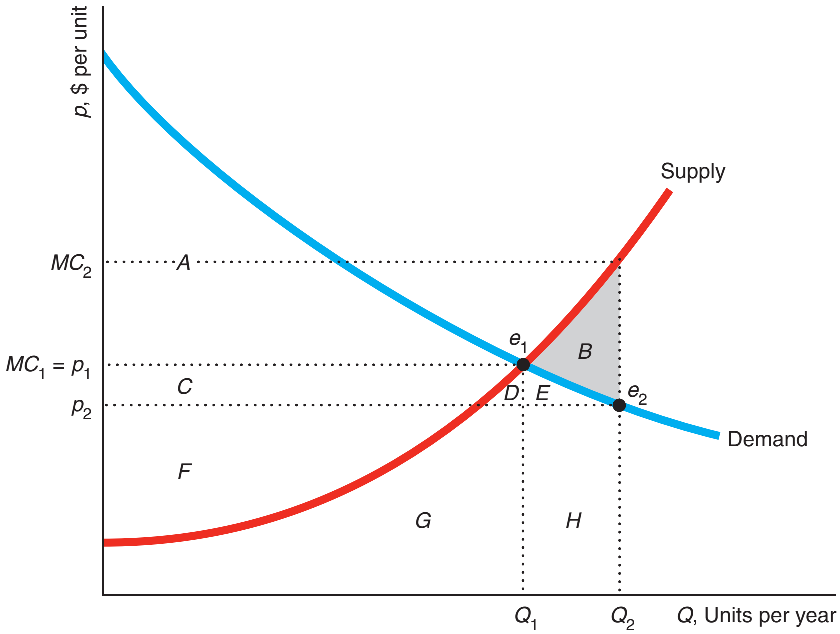
Solved Problem 9.3

Show that increasing output beyond the competitive level decreases welfare because the cost of producing this extra output exceeds the value consumers place on it.

Answer

1. *Illustrate that setting output above the competitive level requires the price to fall for consumers to buy the extra output.* The figure shows the effect of increasing output from the competitive level, Q_1, to Q_2. At the competitive equilibrium, e_1, the price is p_1. For consumers to buy the extra output at Q_2, the price must fall to p_2 at e_2 on the demand curve.

[9]The change in welfare is $\Delta W = W_2 - W_1 = (CS_2 + PS_2) - (CS_1 + PS_1) = (CS_2 - CS_1) + (PS_2 - PS_1) = \Delta CS + \Delta PS$.



	Competitive Output, Q_1	Larger Output, Q_2	Change
Consumer Surplus, *CS*	A	$A + C + D + E$	$C + D + E = \Delta CS$
Producer Surplus, *PS*	$C + F$	$F - B - D - E$	$-B - C - D - E = \Delta PS$
Welfare, $W = CS + PS$	$A + C + F$	$A + C + F - B$	$-B = \Delta W = DWL$

2. *Show how the consumer surplus and producer surplus change when the output level increases.* Because the price falls from p_1 to p_2, consumer surplus rises by $\Delta CS = C + D + E$, which is the area between p_2 and p_1 to the left of the demand curve. At the original price, p_1, producer surplus was $C + F$. The cost of producing the larger output is the area under the supply curve up to Q_2, $B + D + E + G + H$. The firms sell this quantity for only p_2Q_2, area $F + G + H$. Thus, the new producer surplus is $F - B - D - E$. As a result, the increase in output causes producer surplus to fall by $\Delta PS = -B - C - D - E$.

3. *Determine how welfare changes by adding the change in consumer surplus and producer surplus.* Because producers lose more than consumers gain, the deadweight loss is

$$DWL = \Delta W = \Delta CS + \Delta PS = (C + D + E) + (-B - C - D - E) = -B.$$

4. *Explain why welfare changes due to setting the price different than the marginal cost.* The new price, p_2, is less than the marginal cost, MC_2, of producing Q_2. Too much is being produced. A net loss occurs because consumers value the $Q_2 - Q_1$ extra output by only $E + H$, which is less than the extra cost, $B + E + H$, of producing it. The reason that competition maximizes welfare is that price equals marginal cost at the competitive equilibrium. At the competitive equilibrium, demand equals supply, which ensures that price equals marginal cost. When price equals marginal cost, consumers value the last unit of output

by exactly the amount that it costs to produce it. If consumers value the last unit by more than the marginal cost of production, welfare rises if more is produced. Similarly, if consumers value the last unit by less than its marginal cost, welfare is higher at a lower level of production.

Application

Deadweight Loss of Christmas Presents

Just how much did you enjoy the expensive lime green woolen socks with the dancing purple teddy bears that your Aunt Fern gave you? Often the cost of a gift exceeds the value that a recipient places on it.

Until the advent of gift cards, only 10% to 15% of holiday gifts were monetary. A gift of cash typically gives at least as much pleasure to the recipient as a gift that costs the same but can't be exchanged for cash. (So what if giving cash is tacky?) Of course, it's possible that a gift can give more pleasure to the recipient than it cost the giver—but how often does that happen to you?

An *efficient gift* is one that the recipient values as much as the gift costs the giver, or more. If the price of the gift exceeds its value to the recipient, the difference is a deadweight loss to society. Based on surveys of Yale undergraduates, Waldfogel (1993, 2009) estimated that the deadweight loss is between 10% and 33% of the value of gifts. Waldfogel (2005) found that consumers value their own purchases at 10% to 18% more, per dollar spent, than items received as gifts.[10]

Waldfogel found that gifts from friends and "significant others" are most efficient, while noncash gifts from members of the extended family are least efficient (one-third of the value is lost).[11] Luckily, grandparents, aunts, and uncles are most likely to give cash.

[10] Gift recipients may exhibit an endowment effect (Chapter 3) in which their willingness to pay (WTP) for the gift is less than what they would have to be offered to give up the gift, their willingness to accept (WTA). Bauer and Schmidt (2008) asked students at the Ruhr University in Germany their WTP and WTA for three recently received Christmas gifts. On average over all students and gifts, the average WTP was 11% percent below the market price and the WTA was 18% above the market price.

[11] Some people return an unwanted gift to a store. For the 2014 holiday season, gift recipients returned 11% of holiday presents, valued at $65 billion. Other people may deal with a disappointing present by "regifting" it. Some families have been passing the same fruitcake among family members for decades. According to one survey, 33% of women and 19% of men admitted that they pass on an unwanted gift (and 28% of respondents said that they would not admit it if asked whether they had done so).

Waldfogel concluded that a conservative estimate of the deadweight loss of Christmas, Hanukkah, and other holidays with gift-giving rituals is about \$12 billion. (And that's not counting about 2.8 billion hours spent shopping.) However, if the reason others don't give cash or gift cards is that they get pleasure from picking the "perfect" gift, the deadweight loss that adjusts for the pleasure of the giver is lower than these calculations suggest.

The question remains why people don't give cash instead of presents. Indeed, 77% of all Americans and 85% of those 25–34 years old give gift cards. (A gift card is similar to cash, though recipients can use some cards only in a particular store.) By one estimate, 2016 gift card sales were \$140 billion. Indeed, 93% of consumers say that they would prefer receiving a \$25 gift card to a gift that cost \$25. Bah, humbug!

9.5 Policies That Shift Supply and Demand Curves

One of the main reasons that economists developed welfare tools was to predict the impact of government programs that alter a competitive equilibrium. Virtually all government actions affect a competitive equilibrium in one of two ways. Some government policies shift the demand curve or the supply curve, such as a limit on the number of firms in a market. Others, such as sales taxes, create a wedge or gap between price and marginal cost so that they are not equal, even though they were in the original competitive equilibrium.

These government interventions move us from an unconstrained competitive equilibrium to a new, constrained competitive equilibrium. Because welfare was maximized at the initial competitive equilibrium, the examples of government-induced changes that we consider here lower welfare. In later chapters, we show that government intervention may raise welfare in markets in which welfare was not maximized initially.

Although government policies may cause either the supply curve or the demand curve to shift, we concentrate on policies that limit supply because they are used frequently and have clear-cut effects. If a government policy causes the supply curve to shift to the left, consumers make fewer purchases at a higher price and welfare falls. For example, if the supply curve in Figure 9.3 shifts to the left so that it hits the demand curve at e_2, then output falls from Q_1 to Q_2, the price rises from p_1 to p_2, and welfare falls by $C + E$. The only "trick" in this analysis is that we use the original supply curve to evaluate the effects on producer surplus and welfare.[12]

[12]Welfare falls when governments restrict the consumption of competitive products that we all agree are *goods*, such as food and medical services. In contrast, if most of society wants to discourage the use of certain products, such as hallucinogenic drugs and poisons, policies that restrict consumption may increase some measures of society's welfare.

During World War II, most of the nations involved limited the sales of consumer goods so that the nations' resources could be used for the war effort. Similarly, a government may cause a supply curve to shift to the left by restricting the number of firms in a market, such as by licensing taxicabs, psychiatric hospitals, or liquor licenses. We examine the effect of such policies in the Challenge Solution at the end of this chapter.

Entry Barrier

A government may also cause the supply curve to shift to the left by raising the cost of entry. If its cost will be greater than that of firms already in the market, a potential firm might not enter a market even if existing firms are making a profit. Any cost that falls only on potential entrants and not on current firms discourages entry. A long-run **barrier to entry** is an explicit restriction or a cost that applies only to potential new firms—existing firms are not subject to the restriction or do not bear the cost.

barrier to entry
an explicit restriction or a cost that applies only to potential new firms—existing firms are not subject to the restriction or do not bear the cost

At the time they entered, incumbent firms had to pay many of the costs of entering a market that new entrants incur, such as the fixed costs of building plants, buying equipment, and advertising a new product. For example, the fixed cost to McDonald's and other fast-food chains of opening a new fast-food restaurant is about $2 million. These fixed costs are *costs of entry* but are *not* barriers to entry because they apply equally to incumbents and entrants. Costs incurred by both incumbents and entrants do not discourage potential firms from entering a market if existing firms are making money. Potential entrants know that they will do as well as existing firms once they begin operations, so they are willing to enter as long as profit opportunities exist.

Large sunk costs can be barriers to entry under two conditions. First, if capital markets do not work efficiently so that new firms have difficulty raising money, new firms may be unable to enter profitable markets. Second, if a firm must incur a large *sunk* cost, which increases the loss if it exits, the firm may be reluctant to enter a market in which it is uncertain of success.

Application

Welfare Effects of Allowing Fracking

Technological advances have made hydraulic fracturing—fracking—a practical means to extract natural gas as well as oil from shale formations that previously could not be exploited (see the Application "Fracking and Shutdowns" in Chapter 8). Opponents of fracking fear that it pollutes air and water and triggers earthquakes. Due to their opposition, governments limit or prohibit fracking in parts of the United States and Europe.

Hausman and Kellogg (2015) used estimated natural gas supply and demand curves to calculate the welfare effects of permitting fracking. They found that the rightward shift of the supply curve reduced the U.S. natural gas price by 47% in 2013. As a result, consumer surplus increased substantially, particularly in the South Central and Midwestern United States, where the industrial and electric power industries use large quantities of gas. This drop in price was sufficient to reduce producer surplus. Hausman and Kellogg concluded that total surplus of gas consumers and producers increased by $48 billion, but noted that this calculation ignores harmful environmental effects.

Exit Restriction U.S., European, and other governments have laws that delay how quickly some (typically large) firms may go out of business so that workers can receive advance warning that they will be laid off. Although these restrictions keep the number of firms in a market relatively high in the short run, they may reduce the number of firms in a market in the long run.

Why do exit restrictions reduce the number of firms in a market in the long run? Suppose that you are considering starting a construction firm with no capital or other fixed factors. Your firm's only input is labor. You know that the demand for construction services is low during business downturns and in the winter. To avoid paying workers when business is slack, you plan to shut down during those periods. Because you can avoid losses by shutting down during low-demand periods, you enter this market if your expected economic profits during good periods are zero or positive.

A law that requires you to give your workers six months' warning before laying them off prevents your firm from shutting down quickly. You know that you'll regularly suffer losses during business downturns because you'll have to pay your workers for up to six months during periods when you have nothing for them to do. Knowing that you'll incur these regular losses, you are less inclined to enter the market. Unless the economic profits during good periods are much higher than zero—high enough to offset your losses—you will not choose to enter the market. If exit barriers limit the number of firms, the same analysis that we used to examine entry barriers applies. Thus, exit barriers may raise prices, lower consumer surplus, and reduce welfare.

9.6 Policies That Create a Wedge Between Supply and Demand

Never try to kill a government program—you'll only make it mad.

The most common government policies that create a wedge between supply and demand curves are sales taxes (or subsidies) and price controls. Because these policies create a gap between marginal cost and price, the market produces either too little or too much. For example, a tax causes price to exceed marginal cost—consumers value the good more than it costs to produce it—with the result that consumer surplus, producer surplus, and welfare fall.

Welfare Effects of a Sales Tax

A new sales tax causes the price consumers pay to rise (Chapter 3), resulting in a loss of consumer surplus, $\Delta CS < 0$, and a fall in the price firms receive, resulting in a drop in producer surplus, $\Delta PS < 0$. However, the new tax provides the government with new tax revenue, $\Delta T = T > 0$ (if tax revenue was zero before this new tax).

Assuming that the government does something useful with the tax revenue, we should include tax revenue in our definition of welfare: $W = CS + PS + T$. As a result, the change in welfare is

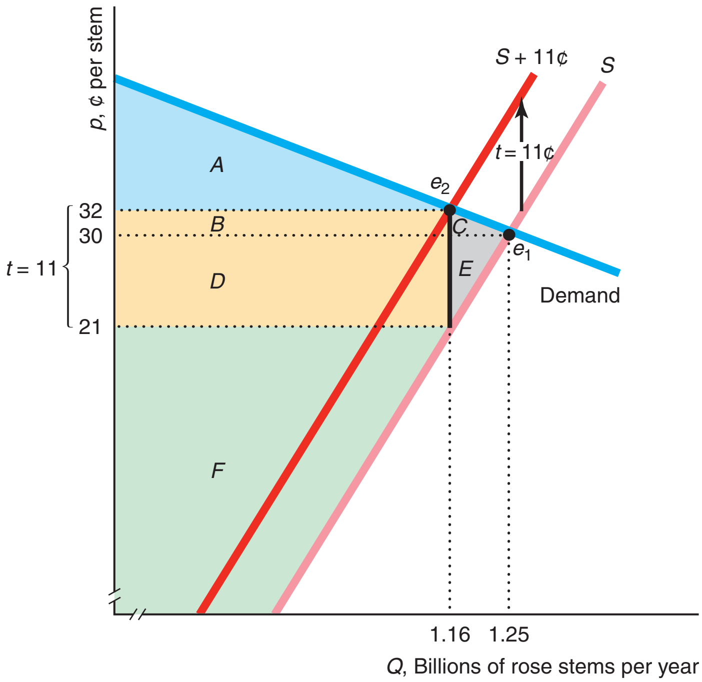
$$\Delta W = \Delta CS + \Delta PS + \Delta T.$$

Even when we include tax revenue in our welfare measure, a specific tax must lower welfare in a competitive market. We show the welfare loss from a specific tax of $t = 11¢$ per rose stem in Figure 9.6.

Without the tax, the intersection of the demand curve, D, and the supply curve, S, determines the competitive equilibrium, e_1, at a price of 30¢ per stem and a quantity of 1.25 billion rose stems per year. Consumer surplus is $A + B + C$, producer surplus is $D + E + F$, tax revenue is zero, and deadweight loss is zero.

Figure 9.6 Effects of a Specific Tax on Roses

The $t = 11¢$ specific tax on roses creates an 11¢ per stem wedge between the price customers pay, 32¢, and the price producers receive, 21¢. Tax revenue is $T = tQ = \$127.6$ million per year. The deadweight loss to society is $C + E = \$4.95$ million per year.



	No Tax	Specific Tax	Change ($ millions)
Consumer Surplus, CS	$A + B + C$	A	$-B - C = -24.1 = \Delta CS$
Producer Surplus, PS	$D + E + F$	F	$-D - E = -108.45 = \Delta PS$
Tax Revenue, $T = tQ$	0	$B + D$	$B + D = 127.6 = \Delta T$
Welfare, $W = CS + PS + T$	$A + B + C + D + E + F$	$A + B + D + F$	$-C - E = -4.95 = DWL$

The specific tax shifts the effective supply curve up by 11¢, creating an 11¢ wedge (Chapter 3) between the price consumers pay, 32¢, and the price producers receive, $32¢ - t = 21¢$. Equilibrium output falls from 1.25 to 1.16 billion stems per year.

The extra 2¢ per stem that buyers pay causes consumer surplus to fall by $B + C = \$24.1$ million per year. Due to the 9¢ drop in the price firms receive, they lose producer surplus of $D + E = \$108.45$ million per year (Solved Problem 9.2). The government gains tax revenue of $tQ = 11¢$ per stem $\times$ 1.16 billion stems per year $= \$127.6$ million per year, area $B + D$.

The combined loss of consumer surplus and producer surplus is only partially offset by the government's gain in tax revenue, so that welfare drops:

$$\begin{aligned}\Delta W &= \Delta CS + \Delta PS + \Delta T\\ &= -\$24.1 - \$108.45 + \$127.6\\ &= -\$4.95 \text{ million per year.}\end{aligned}$$

This deadweight loss is area $C + E$.

Why does society suffer a deadweight loss? The reason is that the tax lowers output from the competitive level where welfare is maximized. An equivalent explanation for this inefficiency or loss to society is that the tax puts a wedge between price and marginal cost. At the new equilibrium, buyers are willing to pay 32¢ for one more rose, while the marginal cost to firms is only 21¢ (= the price minus t). Shouldn't at least one more rose be produced if consumers are willing to pay nearly a third more than the cost of producing it? That's what our welfare study indicates.

Application

The Deadweight Loss from Gas Taxes

The social cost of collecting tax revenue is the deadweight loss that the tax causes. Blundell et al. (2012) found that the deadweight loss per dollar of gasoline tax revenue raised is 4.3% for high-income, 9.2% for middle-income, and 3.9% for low-income U.S. consumers.[13]

Why is a gasoline tax more distorting for middle-income consumers? Part of the explanation is that middle-income U.S. and Canadian consumers are much more responsive to changes in gasoline prices than low-income and high-income consumers. That is, middle-income consumers have a more elastic demand curve. Typically, the more that the quantity demanded falls in response to the tax, the wider is the deadweight loss triangle and the larger is the ratio of deadweight loss to tax revenue, as the next Solved Problem illustrates.

Solved Problem 9.4

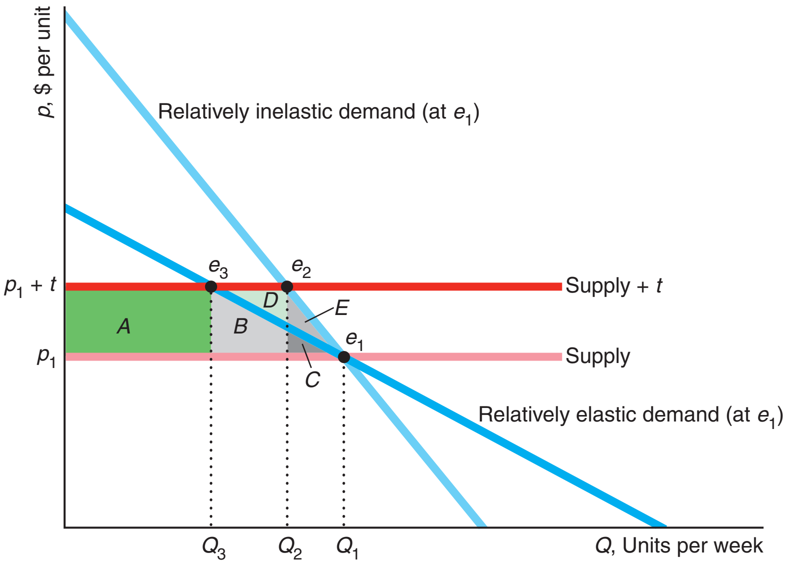
Two linear demand curves go through the initial equilibrium, e_1. One demand curve is less elastic than the other at e_1. The original horizontal supply curve also goes through e_1. For which demand curve is the deadweight loss from a specific tax t greater? For which is the ratio of the deadweight loss (DWL) to the tax revenue (T) greater?

Answer

1. *Draw a horizontal supply curve and add two demand curves with different slopes that intersect at the initial equilibrium, e_1, on the supply curve.* In the figure, the original supply curve and the demand curves intersect at e_1, where the equilibrium quantity is Q_1 and the price is p_1. The flatter demand curve is relatively more elastic as Solved Problem 9.1 explains.

2. *Show how the specific tax shifts the supply curve and determine the new equilibria for the two demand curves.* A specific tax of t shifts up the supply curve by t everywhere (Chapter 3). The relatively inelastic demand curve intersects this new supply curve at e_2, where the new equilibrium quantity is Q_2 and the new equilibrium price is $p_1 + t$. The relatively elastic demand curve intersects the new supply curve at e_3, where the equilibrium quantity is Q_3 and the price is again $p_1 + t$.

[13]The U.S. low-income group consists of the 25% lowest earners and has a median income of \$42,500 per year. The middle-income group has a median income of \$57,500. The median income of the high-income group is \$72,500. These calculations ignore the environmental effects from reduced consumption of gasoline.



	Relatively Elastic Demand Curve	Relatively Inelastic Demand Curve
Deadweight Loss, *DWL*	$B + C$	$C + E$
Tax Revenue, *T*	A	$A + B + D$
$\frac{DWL}{T}$	$\frac{B + C}{A}$	$\frac{C + E}{A + B + D}$

3. *Determine the DWL for the two demand curves.* The *DWL* is the area below the demand curve and above the supply curve between the new equilibrium quantity and Q_1. Thus, the *DWL* is area $C + E$ with the relatively inelastic demand curve and $B + C$ with the relatively elastic demand curve. Given that $B > E$, the *DWL* is greater with the more elastic demand curve.

4. *Determine the DWL/T for the two demand curves.* The tax revenue is t times the equilibrium quantity. Thus, T is the rectangle that lies between the two supply curves, which differ by t, between 0 and the equilibrium quantity. The tax revenue is area A with the relatively elastic demand curve and $A + B + D$ for the relatively inelastic demand curve. Thus, DWL/T is $(B + C)/A$ for the relatively elastic demand curve and $(C + E)/(A + B + D)$ for the relatively inelastic demand curve. Because the *DWL* is greater and T is smaller for the relatively elastic demand curve, DWL/T must be larger for the relatively elastic demand curve.

Welfare Effects of a Subsidy

A subsidy is a negative tax. Thus, its effect on output is the opposite of that of a tax. Nonetheless, a subsidy also lowers welfare. Because the new price is less than the (unsubsidized) marginal cost, a subsidy causes excess production, which lowers welfare as Solved Problem 9.3 shows.

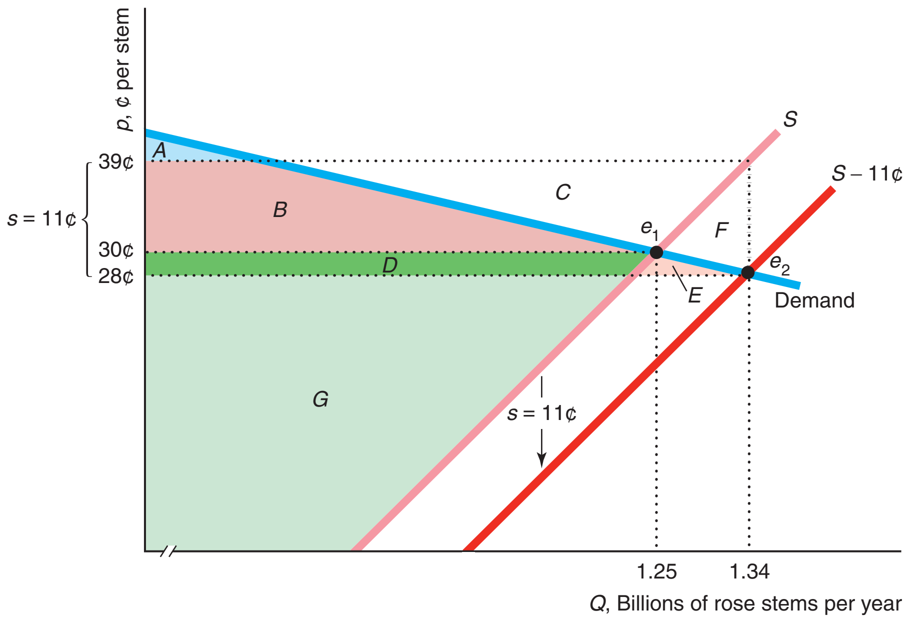
Solved Problem 9.5

Suppose that the government gives rose producers a specific subsidy of $s = 11¢$ per stem. What is the effect of the subsidy on the equilibrium prices and quantity, consumer surplus, producer surplus, government expenditures, welfare, and deadweight loss?

Answer

1. *Show how the subsidy shifts the supply curve and affects the equilibrium.* The specific subsidy shifts the supply curve, S in the figure, down by $s = 11¢$, to the curve labeled $S - 11¢$. Consequently, the equilibrium shifts from e_1 to e_2, so the quantity sold increases (from 1.25 to 1.34 billion rose stems per year), the price that consumers pay falls (from 30¢ to 28¢ per stem), and the amount that suppliers receive, including the subsidy, rises (from 30¢ to 39¢), so that the differential between what the consumer pays and the producers receive is 11¢.

2. *Show that consumers and producers benefit.* Consumers and producers of roses are delighted to be subsidized by other members of society. Because the price drops for customers, consumer surplus increases from $A + B$ to $A + B + D + E$. Because firms receive more per stem after the subsidy, producer surplus rises from $D + G$ to $B + C + D + G$ (the area under the price they receive and above the original supply curve).



	No Subsidy	Subsidy	Change ($ millions)
Consumer Surplus, CS	$A + B$	$A + B + D + E$	$D + E = 116.55 = \Delta CS$
Producer Surplus, PS	$D + G$	$B + C + D + G$	$B + C = 25.9 = \Delta PS$
Government Expense, $-X$	0	$-B - C - D - E - F$	$-B - C - D - E - F = -147.4 = \Delta X$
Welfare, $W = CS + PS - X$	$A + B + D + G$	$A + B + D + G - F$	$-F = -4.95 = DWL$

3. *Show how much government expenditure rises and determine the effect on welfare.* Because the government pays a subsidy of 11¢ per stem for each stem sold, the government's expenditures go from zero to the rectangle

$B + C + D + E + F$, which we call $-X$ in the table. Thus, the new welfare is the sum of the new consumer surplus and producer surplus and the government expenditure: $CS + PS + (-X)$. As the table under the figure shows, welfare falls from $A + B + D + G$ to $A + B + D + G - F$. The deadweight loss, this drop in welfare, $\Delta W = -F$, results from producing too much: The marginal cost to producers of the last stem, 39¢, exceeds the marginal benefit to consumers, 28¢.

Welfare Effects of a Price Floor

No matter what your religion, you should try to become a government program, for then you will have everlasting life. —Lynn Martin (former U.S. Representative)

In some markets, the government sets a *price floor*, or minimum price, which is the lowest price a consumer can pay legally for the good. For example, in most countries the government creates price floors under at least some agricultural prices to guarantee producers that they will receive at least a price of $\underline{p}$ for their good. If the market price is above $\underline{p}$, the support program is irrelevant. If the market price would be below $\underline{p}$, however, the government buys as much output as necessary to drive the price up to $\underline{p}$. Since 1929 (the start of the Great Depression), the U.S. government has used price floors or similar programs to keep prices of many agricultural products above the price that competition would determine in unregulated markets.

Agricultural Price Support Traditionally, the U.S. government has supported agricultural prices by buying some of the crop and storing it. We show the effect of a price support using estimated supply and demand curves for the soybean market (Holt, 1992).[14] The intersection of the market demand curve and the market supply curve in Figure 9.7 determines the competitive equilibrium, e, in the absence of a price support program, where the equilibrium price is $p_1 = \$4.59$ per bushel and the equilibrium quantity is $Q_1 = 2.1$ billion bushels per year.[15]

With a price support on soybeans of $\underline{p} = \$5.00$ per bushel and the government's pledge to buy as much output as farmers want to sell, quantity sold is $Q_s = 2.2$ billion bushels.[16] At $\underline{p}$, consumers buy less output, $Q_d = 1.9$ billion bushels, than the

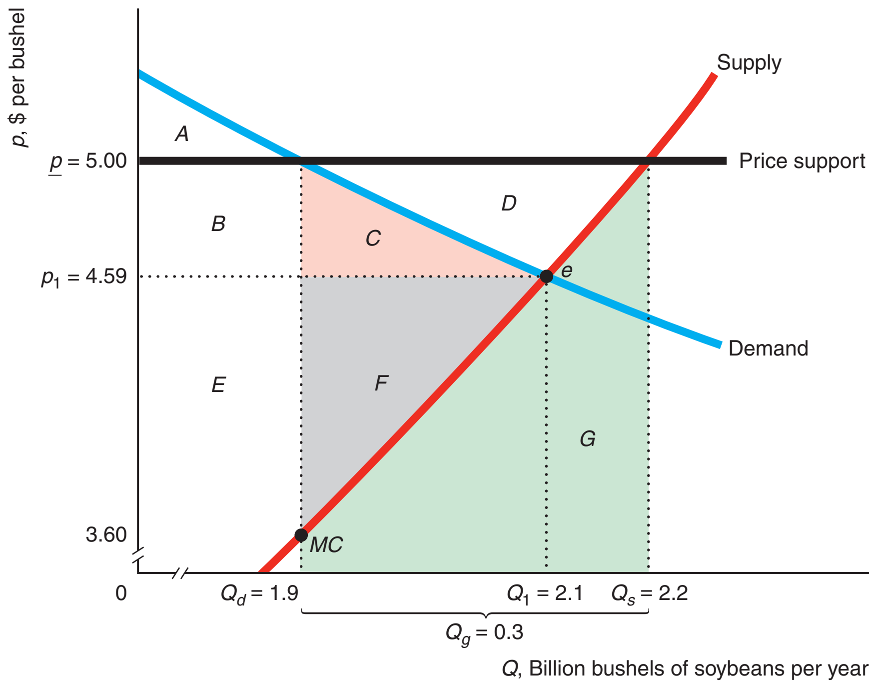
[14]The wool and mohair price support program is my favorite. The U.S. government instituted wool price supports after the Korean War to ensure "strategic supplies" for uniforms. Later, Congress added mohair to the program, even though mohair has no military use. In some years, the extra amount that the government paid for mohair exceeded the amount consumers paid for mohair, and the government payments on wool and mohair reached a fifth of a billion dollars over the first half-century of the program. No doubt the Clinton-era end of these subsidies in 1995 endangered national security. Thanks to Senator Phil Gramm, a well-known fiscal conservative, and other patriots (primarily from Texas, where much mohair is produced), Congress resurrected the program in 2000. Representative Lamar Smith took vehement exception to people who questioned the need for the mohair program: "Mohair is popular! I have a mohair sweater! It's my favorite one!" The 2006 budget called for \$11 million for wool and mohair with a loan rate of \$4.20 per pound. Again in 2011, the program was ended as a cost-cutting measure. However, Congress restored the wool and mohair program in 2012, and the 2014 agricultural bill extends it at least until 2018.

[15]The support or target price has increased slowly over time. It was \$5.02 in 1985 and \$6.00 in 2010–2012. The 2014 Farm Bill set the rate at \$8.40 for 2014–2018.

[16]For most of the last several decades, the soybean price support was around \$5 per bushel; however, it rose to \$6 in 2013.

Figure 9.7 Effects of Price Supports on Soybeans MyLab Economics Video

Without government price supports, the equilibrium is e, where $p_1 = \$4.59$ per bushel and $Q_1 = 2.1$ billion bushels of soybeans per year (based on estimates in Holt, 1992). With the price support at $\underline{p} = \$5.00$ per bushel, output sold increases to Q_s and consumer purchases fall to Q_d, so the government must buy $Q_g = Q_s - Q_d$ at a cost of \$1.283 billion per year. The deadweight loss is $-C - F - G = -\$1.226$ billion per year, not counting storage and administrative costs.



	No Price Support	Price Support	Change ($ millions)
Consumer Surplus, CS	$A + B + C$	A	$-B - C = -864 = \Delta CS$
Producer Surplus, PS	$E + F$	$B + C + D + E + F$	$B + C + D = 921 = \Delta PS$
Government Expense, $-X$	0	$-C - D - F - G$	$-C - D - F - G = -1{,}283 = \Delta X$
Welfare, $W = CS + PS - X$	$A + B + C + E + F$	$A + B + E - G$	$-C - F - G = -1{,}226 = \Delta W = DWL$

Q_1 they would have bought at the market-determined price p_1. As a result, consumer surplus falls by $B + C = \$864$ million. The government buys $Q_g = Q_s - Q_d \approx 0.3$ billion bushels per year, which is the excess supply, at a cost of $X = \underline{p} \times Q_g = C + D + F + G = \$1.283$ billion.

The government cannot resell the output domestically because if it tried to do so, it would succeed only in driving down the price consumers pay. The government stores the output or sends it abroad.

Although farmers gain producer surplus of $B + C + D = \$921$ million, this program is an inefficient way to transfer money to them. Assuming that the government's purchases have no alternative use, the change in welfare is $\Delta W = \Delta CS +$

$\Delta PS - T = -C - F - G = -\$1.226$ billion per year.[17] This deadweight loss reflects two distortions in this market:

- **Excess production.** More output is produced than is consumed, so Q_g is stored, destroyed, or shipped abroad.
- **Inefficiency in consumption.** At the quantity they actually buy, Q_d, consumers are willing to pay \$5 for the last bushel of soybeans, which is more than the marginal cost, $MC = \$3.60$, of producing that bushel.

Alternative Price Support Because of price supports, the government was buying and storing large quantities of food, much of which was allowed to spoil. Consequently, the government started limiting the amount farmers could produce. Because the government is uncertain about how much farmers will produce, it sets quotas or limits on the amount of land farmers may use, so as to restrict their output. Today, the government uses an alternative price support program. The government sets a support price, $\underline{p}$. Farmers decide how much to grow and sell all of their produce to consumers at the price, p, that clears the market. The government then gives the farmers a *deficiency payment* equal to the difference between the support and actual prices, $\underline{p} - p$, for every unit sold, so that farmers receive the support price on their entire crop.

Solved Problem 9.6

MyLab Economics Solved Problem

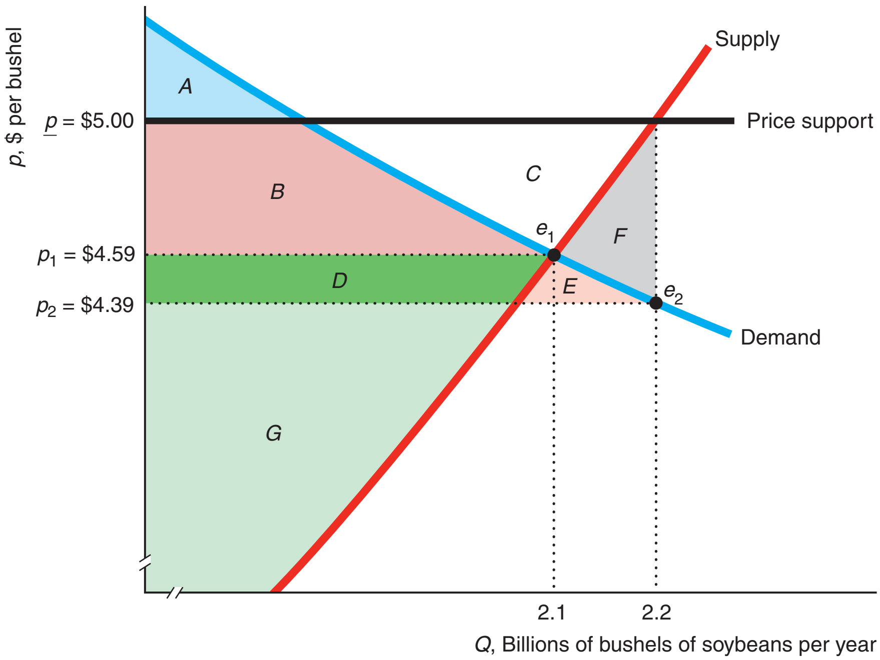
What are the effects in the soybean market of a \$5-per-bushel price support using a deficiency payment on the equilibrium price and quantity, consumer surplus, producer surplus, and deadweight loss?

Answer

1. *Describe how the program affects the equilibrium price and quantity.* Without a price support, the equilibrium is e_1 in the figure, where the price is $p_1 = \$4.59$ and the quantity is 2.1 billion bushels per year. With a support price of \$5 per bushel, the new equilibrium is e_2. Farmers produce at the quantity where the price support line hits their supply curve at 2.2 billion bushels. The equilibrium price is the height of the demand curve at 2.2 billion bushels, or approximately \$4.39 per bushel. Thus, the equilibrium price falls and the quantity increases.

2. *Show the welfare effects.* Because the price consumers pay drops from p_1 to p_2, consumer surplus rises by area $D + E$. Producers now receive $\underline{p}$ instead of p_1, so their producer surplus rises by $B + C$. Government payments are the difference between the support price, $\underline{p} = \$5$, and the price consumers pay, $p_2 = \$4.39$, times the number of units sold, 2.2 billion bushels per year, or the rectangle $B + C + D + E + F$. Because government expenditures exceed the gains to consumers and producers, welfare falls by the deadweight loss triangle F. Compared to the equivalent "buy and store" soybean price support program in Figure 9.7, the deficiency payment approach results in a smaller deadweight loss (less than a tenth of the original one) and lower government expenditures (though the expenditures need not be smaller in general).

[17] This measure of deadweight loss underestimates the true loss. The government also pays storage and administration costs. The U.S. Department of Agriculture, which runs farm support programs, has about 100,000 employees, or one worker for every eight farms that received assistance (although many U.S.D.A. employees have other job responsibilities, such as administering the food stamp program).



	No Price Support	Price Support	Change
Consumer Surplus, *CS*	$A + B$	$A + B + D + E$	$D + E = \Delta CS$
Producer Surplus, *PS*	$D + G$	$B + C + D + G$	$B + C = \Delta PS$
Government Expense, $-X$	0	$-B - C - D - E - F$	$-B - C - D - E - F = \Delta X$
Welfare, $W = CS + PS - X$	$A + B + D + G$	$A + B + D + G - F$	$-F = \Delta W = DWL$

Application

How Big Are Farm Subsidies and Who Gets Them?

Amount the EU paid to businessmen in Serbia–Montenegro for sugar subsidies before realizing that there was no sugar industry there: $1.2 million. —Harper's Index, 2004

Virtually every country in the world showers its farmers with subsidies. Although government support to farmers has fallen in developed countries over the last decade, support remains high. Farmers in developed countries received $585 billion in agricultural producer support payments (subsidies) in 2013–2015.

These payments were 18% of actual farm sales in developed countries in 2015. The percentage of subsidies ranged from 62% in Norway and Switzerland, 43% in Japan, 19% in the European Union, 9% in Canada and the United States, 1.3% in Australia, 0.7% in New Zealand, to 0.5% in Viet Nam.

In 2015, total U.S. agricultural support payments were $39 billion, or 0.5% of the U.S. gross domestic product. Each adult in the United States pays $159 a year to support agriculture. Did you get full value for your money? Some experts predict that the 2014 Farm Bill will lead to even greater federal government outlays.

The lion's share of U.S. farm subsidies goes to large agricultural corporations, not to poor farmers. According to the Environmental Working Group, three-quarters of the payments go to the largest and wealthiest 10% of farm operations and landlords, while nearly two-thirds of farmers receive no direct payments. Indeed, 23 members of Congress received payments; 9 recipients lived in Saudi Arabia, Hong Kong, and the United Kingdom; and $394 million went to absentee landlords who lived in big cities.

Welfare Effects of a Price Ceiling

In some markets, the government sets a *price ceiling*: the highest price that a firm can legally charge. If the government sets the ceiling below the pre-control competitive price, consumers demand more than the pre-control equilibrium quantity and firms supply less than that quantity (Chapter 2). Producer surplus must fall because firms receive a lower price and sell fewer units.

As a result of the price ceiling, consumers buy the good at a lower price but are limited by sellers as to how much they can buy. Because less is sold than at the pre-control equilibrium, society suffers a deadweight loss: Consumers value the good more than the marginal cost of producing extra units.

This measure of the deadweight loss may *underestimate* the true loss for two reasons. First, because consumers want to buy more units than are sold, they may spend additional time searching for a store with units for sale. This (often unsuccessful) search activity is wasteful and thus an additional deadweight loss to society. Deacon and Sonstelie (1989) calculated that for every $1 consumers saved from lower prices due to U.S. gasoline price controls in 1973, they lost $1.16 in waiting time and other factors.[18]

Second, when a price ceiling creates excess demand, the customers who are lucky enough to buy the good may not be the consumers who value it most. In a market without a price ceiling, all consumers who value the good more than the market price buy it, and those who value it less do not, so that those consumers who value it most buy the good. In contrast with a price control where the good is sold on a first-come, first-served basis, the consumers who reach the store first may not be the consumers with the highest willingness to pay. With a price control, if a lucky customer who buys a unit of the good has a willingness to pay of p_1, while someone who cannot buy it has a willingness to pay of $p_2 > p_1$, then the *allocative cost* to society of this unit being sold to the "wrong" consumer is $p_2 - p_1$.[19]

Solved Problem 9.7

MyLab Economics Solved Problem

What is the effect on the equilibrium, consumer surplus, producer surplus, and welfare if the government sets a price ceiling, $\bar{p}$, below the unregulated competitive equilibrium price?

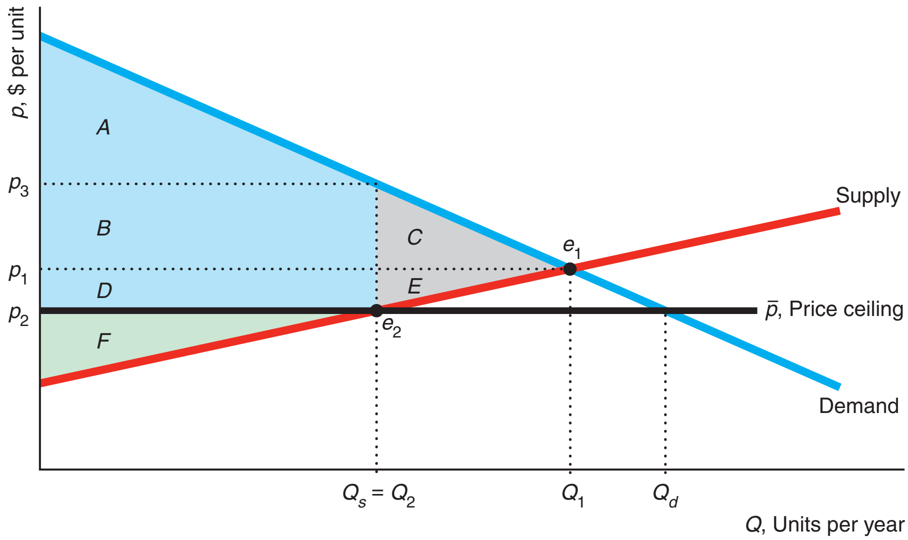
Answer

1. *Show the initial unregulated equilibrium.* The intersection of the demand curve and the supply curve determines the unregulated, competitive equilibrium, e_1, where the equilibrium quantity is Q_1.

[18] Perversely, this type of wasteful search does not occur if the good is efficiently but inequitably distributed to people according to a discriminatory criteria such as race, gender, or attractiveness, because people who are suffering discrimination know it is pointless to search.

[19] This allocative cost will be reduced or eliminated if a resale market exists where consumers who place a high value on the good can buy it from consumers who place a lower value on the good but were lucky enough to be able to buy it initially.

2. *Show how the equilibrium changes with the price ceiling.* Because the price ceiling, $\bar{p}$, is set below the equilibrium price of p_1, the ceiling binds (reduces the price that consumers pay). At this lower price, consumer demand increases to Q_d while the quantity firms are willing to supply falls to Q_s, so only $Q_s = Q_2$ units are sold at the new equilibrium, e_2. Thus, the price control causes the equilibrium quantity and price to fall, but consumers have excess demand of $Q_d - Q_s$.



	No Ceiling	Price Ceiling	Change
Consumer Surplus, CS	$A + B + C$	$A + B + D$	$D - C = \Delta CS$
Producer Surplus, PS	$D + E + F$	F	$-D - E = \Delta PS$
Welfare, $W = CS + PS$	$A + B + C + D + E + F$	$A + B + D + F$	$-C - E = \Delta W = DWL$

3. *Describe the welfare effects.* Because consumers are able to buy Q_s units at a lower price than before the controls, they gain area D. Consumers lose consumer surplus of C, however, because they can purchase only Q_s instead of Q_1 units of output. Thus, consumers gain net consumer surplus of $D - C$. Because they sell fewer units at a lower price, firms lose producer surplus $-D - E$. Part of this loss, D, is transferred to consumers in the form of lower prices, but the rest, E, is a loss to society. The deadweight loss to society is at least $\Delta W = \Delta CS + \Delta PS = -C - E$.

Application

The Social Cost of a Natural Gas Price Ceiling

From 1954 through 1989, U.S. federal law imposed a price ceiling on interstate sales of natural gas. The law did not apply to sales within states in the Southwest that produced the gas—primarily Louisiana, Oklahoma, New Mexico, and Texas. Consequently, consumers in the Midwest and Northeast, where most of the gas was used, were less likely to be able to buy as much natural gas as they wanted, unlike consumers in the Southwest. Because they could not buy natural gas, some consumers who would have otherwise done so did not install natural gas heating. As heating systems last for years, even today, many homes use dirtier fuels such as heating oil due to this decades-old price control.

By comparing consumer behavior before and after the price control period, Davis and Kilian (2011) estimated that demand for natural gas exceeded observed sales of natural gas by an average of 19.4% from 1950 through 2000. They calculated that the allocative cost averaged \$3.6 billion annually during this half century. This additional loss is a third of the estimated annual deadweight loss from the price control of \$10.5 billion (MacAvoy, 2000). The total loss is \$14.1 (= \$10.5 + \$3.6) billion.[20]

9.7 Comparing Both Types of Policies: Imports

Traditionally, most of the United States' imports come from overseas.

We've examined examples of government policies that shift supply or demand curves and policies that create a wedge between supply and demand. Governments use both types of policies to control international trade.

Allowing imports of foreign goods benefits the importing country. If a government reduces imports of a good, the domestic price rises; the profits of domestic firms that produce the good increase, but domestic consumers are hurt. Our analysis will show that the loss to consumers exceeds the gain to producers.

The government of the (potentially) importing country can use one of four import policies:

tariff (*duty*) a tax on only imported goods

- **Allow free trade.** Any firm can sell in this country without restrictions.
- **Ban all imports.** The government sets a quota of zero on imports.
- **Set a tariff.** The government imposes a tax called a **tariff (or a *duty*)** on only imported goods.
- **Set a positive quota.** The government limits imports to $\overline{Q}$.

We compare welfare under free trade to welfare under bans and quotas, which change the supply curve, and to welfare under tariffs, which create a wedge between supply and demand.

To illustrate the differences in welfare under these various policies, we use estimated U.S. crude oil supply and demand curves.[21] To examine the U.S. market, we make two assumptions for the sake of simplicity: We assume that transportation costs are zero and the supply curve of the potentially imported good is horizontal at the world price p^*. Given these assumptions, the importing country, the United States, can buy as much of this good as it wants at p^* per unit: It is a price taker in the world market because its demand is too small to influence the world price.

Free Trade Versus a Ban on Imports

No nation was ever ruined by trade. —Benjamin Franklin

Preventing imports into the domestic market raises the price. We now compare the U.S. crude oil market equilibrium with and without free trade.

Although the estimated U.S. domestic supply curve, S^a, in Figure 9.8 is upward sloping, the foreign supply curve is horizontal at the world price of \$60. The total

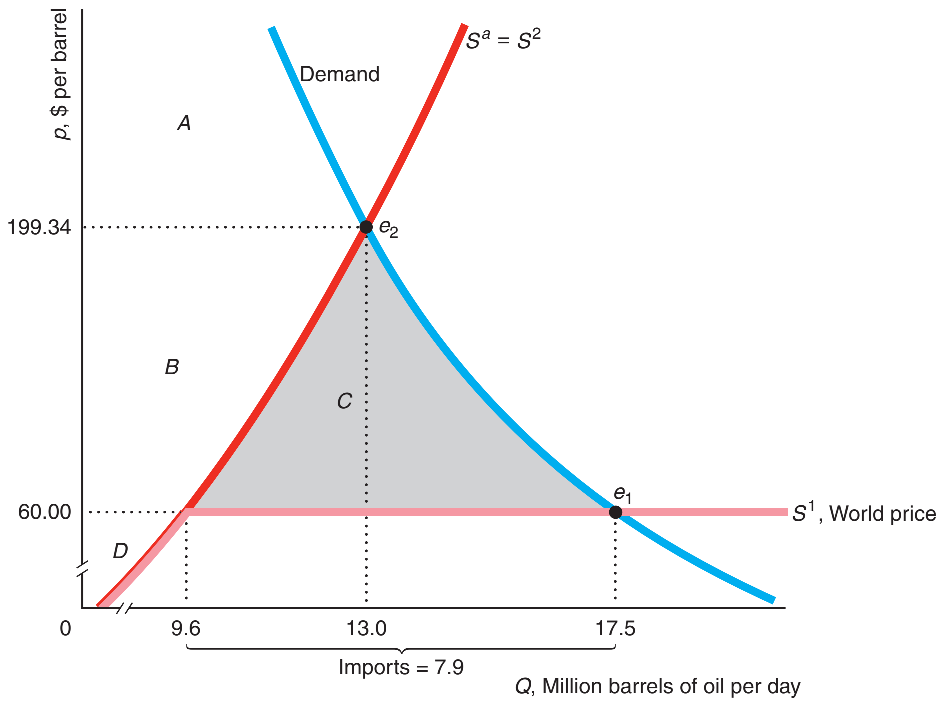
---

[20] Consumers' share of the deadweight loss, area C in the figure in Solved Problem 9.7, is \$9.3 billion annually; the sellers' share, area E, is \$1.2 billion; so the entire deadweight loss is \$10.5 billion. Consumers benefit by area D = \$6.9 billion from paying a lower price, which represents a transfer from sellers. Thus, altogether consumers lose \$6.0 (= \$9.3 + \$3.6 − \$6.9) billion and firms lose \$8.1 (= \$1.2 + \$6.9) billion.

[21] These constant-elasticity supply and demand equations for crude oil are based on Baumeister and Peersman (2013), with rounding and updating using 2015 production and import data.

Figure 9.8 Loss from Eliminating Free Trade MyLab Economics Video

Because the supply curve that foreigners face is horizontal at the world price of \$60, the total U.S. supply curve of crude oil is S^1 with free trade. The free-trade equilibrium is e_1. With a ban on imports, the equilibrium e_2 occurs where the domestic supply curve, $S^a = S^2$, intersects the demand curve. The ban increases producer surplus by B = \$1,606 million per day and decreases consumer surplus by $B + C$ = \$2,045 million per day, so the deadweight loss is C = \$439 million per day or about \$160 billion per year.



U.S.	Free Trade	U.S. Import Ban	Change (\$ millions)
Consumer Surplus, CS	$A + B + C$	A	$-B - C = -2{,}045 = \Delta CS$
Producer Surplus, PS	D	$B + D$	$B = 1{,}606 = \Delta PS$
Welfare, $W = CS + PS$	$A + B + C + D$	$A + B + D$	$-C = -439 = \Delta W = DWL$

U.S. supply curve, S^1, is the horizontal sum of the domestic supply curve and the foreign supply curve. Thus, S^1 is the same as the upward-sloping domestic supply curve for prices below \$60 and is horizontal at \$60. Under free trade, the United States imports crude oil if its domestic price in the absence of imports would exceed the world price, \$60 per barrel.

The intersection of S^1 and the demand curve determines the free-trade equilibrium, e_1, where the U.S. price equals the world price, \$60. U.S. consumers demand 17.5 million barrels per day at that price. Because domestic firms produce only 9.6 million barrels, imports are $17.5 - 9.6 = 7.9$ million barrels per day. U.S. consumer surplus is $A + B + C$, U.S. producer surplus is D, and U.S. welfare is $A + B + C + D$. Throughout our discussion of trade, we ignore welfare effects in other countries.

If the U.S. government were to ban imports, the total U.S. supply curve, S^2, is the American domestic supply curve, S^a. The intersection of S^2 and the demand curve determines the no-trade equilibrium at e_2, where the equilibrium price is \$199.34 and the quantity is 13.0. Consumer surplus is only A, producer surplus is $B + D$, and welfare is $A + B + D$.

Thus, the ban harms consumers by more than it helps producers. Because of the higher price, domestic firms gain producer surplus of $\Delta PS = B$. The change in consumer surplus is $\Delta CS = -B - C$. The change in total welfare, ΔW, is the difference between the gain to producers and the loss to consumers, $\Delta W = \Delta PS + \Delta CS = -C$, so the ban hurts society.

Application

Russian Food Ban

Starting in 2014, many Western nations imposed a variety of sanctions on Russia because of its military activities in Ukraine. In retaliation, Russia banned imports of meat, poultry, fish, seafood, dairy and meat products, vegetables, and fruit and nuts from the United States, the European Union, Canada, and other countries.

Russians, particularly in prosperous cities such as Moscow, depend heavily on imported foods from the West. The previous year, 2013, Russian agricultural imports were about \$1 billion from the United States and 11.8 billion euros (\$15.7 billion) from the European Union.

The ban imposes substantial costs on Russian consumers. In 2014, food prices soared 11.5%, which was 5.8% higher than the overall inflation rate. Prices for some types of food shot up by even more. Meat and poultry prices rose 18% over the previous year, while the price of butter shot up by 17%.

The ban had less of an effect on firms in exporting nations, which could sell their products elsewhere. However, foreign firms that catered to the Russian market were hurt, such as the Finnish dairy company Valio, which closed its plant near Russia in 2016.

Of course, Russian food-producing firms benefited, as Russia's Agriculture Minister claimed in 2016. For example, in the first quarter after the ban went into effect, the profit of Cherkizovo, a Russian producer of meat, rose eightfold from the previous year. However, in 2016, Russia announced that it was moderating the ban, allowing in products intended for children's consumption and fruit from Turkey.

Free Trade Versus a Tariff

TARIFF, n. A scale of taxes on imports, designed to protect the domestic producer against the greed of his customers. —Ambrose Bierce

Governments use *specific tariffs* (t dollars per unit) and *ad valorem tariffs* (v percent of the sales price). Governments around the world use tariffs, particularly on agricultural products.[22] American policymakers have frequently debated the

[22]After World War II, most trading nations signed the General Agreement on Tariffs and Trade (GATT), which limited their ability to subsidize exports or limit imports using quotas and tariffs. The rules prohibited most export subsidies and import quotas, except when imports threatened "market disruption" (a term that unfortunately was not defined). The GATT also required that any new tariff be offset by a reduction in other tariffs to compensate the exporting country. Modifications of the GATT and agreements negotiated by its successor, the World Trade Organization, have reduced or eliminated many tariffs.

optimal tariff on crude oil as a way to raise revenue or to reduce dependence on foreign oil.

You may be asking yourself, "Why should we study tariffs if we've already looked at taxes? Isn't a tariff just another tax?" Good point! Tariffs are just taxes. If the only goods sold in the market were imported, the effect of a tariff in the importing country would be the same as for a sales tax. We study tariffs separately because a government applies a tariff to only imported goods, so it affects domestic and foreign producers differently. Because tariffs apply to only imported goods, all else the same, they do not raise as much tax revenue or affect equilibrium quantities as much as taxes applied to all goods in a market.

To illustrate the effect of a tariff clearly in our figures, we suppose that the U.S. government imposes a large specific tariff of $t = \$40$ per barrel of crude oil. Consequently, firms import oil into the United States only if the U.S. price is at least $40 above the world price. The tariff creates a wedge between the world price, $60, and the U.S. price, $100. The tariff shifts the U.S. total supply curve upward from S^1 to S^3 in Figure 9.9 so that S^3 equals the domestic supply curve for prices below $100 and is horizontal at $100.

To illustrate the effect of a tariff, suppose that the government imposes a specific tariff of $t = \$40$ per barrel of crude oil. Given this tariff, firms will not import oil into the United States unless the U.S. price is at least $40 above the world price, $93. The tariff creates a wedge between the world price and the U.S. price. This tariff causes the total supply curve to shift from S^1 to S^3 in Figure 9.9. As the world supply curve for oil is horizontal at a price of $93, adding a $40 tariff shifts this supply curve upward so that it is horizontal at $133. That is, the rest of the world will supply an unlimited amount of oil at $133 inclusive of the tariff. As a result, the total U.S. supply curve with the tariff, S^3, equals the domestic supply curve for prices below $133 and is horizontal at $133.

In the new equilibrium, e_3, where S^3 intersects the demand curve, the equilibrium price is $100 and the quantity is 15.4 million barrels of oil per day. Domestic firms supply 10.9 million barrels, so imports are 4.5.

The tariff *protects* American producers from foreign competition. The larger the tariff, the fewer the imports, and hence the higher the price that domestic firms can charge. (With a large enough tariff, firms import nothing and the price rises to the no-trade level, $199.34.) With a tariff of $40, domestic firms' producer surplus increases by area $B = \$412$ million per day.

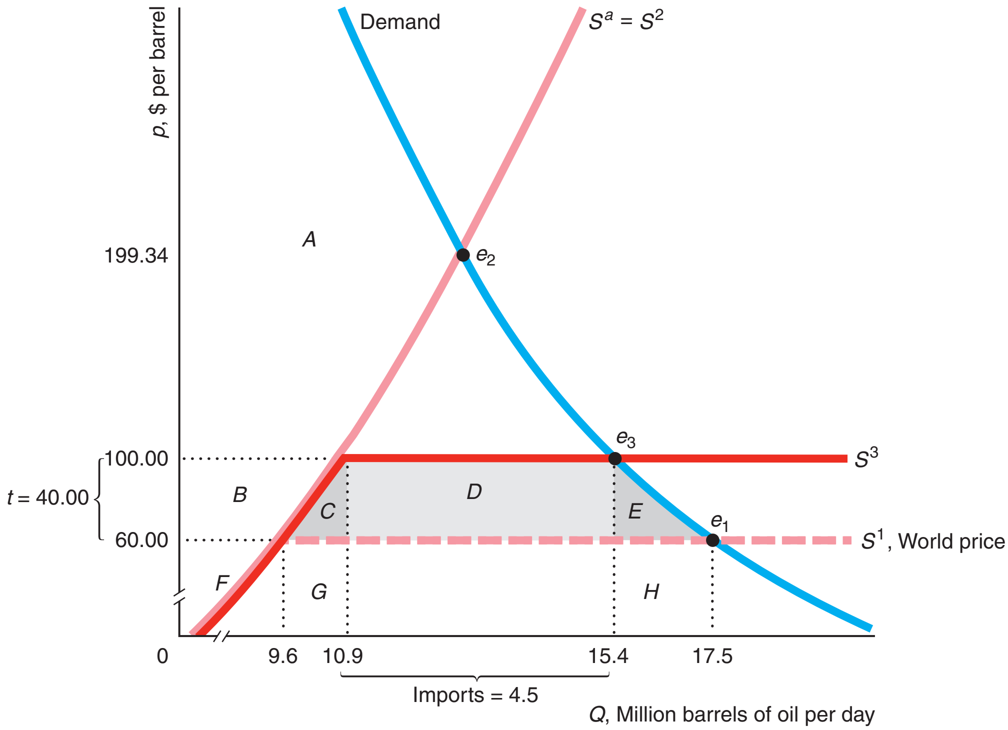
Because the U.S. price rises from $60 to $100, consumer surplus falls by $B + C + D + E = \$654$ million per day. The government receives tariff revenues, T, equal to area $D = \$180$ million per day, which is $t = \$40$ times the quantity imported, 4.5 million.

The deadweight loss is the loss of consumer surplus, $B + C + D + E$, minus the tax revenue, D, minus the producer surplus gain, B. That is, the deadweight loss is $C + E = \$62$ million per day, or $22.6 billion per year. This deadweight loss is 15% of the gain to producers. Consumers lose $1.59 for each $1 that domestic producers gain. Because the tariff does not eliminate all imports, the welfare loss is smaller than from an import ban.

This deadweight loss has two components. First, C is the *production distortion lost* from U.S. firms producing 10.9 million barrels per day instead of 9.6 million barrels per day. Domestic firms produce this extra output because the tariff drives up the price from $60 to $100. The cost of producing these extra 1.3 million barrels of oil per day domestically is $C + G$, the area under the domestic supply curve, S^a, between 9.6 and 10.9 million barrels per day. Had Americans bought this oil at the world price, the cost would have been only G. Thus, C is the additional cost of producing the extra 1.3 million barrels of oil per day domestically instead of importing it.

Figure 9.9 Effects of a Tariff (or Quota) MyLab Economics Video

A tariff of $t = \$40$ per barrel of oil imported or a quota of $Q = 3.4$ million barrels per day drives the U.S. price of crude oil to \$100, which is \$40 more than the world price of \$60. Under the tariff, the intersection of the total U.S. supply curve, S_3, and the demand curve determine the equilibrium, e_3. Under the quota, e_3 is determined by a quantity wedge of 4.5 million barrels per day between the quantity demanded, 15.4 million barrels per day, and the quantity supplied, 10.9 million barrels per day. Compared to free trade, producers gain $B = \$412$ million per day and consumers lose $B + C + D + E = \$654$ million per day from the tariff or quota. The deadweight loss under the quota is $C + D + E = \$242$ million per day. With a tariff, the government's tariff revenue increases by $D = \$180$ million a day, so the deadweight loss is less—"only" $C + E = \$62$ million per day.



U.S.	Free Trade	U.S. Tariff or Quota	Change (\$ millions)
Consumer Surplus, *CS*	$A + B + C + D + E$	A	$-B - C - D - E = -654$
Producer Surplus, *PS*	F	$B + F$	$B = 412$
Tariff Revenues, *T*	0	D (tariff) 0 (quota)	$D = 180$ (tariff) 0 (quota)
Welfare from a Tariff, $W = CS + PS + T$	$A + B + C + D + E + F$	$A + B + D + F$	$-C - E = -62 = DWL$
Welfare from a Quota, $W = CS + PS$	$A + B + C + D + E + F$	$A + B + F$	$-C - D - E = -242 = DWL$

Second, E is a *consumption distortion loss* from U.S. consumers' buying too little oil, 15.4 instead of 17.5 million barrels, because the tariff increases the price from \$60 to \$100.[23] U.S. consumers place a value on this extra output of $E + H$, the area under their demand curve between 15.4 and 17.5 million barrels per day. The

[23]This analysis ignores the effect of oil consumption on the environment. We address that issue in Chapter 18.

cost of buying this extra oil from the world market is only *H*, the area below the line at $60 between 15.4 and 17.5. Thus, *E* is the difference between the value at the world price and the value U.S. consumers place on this extra 2.1 million barrels per day.

Free Trade Versus a Quota

Many countries use quotas instead of tariffs, which may lead to a false belief:

Common Confusion: Quotas are preferable to tariffs.

Although politicians have a variety of reasons to prefer quotas, countries usually benefit from employing a tariff rather than an equivalent quota—that reduces imports by the same amount—because only the tariff produces revenue for the government. (Of course, countries would generally be better off using neither.)

The market effects of a quota are similar to those of a tariff. In Figure 9.9, if the government limits imports to $\overline{Q} = 4.5$ million barrels per day, the quota is binding because firms import 7.9 million barrels per day under free trade. This quota on imports of 4.5 million barrels per day leads to the same equilibrium, e_3 in Figure 9.8, as a tariff of $40. Given this binding quota, the equilibrium price is $100. The quantity imported, 4.5 million barrels per day, equals the quantity demanded, 15.4 million barrels per day, minus the quantity supplied by domestic producers, 10.9 million barrels per day.

The gain to domestic producers, *B*, and the loss to consumers, $B + C + D + E$, are the same as those with a tariff. Suppose the government gives licenses to some foreign sellers that permit them to sell the quota amount in the United States. As a result, the government does not receive any revenue from the quota, unlike with a tariff. Thus, the deadweight loss with this quota, $C + D + E = \$142$ million per day, is greater than the deadweight loss with the equivalent tariff, $C + E = \$62$ million per day. The extra deadweight loss from using the quota instead of a tariff is the forgone government tariff revenues, $D = \$180$ million per day, which goes to foreign firms.

Thus, the importing country fares better using a tariff than setting this quota that reduces imports by the same amount. Consumers and domestic firms do as well under the two policies, but the government gains tariff revenues, *D*, only when the tariff is used.

However, if the government gives the quota licenses to domestic importing firms, then the United States does not lose *D*. Similarly, if the government sells the quota licenses to firms at a price equal to the tariff, then the government gets *D*.

Rent Seeking

Given that tariffs and quotas hurt the importing country, why do the Japanese, U.S., and other governments impose tariffs, quotas, or other trade barriers? The reason is that domestic producers stand to make large gains from such government actions, so the producers lobby the government to enact these trade policies. Although consumers as a whole suffer large losses, the loss to any one consumer is usually small. Moreover, consumers rarely organize to lobby the government about trade issues. Thus, in most countries, producers are often able to convince (cajole, influence, or bribe) legislators or government officials to aid them, even though the loss to consumers exceeds the gain to domestic producers.

If domestic producers can talk the government into a tariff, quota, or other policy that reduces imports, they gain extra producer surplus (rents), such as area *B* in

rent seeking
efforts and expenditures to gain a rent or a profit from government actions

Figures 9.8 and 9.9. Economists call efforts and expenditures to gain a rent or a profit from government actions **rent seeking**. If producers or other interest groups bribe legislators to influence policy, the bribe is a transfer of income and hence does not increase deadweight loss (except to the degree that a harmful policy is chosen). However, if this rent-seeking behavior—such as hiring lobbyists and engaging in advertising to influence legislators—uses up resources, the deadweight loss from tariffs and quotas understates the true loss to society. The domestic producers may spend up to the gain in producer surplus to influence the government.[24]

Indeed, some economists argue that the government revenues from tariffs are completely offset by administrative costs and rent-seeking behavior. If so (and if the tariffs and quotas do not affect world prices), the loss to society from tariffs and quotas is all of the change in consumer surplus, such as areas $B + C$ in Figure 9.8 and areas $B + C + D + E$ in Figure 9.9.

Lopez and Pagoulatos (1994) estimated the deadweight loss and the additional losses due to rent-seeking activities in the United States in food and tobacco products. They estimated that the deadweight loss was \$18.3 billion (in 2015 dollars), which was 2.6% of the domestic consumption of these products. The largest deadweight losses were in milk products and sugar manufacturing, which primarily use import quotas to raise domestic prices. The gain in producer surplus is \$66.2 billion, or 9.5% of domestic consumption. The government obtained \$2.7 billion in tariff revenues, or 0.4% of consumption. If all of producer surplus and government revenues were expended in rent-seeking behavior and other wasteful activities, the total loss would be \$68.9 billion, or 12.5% of consumption, which is 4.75 times larger than the deadweight loss alone. In other words, the loss to society is somewhere between the deadweight loss of \$18.3 billion and \$87.2 billion.

Challenge Solution

Liquor Licenses

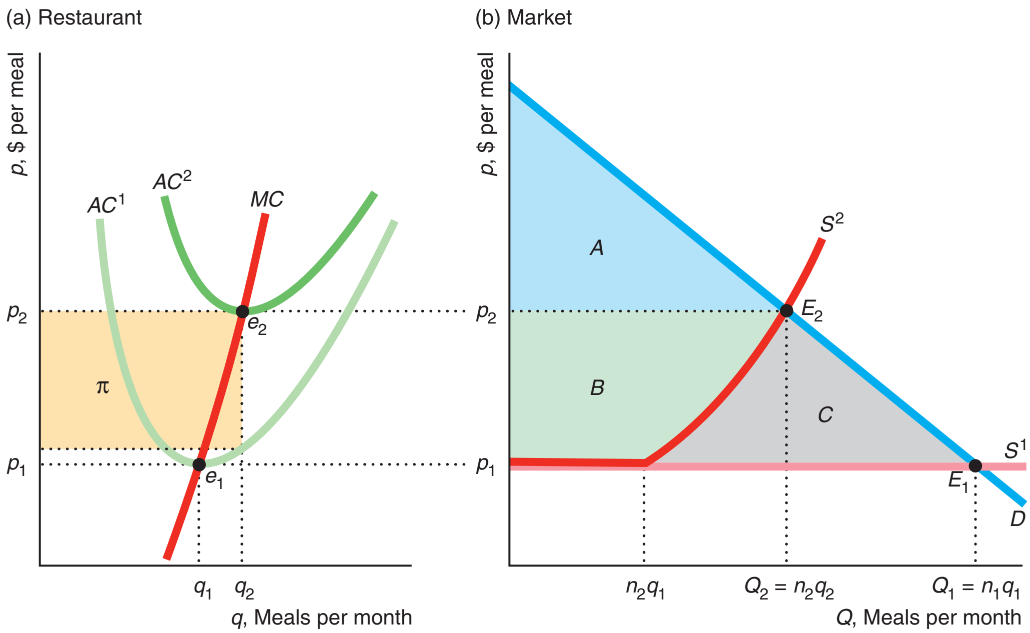
We can now answer the Challenge questions from the beginning of the chapter: What effect does setting a quota on the number of liquor licenses have on the price of meals (including liquor)? What determines the value of a license? How much profit beyond the cost of the license can a restaurant earn? Who benefits and who loses from limiting the number of liquor licenses?

By limiting the number of liquor licenses, the government causes the supply curve of restaurant meals to shift leftward or become steeper. As a result, the equilibrium price of a meal rises and the equilibrium quantity falls. The quota harms consumers: They do not buy as much as they would at lower prices. Restaurants that are in the market when the government first imposes the limits benefit from higher profits.

For simplicity, we'll assume that all restaurants have identical costs and produce identical meals. Panel a of the figure shows a typical restaurant owner's marginal cost curve, MC, and average cost curve, AC^1. Without a quota on the number of liquor licenses, a virtually unlimited number of potential restaurants could enter the market freely. As result, the long-run supply curve of restaurant meals, S^1 in panel b, is horizontal at the minimum of AC^1 (Chapter 8).

Given the market demand curve in the figure, the equilibrium is E_1 where the equilibrium price, p_1, equals the minimum of AC^1 of a typical restaurant. The total number of meals is $Q_1 = n_1 q_1$, where n_1 is the equilibrium number of restaurants and q_1 is the number of meals per month provided by a restaurant.

[24] This argument is made in Tullock (1967) and Posner (1975). Fisher (1985) and Varian (1989) argue that the expenditure is typically less than the producer surplus.



	No Limit	Limit	Change
Consumer Surplus, CS	$A + B + C$	A	$-B - C = \Delta CS$
Producer Surplus, PS	0	B	$B = \Delta PS$
Welfare, $W = CS + PS$	$A + B + C$	$A + B$	$-C = \Delta W = DWL$

Consumer surplus, $A + B + C$, is the area under the market demand curve above p_1 up to Q_1. Restaurants receive no producer surplus because the supply curve is horizontal at the market price, which equals marginal and average cost. Thus, welfare is the same as consumer surplus.

If the government issues only n_2 licenses, the number of restaurants is only $n_2 < n_1$. The market supply curve, S^2, is the horizontal sum of the marginal cost curves above the corresponding minimum average cost curves of the n_2 restaurants in the market. For the market to produce more than n_2q_1 meals, the price must rise to induce the n_2 restaurants to supply more.

With the same demand curve as before, the equilibrium market price rises to p_2. At this higher price, each restaurant produces more meals, $q_2 > q_1$, but the total number of meals, $Q_2 = n_2q_2$, falls because the number of restaurants, n_2, drops. Consumer surplus is A, producer surplus is B, and welfare is $A + B$.

Thus, because of the higher prices under a quota system, consumer surplus falls: $\Delta CS = -B - C$. The producer surplus of the lucky license holders rises by $\Delta PS = B$. As a result, total welfare falls:

$$\Delta W = \Delta CS + \Delta PS = (-B - C) + B = -C,$$

where C is the deadweight loss.

If a state prevents new restaurants from entering the market by limiting liquor licenses, it creates economic profit, the area labeled π in panel a, for each license holder. If a license holder may sell a license, the owner of the scarce resource—the license—can capture this unusual profit. The license sells at a price that captures the current value of all future profits. The government causes the license to have this value by creating an artificial scarcity of licenses. The new owner's average cost rises to AC^2. Because the fee is a fixed cost that is unrelated to output, it does not affect the marginal cost. The new owner earns zero economic profits because the market price, p_2, equals the minimum of AC^2. The producer surplus, B, created by the limits on entry goes to the original owners of the licenses rather than to the current owners. Thus, the original license holders are the *only* ones who benefit from the restrictions, and their gains are less than the losses to others.

Summary

1. **Zero Profit for Competitive Firms in the Long Run.** Although competitive firms may make profits or losses in the short run, they earn zero economic profit in the long run. If necessary, the prices of scarce inputs adjust to ensure that competitive firms make zero long-run profit. Because profit-maximizing firms just break even in the long run, firms that do not try to maximize profits will lose money. Competitive firms must maximize profit to survive.
2. **Consumer Welfare.** The pleasure a consumer receives from a good in excess of its cost is called *consumer surplus*. Consumer surplus equals the area under the consumer's demand curve above the market price up to the quantity that the consumer buys. How much consumers are harmed by an increase in price is measured by the change in consumer surplus.
3. **Producer Welfare.** A firm's gain from trading is measured by its *producer surplus*. Producer surplus is the largest amount of money that could be taken from a firm's revenue and still leave the firm willing to produce. That is, the producer surplus is the amount the firm is paid minus its variable cost of production, which is profit in the long run. It is the area below the price and above the supply curve up to the quantity that the firm sells. The effect of a change in a price on a supplier is measured by the change in producer surplus.
4. **Competition Maximizes Welfare.** One standard measure of welfare is the sum of consumer surplus and producer surplus. The more price exceeds marginal cost, the lower this measure of welfare. In the competitive equilibrium, in which price equals marginal cost, welfare is maximized.
5. **Policies That Shift Supply and Demand Curves.** Governments frequently limit the number of firms in a market directly, by licensing them, or indirectly, by raising the costs of entry to new firms or raising the cost of exiting. A reduction in the number of firms in a competitive market raises price, hurts consumers, helps producing firms, and lowers the standard measure of welfare. This reduction in welfare is a deadweight loss: The gain to producers is less than the loss to consumers.
6. **Policies That Create a Wedge Between Supply and Demand.** Taxes, price ceilings, and price floors create a gap between the price consumers pay and the price firms receive. These policies force price above marginal cost, which raises the price to consumers and lowers the amount consumed. The wedge between price and marginal cost results in a deadweight loss: The loss of consumer surplus and producer surplus is not offset by increased taxes or by benefits to other groups.
7. **Comparing Both Types of Policies: Imports.** A government may use either a quantity restriction such as a quota, which shifts the supply curve, or a tariff, which creates a wedge, to reduce imports or achieve other goals. These policies may have different welfare implications. A tariff that reduces imports by the same amount as a quota has the same harms—a larger loss of consumer surplus than increased domestic producer surplus—but has a partially offsetting benefit—increased tariff revenues for the government. Rent-seeking activities are attempts by firms or individuals to influence a government to adopt a policy that favors them. By using resources, rent seeking exacerbates the welfare loss beyond the deadweight loss caused by the policy itself. In a perfectly competitive market, government policies frequently lower welfare. However, as we show in later chapters, government policies may increase welfare in markets that are not perfectly competitive.

Questions

Select questions are available on MyLab Economics*; * = answer appears at the back of this book;* **A** *= algebra problem.*

1. Zero Profit for Competitive Firms in the Long Run

1.1 Only a limited amount of high-quality wine-growing land is available. The firms that farm the land are identical. Because the demand curve hits the market supply curve in its upward-sloping section, the firms initially earn positive profit.

a. The owners of the land charge a higher rent so as to capture the profit. Show how the market supply curve changes (if at all).

b. Some farming firms own the land and some rent. Does this difference affect their shutdown decision or other decisions?

1.2 The Application "What's a Name Worth?" notes that a successful entertainer such as Taylor Swift earns astronomical amounts. Other entertainers earn just a relatively small wage. Discuss why Taylor Swift's earnings consist of the basic wage plus a rent.

2. Consumer Welfare

*2.1 If the inverse demand function for toasters is $p = 60 - Q$, what is the consumer surplus if price is 30? **A**

2.2 If the inverse demand function for radios is $p = a - bQ$, what is the consumer surplus if price is $a/2$? **A**

2.3 Use the numbers for the alcohol and tobacco category from the table in the Application "Goods with a Large Consumer Surplus Loss from Price Increases" to draw a figure that illustrates the roles that the revenue and the elasticity of demand play in determining the loss of consumer surplus due to an increase in price. Indicate how the various areas of your figure correspond to the equation derived in Appendix 9A and the discussion in this chapter about how a more elastic demand curve would affect consumer surplus. (*Hint*: See Solved Problem 9.1.) **A**

2.4 Go to **eBay.com** or another auction site. Find the bidding record from a completed (or nearly completed) auction on a type of good that interests you.

a. Create a table showing each bidder's maximum bid. (*Note*: Some people may have bid more than once. Use only each person's highest bid.)

b. Use the table to plot the demand curve for the item in this auction.

c. Determine the bidders' total willingness to pay for the good. (*Hint*: see the Application "Willingness to Pay on eBay.")

3. Producer Welfare

3.1 Does a firm's *producer surplus* differ from its *profit* if it has no fixed cost?

3.2 If the supply function is $Q = 10 + p$, what is the producer surplus if price is 20? (*Hint*: See Solved Problem 9.2.) **A**

4. Competition Maximizes Welfare

4.1 If society cares only about the well-being of consumers so that it wants to maximize consumer surplus, would a competitive market achieve that goal given that the government cannot force or bribe firms to produce more than the competitive level of output? How would your answer change if society cares only about maximizing producer surplus? (*Hint*: See the discussion of Figure 9.5 and Solved Problem 9.3.)

4.2 Use an indifference curve diagram (gifts on one axis and all other goods on the other) to illustrate that one is better off receiving cash than a gift. (*Hint*: See the discussion of gifts in this chapter and the discussion of food stamps in Chapter 4.) Relate your analysis to the "Deadweight Loss of Christmas Presents" Application.

4.3 An April 2014 study by researchers at University College London found that people who ate seven or more portions of fruits and vegetables a day had a 42% lower risk of death at any point in time compared to those who ate less than one portion. The research also showed that vegetables have significantly higher health benefits than fruits (UCL News, 2014; Oyebode et al., 2014). Use a diagram to show what would happen to equilibrium price and quantity if consumers followed that advice and increased their daily consumption of vegetables. Illustrate as well how consumer surplus and producer surplus would change.

5. Policies That Shift Supply and Demand Curves

5.1 The government imposes a restriction on firms that shifts the supply curve in Figure 9.5 so that it intersects the demand curve at e_2. Discuss the effects on CS, PS, welfare, and DWL. (*Hint*: Use the original supply curve in determining the welfare effects.)

5.2 Airport owners grant airlines the right to schedule a landing or departure during a specific period of time in accordance with guidelines set out by the International Air Transport Association. Once allocated, landing slots can be traded between airlines. It was reported in February 2016, for example, that Oman Air paid Air France-KLM \$75 million for a landing slot at London, Heathrow. What are the welfare implications of landing slot agreements for new entrants and for existing right holders?

5.3 The consumption of alcohol is illegal for Afghan citizens, but the government does allow the sale of alcohol under licence to foreigners. Use a diagram to show how such a prohibition would affect the price of alcohol. Who wins and who loses? Consider the effects of the ban on consumers, producers, and society.

5.4 The Application "Welfare Effects of Allowing Fracking" says that allowing fracking resulted in a loss in producer surplus and an increase in consumer surplus and welfare in the natural gas market. Illustrate these results in a figure.

5.5 Suppose that the inverse market demand for silicone replacement tips for Sony EX71 earbud headphones is $p = p_N - 0.1Q$, where p is the price per pair of replacement tips, p_N is the price of a new pair of headphones, and Q is the number of tips per week. Suppose that the inverse supply function of the replacement tips is $p = 2 + 0.012Q$.

a. Find the effect of a change in the price of a new pair of headphones on the equilibrium price of replacement tips at the equilibrium, $\mathrm{d}p/\mathrm{d}p_N$.

b. If $p_N = \$30$, what are the equilibrium p and Q? What is the consumer surplus? What is the producer surplus? **A**

6. Policies That Create a Wedge Between Supply and Demand

6.1 If the inverse demand function for books is $p = 60 - Q$ and the supply function is $Q = p$, what is the initial equilibrium? What is the welfare effect of a specific tax of $t = \$2$? **A**

6.2 Suppose that the demand curve for wheat is $Q = 100 - 10p$ and the supply curve is $Q = 10p$. The government imposes a specific tax of $t = 1$ per unit.

a. How do the equilibrium price and quantity change?

b. What effect does this tax have on consumer surplus, producer surplus, government revenue, welfare, and deadweight loss? **A**

6.3 What is the welfare effect of an ad valorem sales tax, v, assessed on each competitive firm in a market?

6.4 How would the analysis in Solved Problem 9.4 change if the supply curve were upward-sloping instead of horizontal?

6.5 In Solved Problem 9.4, what is the relationship between lost consumer surplus due to the tax, deadweight loss, and tax revenue? Discuss and reconcile the different results in Solved Problems 9.1 and 9.4.

*6.6 What is the welfare effect of a lump-sum tax assessed on each competitive firm in a market? (*Hint*: See the Challenge Solution in Chapter 8.)

*6.7 What is the long-run welfare effect of a profit tax (the government collects a specified percentage of a firm's profit) assessed on each competitive firm in a market?

6.8 The government wants to drive the price of soybeans above the equilibrium price, p_1, to p_2. It offers growers a lump-sum payment to reduce their output from Q_1 (the equilibrium level) to Q_2, which is the quantity demanded by consumers at p_2. Use a figure to show how large this payment must be (an area in the figure) for growers to reduce output to this level. What are the effects of this program on consumers, farmers, and total welfare? Compare this approach to (a) offering a price support of p_2, (b) offering a price support and a quota set at Q_1, and (c) offering a price support and a quota set at Q_2.

6.9 Suppose that the demand curve for wheat is $Q = 100 - 10p$ and the supply curve is $Q = 10p$. The government provides producers with a specific subsidy of $s = 1$ per unit.

a. How do the equilibrium price and quantity change?

b. What effect does this tax have on consumer surplus, producer surplus, government revenue, welfare, and deadweight loss? (*Hint*: See Solved Problem 9.5.) **A**

6.10 Use diagrams to compare the welfare implications of the traditional agricultural price support program and the deficiency payment program if both set the same price floor, $\underline{p}$. Under what circumstances would farmers, consumers, or taxpayers prefer one program to the other? (*Hint*: See Solved Problem 9.6.)

*6.11 Suppose that the demand curve for wheat is $Q = 100 - 10p$ and the supply curve is $Q = 10p$. The government imposes a price support at $\underline{p} = 6$ using a deficiency payment program.

a. What are the quantity supplied, the price that clears the market, and the deficiency payment?

b. What effect does this program have on consumer surplus, producer surplus, welfare, and deadweight loss? (*Hint*: See Solved Problem 9.7 and the Application "The Social Cost of a Natural Gas Price Ceiling.") **A**

6.12 The government sets a minimum wage above the current equilibrium wage. What effect does the minimum wage have on the market equilibrium? What are its effects on consumer surplus, producer surplus, and total surplus? Who are the consumers and who are the producers? (*Hint*: See the treatment of a ceiling in Solved Problem 9.7.)

6.13 A mayor wants to help renters in her city. She considers two policies that will benefit renters equally. One policy is *rent control*, which places a price ceiling, $\bar{p}$, on rents. The other is a government housing subsidy of s dollars per month that lowers the amount renters pay (to $\bar{p}$). Who benefits and who loses from these policies? Compare the two policies' effects on the quantity of housing consumed, consumer surplus, producer surplus, government expenditure, and deadweight loss. Does the comparison of deadweight loss depend on the elasticities of supply and demand? (*Hint*: Consider extreme cases.) If so, how? (*Hint*: See Solved Problem 9.7.)

6.14 Suppose that the demand curve is $Q = 100 - 10p$ and the supply curve is $Q = 10p$. The government imposes a price ceiling of $p = 3$.

a. Describe how the equilibrium changes.

b. What effect does this ceiling have on consumer surplus, producer surplus, and deadweight loss? (*Hint*: See Solved Problem 9.7.) **A**

6.15 Australia has over 500 national parks in which commercial activities are prohibited and human activity is strictly monitored. In the interests of visitor safety or environmental sustainability, access to national parks may be restricted at certain times. Suppose the Australian National Parks and Wildlife Service decides to limit the number of visitors to Bald Rock National Park in New South Wales to Q^*, which is less than the current number of visitors to the park Q_0. It considers two policies for accomplishing this: (a) raising the park entry fee to P_1 and (b) setting a quota on the number of visitors. Assuming the marginal cost of a visit is constant and equal to the current price of admission, P_0, compare the effects of these two policies on consumer surplus, producer surplus, and welfare. Use a diagram to show which policy is superior according to the welfare criterion.

6.16 Firms and consumer groups often try to influence government to adopt a policy that favors them. Others who do not support lobbying arguing that, left alone, competitive markets maximize societal welfare. Which of the following policies would each of these three groups favor: a price ceiling, a price floor, a quota on the supply of a good, or no intervention? (*Hint*: Consider how the shapes of the supply and demand curves might affect your answers.)

6.17 At the unregulated, competitive equilibrium, the demand curve of gasoline becomes more elastic over time as people can react to a higher gasoline price by buying more fuel-efficient cars, moving closer to work, or making other changes than they cannot make in the short run. Given a binding price ceiling, the gasoline market has a shortage. As the demand curve becomes less elastic in the long run, what happens to the size of the shortage and the deadweight loss?

6.18 Milk quotas in the European Union were abolished on April 1, 2015. They had been introduced in 1984 under the Common Agricultural Policy to manage excess supply. What are the effects of a limit on the amount of a good that competitive firms may sell on consumer surplus, producer surplus, and welfare? Who benefits from such a rule?

6.19 Use a figure to illustrate why technological progress (which causes the supply curve to shift rightward) hurts the government and consumers in an agricultural market in which the government maintains a traditional price support as in Figure 9.8. Compare that result with the effects of technological progress if the government uses the alternative price support in Solved Problem 9.6.

7. Comparing Both Types of Policies: Imports

7.1 Show that if the importing country faces an upward-sloping foreign supply curve (excess supply curve), a tariff may raise welfare in the importing country.

7.2 Given that the world supply curve is horizontal at the world price for a given good, can a subsidy on imports raise welfare in the importing country? Explain your answer.

7.3 The WTO agreement allows members to apply anti-dumping measures when dumping (selling at a price less than the cost of production) by one country that may result in material injury to the competing domestic industry in the member country. Regardless, many economists criticize anti-dumping measures, and some advocate their complete elimination. If a producer in one country exports its product at such an unfair price, who wins and who loses in the importing country? What would be the effect on consumers, producers, the government, and society of a tariff that removes the injury to domestic industry in the importing country (that is, it restores the market price to its pre-dumping level)?

7.4 Contracting parties to the General Agreement on Tariffs and Trade (1947–94) could use quotas to limit imports of agricultural products from other countries. Under the Uruguay Round (1986–94), import quotas were replaced by tariffs (a process referred to as "tariffication"). Use diagrams to compare the welfare implications of an import quota versus an equivalent tariff on consumers, producers, and the government of the importing country. Use another diagram to show how welfare in the market for a substitute would be affected.

7.5 A government is considering a quota or a tariff, both of which will reduce imports by the same amount. Which does the government prefer, and why? Explain how your answer depends on the way that the quota is allocated.

7.6 Economic sanctions are sometimes imposed on one country by another country (or a group of countries) to achieve an economic, political, or social objective. They may take a variety of forms, including a ban on trade with the country being targeted, and it can be quite costly for the country that imposes these sanctions. Use diagrams to show the welfare effects on the country that imposes a ban on imports of a good from the targeted country and exports of a different good to the targeted country. Assume the ban eliminates all trade in the two markets.

7.7 After Mexico signed the North American Free Trade Agreement (NAFTA) with the United States in 1994, corn imports from the United States doubled within a year, and, in some recent years, U.S. imports have approached half of the amount of corn consumed in Mexico. According to Oxfam (2003), the price of Mexican corn fell more than 70% in the first decade after NAFTA took effect. Part of the reason for this flow south of our border is that the U.S. government subsidizes corn production to the tune of $10 billion a year. According to Oxfam, the 2002 U.S. cost of production was $3.08 per bushel, but the export price was $2.69 per bushel, with the difference reflecting an export subsidy of 39¢ per bushel. The U.S. exported 5.3 million metric tons. Use graphs to show the effect of such a subsidy on the welfare of various groups and on government expenditures in the United States and Mexico.

7.8 Canada has 20% of the world's known freshwater resources, yet many Canadians believe that the country has little or none to spare. Over the years, U.S. and Canadian firms have struck deals to export bulk shipments of water to drought-afflicted U.S. cities and towns. Provincial leaders have blocked these deals in British Columbia and Ontario. Use graphs to show the likely outcome of such barriers to exports on the price and quantity of water used in Canada and in the United States if markets for water are competitive. Show the effects on consumer and producer surplus in both countries.

7.9 The U.S. Supreme Court ruled in May 2005 that people can buy wine directly from out-of-state wineries. Previously, some states had laws that required people to buy directly from wine retailers located in the state.

a. Suppose the market for wine in New York is perfectly competitive both before and after the Supreme Court decision. Use the analysis in Section 9.7 to evaluate the effect of the Court's decision on the price of wine in New York.

b. Evaluate the increase in New York consumer surplus.

c. How does the increase in consumer surplus depend on the price elasticity of supply and demand?

8. Challenge

8.1 A city may limit the number of liquor licenses for restaurants in many ways. It could issue a license that the owner keeps forever and that can be resold. Or, it could charge a high license fee each year, which is equivalent to the city's issuing a license that lasts only a year. A third option is to charge a daily tax on restaurants that serve liquor. Use figures to compare and contrast the equilibrium under each of these approaches. Discuss who wins and who loses from each plan, considering consumers, drivers, the city, and (if relevant) medallion owners.

General Equilibrium and Economic Welfare 10

Let the good of the people be the supreme law. —Cicero

Challenge

Anti-Price Gouging Laws

After a disaster strikes, prices tend to rise. The average U.S. gasoline price increased by 46¢ per gallon after Hurricane Katrina in 2005 damaged most Gulf Coast oil refineries. Many state governments enforce anti-price gouging laws to prevent prices from rising, while prices may be free to adjust in neighboring states. For example, Louisiana's anti-price gouging law went into effect when Governor Bobby Jindal declared a state of emergency in response to the 2010 BP oil spill that endangered Louisiana's coast.

On average, gasoline prices rose by a few cents immediately after Superstorm Sandy in October 2012; however, some stations increased the retail markup over the wholesale prices by up to 135%. The New York Attorney General's office received over 500 consumer complaints about price gouging within a week of the storm. The attorney general pursued price gouging cases against 25 gas stations. States that declared storm emergencies that triggered anti-price gouging laws include Virginia in 2014; Kentucky, New York, Pennsylvania, and West Virginia in 2015; and Oklahoma in 2016.

The District of Columbia and 34 states have anti-price gouging laws. Arkansas, California, Maine, New Jersey, Oklahoma, Oregon, and West Virginia have set a "percentage increase cap limit" on price increases after a disaster, ranging from 10% to 25% of the price before the emergency. Sixteen states prohibit "unconscionable" price increases. Connecticut, Georgia, Hawaii, Kentucky, Louisiana, Mississippi, and Utah have outright bans on price increases during an emergency.

Generally, legislatures pass these laws after a major natural disaster.[1] California passed its law in 1994 after the Northridge earthquake. Georgia enacted its anti-price gouging statute after a 500-year flood in 1994. Consequently, often a state hit by a recent disaster has such a law while a neighboring state does not.

In Chapter 2, we showed that a national price control causes shortages. However, does a binding price control that affects one state, but not a neighboring state, cause shortages? How does it affect prices and quantities sold in the two states? Which consumers benefit from these laws?

---

[1]Governments pass anti-price gouging laws because they are popular. After the post-Katrina gas price increases, an ABC News/*Washington Post* poll found that only 16% of respondents thought that the price increase was "justified," 72.7% thought that "oil companies and gas dealers are taking unfair advantage," 7.4% said both views were true, and the rest held another or no opinion.

In addition to natural disasters, a change in government policies or other shocks often affect equilibrium price and quantity in more than one market. To determine the effects of such a change, we examine the interrelationships among markets.

In this chapter, we extend our analysis of equilibrium in a single market to equilibrium in all markets. We then examine how a society decides whether a particular equilibrium (or change in equilibrium) in all markets is desirable. To do so, society must answer two questions: "Is the equilibrium efficient?" and "Is the equilibrium equitable?"

For the equilibrium to be efficient, both consumption and production must be efficient. Production is efficient only if it is impossible to produce more output at current cost given current knowledge (Chapter 7). Consumption is efficient only if goods cannot be reallocated across people so that at least someone is better off and no one is harmed. In this chapter, we show how to determine whether consumption is efficient.

Whether the equilibrium is efficient is a scientific question. It is possible that all members of society could agree on how to answer scientific questions concerning efficiency.

To answer the equity question, society must make a value judgment as to whether each member of society has his or her "fair" or "just" share of all the goods and services. A common view in individualistic cultures is that each person is the best—and possibly the only legitimate—judge of his or her own welfare. Nonetheless, to make social choices about events that affect more than one person, we have to make interpersonal comparisons, through which we decide whether one person's gain is more or less important than another person's loss. For example, in Chapter 9 we argued that a price ceiling lowers a measure of total welfare given the value judgment that the well-being of consumers (consumer surplus) and the well-being of the owners of firms (producer surplus) should be weighted equally. People of goodwill—and others—may disagree greatly about equity issues.

As a first step in studying welfare issues, many economists use a narrow value criterion, called the *Pareto principle* (after an Italian economist, Vilfredo Pareto), to rank different allocations of goods and services for which no interpersonal comparisons need to be made. According to this principle, a change that makes one person better off without harming anyone else is desirable. An allocation of goods and services is **Pareto efficient** if any possible reallocation would harm at least one person.

Pareto efficient
describing an allocation of goods and services such that any reallocation harms at least one person

Presumably, you agree that any government policy that makes all members of society better off is desirable. Do you also agree that a policy that makes some members better off without harming others is desirable? What about a policy that helps one group more than it hurts another group? What about a policy that hurts another group more than it helps your group? It is very unlikely that all members of society will agree on how to answer these questions—much less on the answers.

The efficiency and equity questions arise even in small societies, such as your family. Suppose that your family has gathered together in November, and everyone wants pumpkin pie. How much pie you get will depend on the answer to efficiency and equity questions: "How can we make the pie as large as possible with available resources?" and "How should we divide the pie?" It is probably easier to get agreement about how to make the largest possible pie than about how to divide it equitably.

Economists primarily use economic theory to answer scientific efficiency questions because they can do so without making value judgments. To examine equity questions, they must make value judgments, as we did in Chapter 9's welfare analysis. (Strangely, most members of our society seem to believe that economists are no better at making value judgments than anyone else.) In this chapter, we examine various views on equity.

In this chapter, we examine five main topics

1. **General Equilibrium.** The effects of a new government policy or other shock differ if the shock affects several markets rather than just one.
2. **Trading Between Two People.** Where two people have goods but cannot produce more goods, both parties benefit from mutually agreed trades.
3. **Competitive Exchange.** The competitive equilibrium has two desirable properties: Any competitive equilibrium is Pareto efficient, and any Pareto-efficient allocation can be obtained by using competition, given an appropriate income distribution.
4. **Production and Trading.** The benefits from trade continue to hold when production is introduced.
5. **Efficiency and Equity.** A society uses its views about equity to choose among Pareto-efficient allocations.

10.1 General Equilibrium

partial-equilibrium analysis
an examination of equilibrium and changes in equilibrium in one market in isolation

So far we have used a **partial-equilibrium analysis:** an examination of equilibrium and changes in equilibrium in one market in isolation. In a partial-equilibrium analysis in which we hold the prices and quantities of other goods fixed, we implicitly ignore the possibility that events in this market affect other markets' equilibrium prices and quantities.

When stated this baldly, partial-equilibrium analysis sounds foolish. However, it needn't be. Suppose that the government puts a tax on hula hoops. If the tax is sizable, it will dramatically affect the sales of hula hoops. Even a very large tax on hula hoops is unlikely to affect the markets for automobiles, doctor services, or orange juice. Indeed, it is unlikely to affect the demand for other toys greatly. Thus, a partial-equilibrium analysis of the effect of such a tax should serve us well. Studying all markets simultaneously to analyze this tax would be unnecessary at best and confusing at worst.

general-equilibrium analysis
the study of how equilibrium is determined in all markets simultaneously

Sometimes, however, we need to use a **general-equilibrium analysis**: the study of how equilibrium is determined in all markets simultaneously. For example, the discovery of a major oil deposit in a small country raises the income of its citizens, and the increased income affects all that country's markets. Sometimes economists model many markets in an economy and solve for the general equilibrium in all of them simultaneously, using computer models.

Frequently, economists look at equilibrium in several—but not all—markets simultaneously. We would expect a tax on comic books to affect the price of comic books, which in turn would affect the price of video games because video games are substitutes for comics for some people. But we would not expect a tax on comics to have a measurable effect on the demand for washing machines. Therefore, it is reasonable to conduct a multimarket analysis of the effects of a tax on comics by looking only at the markets for comics, video games, and a few other closely related markets such as those for movies and trading cards. That is, a multimarket-equilibrium analysis covers the relevant markets, but not all markets, as a general equilibrium analysis would.

Markets are closely related if an increase in the price in one market causes the demand or supply curve in another market to shift measurably. Suppose that a tax on coffee causes the price of coffee to increase. The rise in the price of coffee causes the demand curve for tea to shift outward (more is demanded at any given price of tea) because tea and coffee are substitutes. The coffee price increase also causes the demand curve for cream to shift inward because coffee and cream are complements.

Similarly, supply curves in different markets may be related. If a farmer produces corn and soybeans, an increase in the price of corn will affect the relative amounts of both crops the farmer chooses to produce.

Markets may also be linked if the output of one market is an input in another market. A shock that raises the price of computer chips will also raise the price of computers.

Thus, an event in one market may have a spillover effect on other related markets for various reasons. Indeed, a single event may initiate a chain reaction of spillover effects that reverberates back and forth between markets.

Feedback Between Competitive Markets

To illustrate the feedback of spillover effects between markets, we examine the corn and soybean markets using supply and demand curves estimated by Holt (1992). Consumers and producers substitute between corn and soybeans, so the supply and demand curves in these two markets are related. The quantity of corn demanded and the quantity of soybeans demanded both depend on the price of corn, the price of soybeans, and other variables. Similarly, the quantities of corn and soybeans supplied depend on their relative prices.

We can demonstrate the effect of a shock in one market on both markets by tracing the sequence of events in the two markets. Whether these steps occur nearly instantaneously or take some time depends on how quickly consumers and producers react.

The initial supply and demand curves for corn, S_0^c and D_0^c, intersect at the initial equilibrium for corn, e_0^c, in panel a of Figure 10.1.[2] The price of corn is \$2.15 per bushel, and the quantity of corn is 8.44 billion bushels per year. The initial supply and demand curves for soybeans, S_0^s and D_0^s, intersect at e_0^s in panel b, where price is \$4.12 per bushel and quantity is 2.07 billion bushels per year. The first row of Table 10.1 shows the initial equilibrium prices and quantities in these two markets.

Now suppose that the foreign demand for U.S. corn decreases, causing the export of corn to fall by 10% and the total U.S. demand for corn to shift from D_0^c to D_1^c in panel a. The new equilibrium is at e_1^c, where D_1^c intersects S_0^c. The price of corn falls by nearly 11% to \$1.9171 per bushel, and the quantity falls 2.5% to 8.227 billion bushels per year, as the Step 1 row of the table shows.

If we were conducting a partial-equilibrium analysis, we would stop here. In a general-equilibrium analysis, however, we next consider how this shock to the corn market affects the soybean market. Because this shock initially causes the price of corn to fall relative to the price of soybeans (which stays constant), consumers substitute toward corn and away from soybeans: The demand curve for soybeans shifts to the left from D_0^s to D_2^s in panel b.

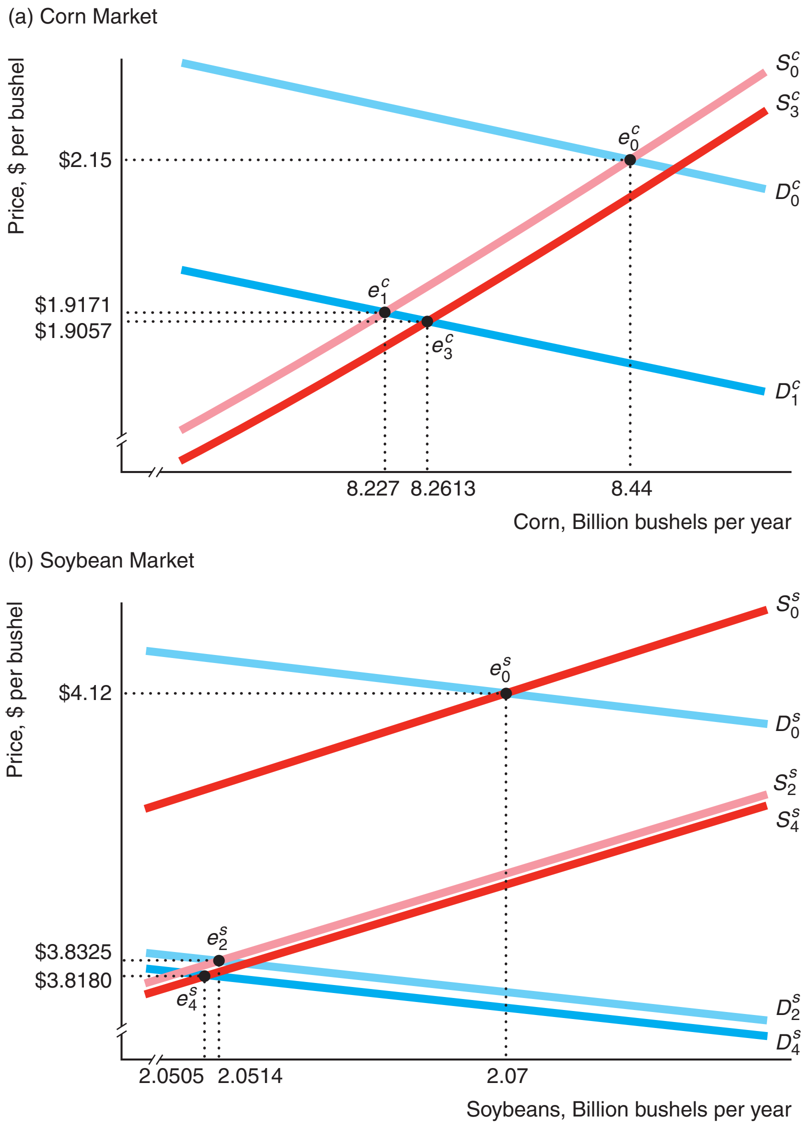
In addition, because the price of corn falls relative to the price of soybeans, farmers produce more soybeans at any given price of soybeans: The supply curve for soybeans shifts outward to S_2^s. The new soybean demand curve, D_2^s, intersects the new soybean supply curve, S_2^s, at the new equilibrium e_2^s, where price is \$3.8325 per bushel, a fall of 7%, and quantity is 2.0514 billion bushels per year, a drop of less than 1% (Step 2 in the table).

As it turns out, this fall in the price of soybeans relative to the price of corn causes essentially no shift in the demand curve for corn (panel a shows no shift) but shifts the supply curve of corn, S_3^c, to the right. The new equilibrium is e_3^c, where S_3^c and

[2]Until recently, the corn and soybean markets were subject to price controls (Chapter 9). However, we use the estimated supply and demand curves to ask what would happen in these markets in the absence of price controls.

Figure 10.1 Relationship Between the Corn and Soybean Markets MyLab Economics Video

Supply and demand curves in the corn and soybean markets (as estimated by Holt, 1992) are related.



D_1^c intersect. Price falls to \$1.9057 per bushel of corn and quantity to 8.2613 billion bushels per year (Step 3).

This new fall in the relative price of corn causes the soybean demand curve, D_4^s, to shift farther to the left and the supply curve, S_4^s, to shift farther to the right in panel b. At the new equilibrium at e_4^s, where D_4^s and S_4^s intersect, the price and quantity of soybeans fall slightly to \$3.818 per bushel and 2.0505 billion bushels per year, respectively (Step 4).

These reverberations between the markets continue, with additional smaller shifts of the supply and demand curves. Eventually, a final equilibrium is reached at which none of the supply and demand curves will shift further. The final equilibria in these two markets (last row of Table 10.1) are virtually the same as e_3^c in panel a and e_4^s in panel b.

Table 10.1 Adjustment in the Corn and Soybean Markets

	Corn		Soybeans	
Step	*Price*	*Quantity*	*Price*	*Quantity*
Initial (0)	2.15	8.44	4.12	2.07
1	1.9171	8.227		
2			3.8325	2.0514
3	1.9057	8.2613		
4			3.818	2.0505
5	1.90508	8.26308		
6			3.81728	2.05043
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
Final	1.90505	8.26318	3.81724	2.05043

If we were interested only in the effect of the shift in the foreign demand curve on the corn market, would we rely on a partial-equilibrium analysis? According to the partial-equilibrium analysis, the price of corn falls 10.8% to \$1.9171. In contrast, in the general-equilibrium analysis, the price falls 11.4% to \$1.905, which is 1.2¢ less per bushel. Thus, the partial-equilibrium analysis underestimates the price effect by 0.6 percentage points. Similarly, the fall in quantity is 2.5% according to the partial-equilibrium analysis and only 2.1% according to the general-equilibrium analysis. In this market, then, the biases from using a partial-equilibrium analysis are small.[3]

Solved Problem 10.1

MyLab Economics Solved Problem

Because many consumers choose between coffee and tea, the coffee and tea demand functions depend on both prices. Suppose the demand curves for coffee and tea are

$$Q_c = 120 - 2p_c + p_t,$$
$$Q_t = 90 - 2p_t + p_c,$$

where Q_c is the quantity of coffee, Q_t is the quantity of tea, p_c is the price of coffee, and p_t is the price of tea. These crops are grown in separate parts of the world, so their supply curves are not interrelated. We assume that the short-run, inelastic supply curves for coffee and tea are $Q_c = 45$ and $Q_t = 30$. Solve for the equilibrium prices and quantities. Now suppose that a freeze shifts the short-run supply curve of coffee to $Q_c = 30$. How does the freeze affect the prices and quantities?

Answer

1. *Equate the quantity demanded and supplied for both markets.* Equating the right sides of the coffee demand and supply functions, we obtain $120 - 2p_c + p_t = 45$, or $p_t = 2p_c - 75$. For the tea market, $90 - 2p_t + p_c = 30$, or $p_c = 2p_t - 60$. That leaves us with two equations and two unknowns, p_t and p_c.

[3]For an example where the bias from using a partial-equilibrium analysis instead of a general-equilibrium analysis is large, see MyLab Economics, Chapter 10, "Sin Taxes."

2. *Substitute the expression for p_t from the coffee equation into the tea equation and solve for the price of coffee, then use that result to obtain p_t.* By substituting $p_t = 2p_c - 75$ into $p_c = 2p_t - 60$, we find that $p_c = 4p_c - 150 - 60$. Solving this expression for p_c, we find that $p_c = 70$. Substituting $p_c = 70$ into the coffee equation, we learn that $p_t = 2p_c - 75 = 140 - 75 = 65$. Substituting these prices into the demand equations, we confirm that the equilibrium quantities equal the fixed supplies: $Q_c = 45$ and $Q_t = 30$.

3. *Repeat the analysis for $Q_c = 30$.* The new quantity changes the coffee market equilibrium condition to $120 - 2p_c + p_t = 30$, or $p_t = 2p_c - 90$. Substituting this expression into the tea equilibrium condition, we find that $p_c = 4p_c - 180 - 60$, so $p_c = 240/3 = 80$, and hence $p_t = 2p_c - 90 = 70$. Thus, the price of coffee rises by 10 and the price of tea by 5 in response to the coffee freeze, which reduces Q_c by 15 while leaving Q_t unaffected.

Minimum Wages with Incomplete Coverage

We used a partial-equilibrium analysis in Chapter 2 to examine the effects of a minimum wage law that holds throughout the entire labor market. The minimum wage causes the quantity of labor demanded to be less than the quantity of labor supplied. Workers who lose their jobs cannot find work elsewhere, so they become unemployed.

Many people are familiar with that reasoning and overgeneralize:

Common Confusion: A minimum wage must cause unemployment.

This result follows logically if the minimum wage applies to the entire work force. However, the minimum wage may not cause unemployment if the minimum wage law applies to workers in only some sectors of the economy, as we show using a general-equilibrium analysis.[4]

When a minimum wage applies to only a covered sector of the economy, the resulting increase in the wage causes the quantity of labor demanded in that sector to fall. Workers who lose their jobs in the covered sector move to the uncovered sector, driving down the wage in that sector.

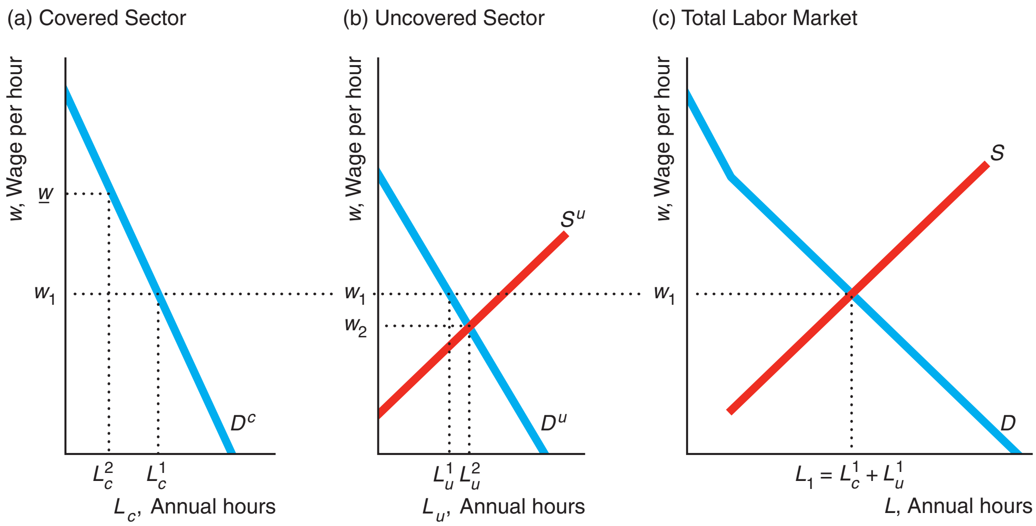
When the U.S. minimum wage law was first passed in 1938, some economists joked that its purpose was to maintain family farms. The law drove workers out of manufacturing and other covered industries into agriculture, which the law did not cover. Over time, the minimum wage law has covered more and more sectors of the economy.

Figure 10.2 shows the effect of a minimum wage law when coverage is incomplete. The total demand curve, D in panel c, is the horizontal sum of the demand curve for labor services in the covered sector, D^c in panel a, and the demand curve in the uncovered sector, D^u in panel b. In the absence of a minimum wage law, the wage in both sectors is w_1, which is determined by the intersection of the total demand curve, D, and the total supply curve, S. At that wage, L_c^1 annual hours of work are hired

[4]Also, in a market with a single employer, a *monopsony*, a minimum wage increases employment and does not cause unemployment (see Solved Problem 15.4).

Figure 10.2 Minimum Wage with Incomplete Coverage

In the absence of a minimum wage, the equilibrium wage is w_1. Applying a minimum wage, $\underline{w}$, to only one sector causes the quantity of labor services demanded in the covered sector to fall. The extra labor moves to the uncovered sector, driving the uncovered sector wage down to w_2.



in the covered sector, L_u^1 annual hours in the uncovered sector, and $L_1 = L_c^1 + L_u^1$ in the entire market.

If a minimum wage of $\underline{w}$ is set in only the covered sector, employment in that sector falls to L_c^2. To determine the wage and level of employment in the uncovered sector, we first need to determine how much labor service is available to that sector.

Anyone who can't find work in the covered sector goes to the uncovered sector. The supply curve of labor to the uncovered sector in panel b is a residual supply curve: the quantity the market supplies that is not met by demanders in other sectors at any given wage (Chapter 8). With a binding minimum wage in the covered sector, the residual supply function in the uncovered sector is[5]

$$S^u(w) = S(w) - D^c(\underline{w}).$$

Thus, the residual supply to the uncovered sector, $S^u(w)$, is the total supply, $S(w)$, at any given wage w minus the amount of labor used in the covered sector, $L_c^2 = D^c(\underline{w})$.

The intersection of D^u and S^u determines w_2, the new wage in the uncovered sector, and L_u^2, the new level of employment.[6] This general-equilibrium analysis shows that a minimum wage causes employment to drop in the covered sector, employment to rise (by a smaller amount) in the uncovered sector, and the wage in the uncovered

---

[5]Without a minimum wage, the residual supply curve for the uncovered sector is $S^u(w) = S(w) - D^c(w)$.

[6]This analysis is incomplete if the minimum wage causes the price of goods in the covered sector to rise relative to those in the uncovered sector, which in turn causes the demands for labor in those two sectors, D^c and D^u, to shift. Ignoring that possibility is reasonable if labor costs are a small fraction of total costs (hence the effect of the minimum wage is minimal on total costs) or if the demands for the final goods are relatively price insensitive.

sector to fall below the original competitive level. Thus, a minimum wage law with only partial coverage affects wage levels and employment levels in various sectors but need not create unemployment.

When the U.S. minimum wage was first passed in 1938, only 56% of workers were employed in covered firms (see MyLab Economics, Chapter 10, "U.S. Minimum Wage Laws and Teenagers"). Even today, many state minimum wage laws provide incomplete coverage.

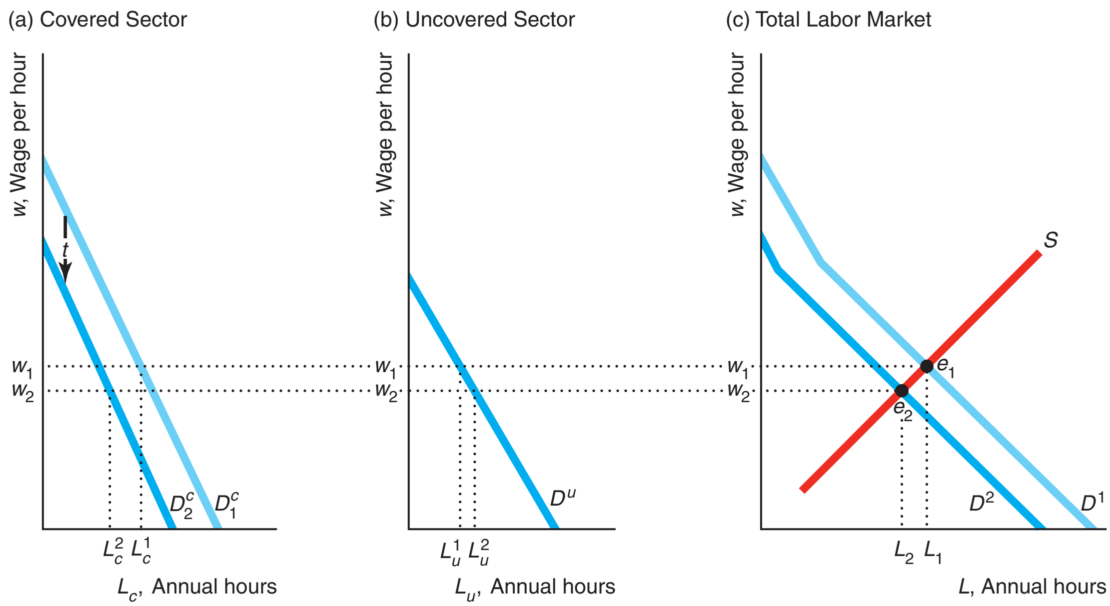
More than 145 U.S. cities and counties now have living-wage laws, a new type of minimum wage legislation that mandates a minimum wage high enough to allow a fully employed person to live above the poverty level in a given locale. Living-wage laws provide incomplete coverage, typically applying only to government employees or to firms that contract with that government.

Solved Problem 10.2

Initially, all workers are paid a wage of w_1 per hour. The government taxes the cost of labor by t per hour only in the *covered* sector of the economy. That is, if workers receive a wage of w_2 per hour, firms in the covered sector pay $w_2 + t$ per hour. Show how the wages in the covered and uncovered sectors are determined in the post-tax equilibrium. What effect does the tax have on total employment, L, employment in the covered sector, L_c, and employment in the uncovered sector, L_u?

Answer

1. *Determine the original equilibrium.* In the diagram, the intersection of the total demand curve, D^1, and the total supply curve of labor, S, determines the original equilibrium, e_1, where the wage is w_1, employment in the covered sector is L_c^1, employment in the uncovered sector is L_u^1, and total employment is $L_1 = L_c^1 + L_u^1$. The total demand curve is the horizontal sum of the demand curves in the covered, D_1^c, and uncovered, D^u, sectors.



2. *Show the shift in the demand for labor in the covered sector and the resulting shift in the total demand curve.* The tax causes the demand curve for labor in the covered sector to shift downward from D_1^c to D_2^c. As a result, the total demand curve shifts inward to D^2.

3. *Determine the equilibrium wage using the total supply and demand curves, and then determine employment in the two sectors.* Workers shift between sectors until the new wage is equal in both sectors at w_2, which is determined by the intersection of the new total demand curve, D^2, and the total supply curve, S. Employment in the covered sector is L_c^2, and employment in the uncovered sector L_u^2.

4. *Compare the equilibria.* The tax causes the wage, total employment, and employment in the covered sector to fall and employment in the uncovered sector to rise.

Application

Urban Flight

Philadelphia and some other cities tax wages, while suburban areas do not (or they set much lower rates). Philadelphia collects a wage tax from residents whether or not they work in the city and from nonresidents who work in the city. Unfortunately, this situation drives people and jobs from Philadelphia to the suburbs. To offset such job losses, the city has gradually reduced the wage tax from a high of 4.96% from 1983–1995 to 3.9004% for Philadelphia residents and 3.4741% for nonresidents in 2016.

A study conducted for Philadelphia estimated that if the city were to lower the wage tax by 0.4175 percentage points, 30,500 more people would work in the city. Local wage tax cuts have greater effects than a federal income tax cut. Workers rarely leave the country to avoid taxes, but many will move to the suburbs to avoid a city tax. Indeed, growth over many years has been greater on the suburban side of City Line Avenue, which runs along Philadelphia's border, than on the side within city limits.

10.2 Trading Between Two People

Tariffs, quotas, and other restrictions on trade usually harm both importing and exporting nations (Chapter 9). The reason is that both parties to a voluntary trade benefit from that trade—otherwise they would not have traded. Using a general-equilibrium model, we will show that free trade is Pareto efficient. After all voluntary trades have occurred, we cannot reallocate goods to make one person better off without harming another person. We first demonstrate that trade between two people has this Pareto property. We then show that the same property holds when many people trade using a competitive market.

Endowments

Suppose that Jane and Denise live near each other in the wilds of Massachusetts. A nasty snowstorm hits, isolating them. They must either trade with each other or consume only what they have at hand.

endowment an initial allocation of goods

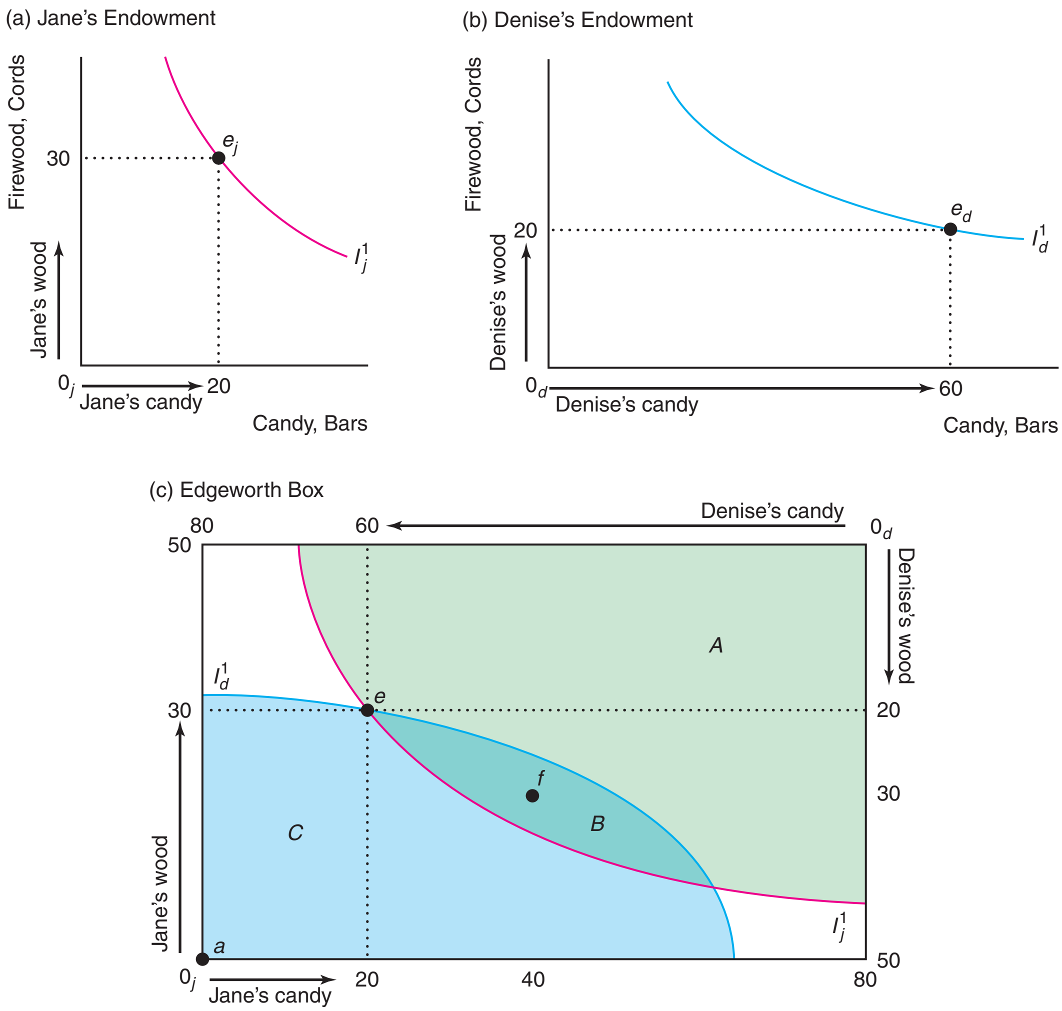
Collectively, they have 50 piles of firewood and 80 bars of candy and no way of producing more of either good. Jane's **endowment**—her initial allocation of goods—is 30 piles of firewood and 20 candy bars. Denise's endowment is 20 (= 50 − 30) piles of firewood and 60 (= 80 − 20) candy bars. So Jane has relatively more wood, and Denise has relatively more candy.

We show these endowments in Figure 10.3. Panels a and b are typical indifference curve diagrams (Chapters 4 and 5) in which we measure piles of firewood on the vertical axis and candy bars on the horizontal axis. Jane's endowment is e_j (30 piles of firewood and 20 candy bars) in panel a, and Denise's endowment is e_d in panel b. Both panels show the indifference curve through the endowment.

If we take Denise's diagram, rotate it 180 degrees, and place it on top of Jane's diagram, we obtain the box in panel c. This type of figure, called an Edgeworth box

Figure 10.3 Endowments in an Edgeworth Box

(a) Jane's endowment is e_j—she has 20 candy bars and 30 piles of firewood. She is indifferent between that bundle and the others that lie on her indifference curve I_j^1. (b) Denise is indifferent between her endowment, e_d (60 candy bars and 20 piles of wood), and the other bundles on I_d^1. (c) Their endowments are at e in the Edgeworth box formed by combining panels a and b. Jane prefers bundles in A and B to e. Denise prefers bundles in B and C to e. Thus, both prefer any bundle in area B to e.



(after an English economist, Francis Ysidro Edgeworth), illustrates trade between two people with fixed endowments of two goods. We use this Edgeworth box to illustrate a general-equilibrium model in which we examine simultaneous trade in firewood and in candy.

The height of the Edgeworth box represents 50 piles of firewood, and the length represents 80 candy bars, which are the combined endowments of Jane and Denise. Bundle *e* shows both endowments. Measuring from Jane's origin, 0_j, at the lower left of the diagram, we see that Jane has 30 piles of firewood and 20 candy bars at endowment *e*. Similarly, measuring from Denise's origin, 0_d, at the upper-right corner, we see that Denise has 60 bars of candy and 20 piles of firewood at *e*.

Mutually Beneficial Trades

Should Jane and Denise trade? The answer depends on their tastes, which are summarized by their indifference curves. We make four assumptions about their tastes and behavior:

- **Utility maximization.** Each person *maximizes* her *utility*.
- **Usual-shaped indifference curves.** Each person's indifference curves have the usual convex shape, so each person has strictly positive *marginal utility* for each good (is never satiated).
- **No interdependence.** Neither person's utility depends on the other's consumption (neither person gets pleasure or displeasure from the other's consumption), and neither person's consumption harms the other (one person's consumption of firewood does not cause smoke pollution that bothers the other person).

Figure 10.3 reflects these assumptions. In panel a, Jane's indifference curve, I_j^1, through her endowment point, e_j, is convex to her origin, 0_j. Jane is indifferent between e_j and any other bundle on I_j^1. She prefers bundles that lie above I_j^1 to e_j and prefers e_j to points that lie below I_j^1. Panel c also shows her indifference curve, I_j^1. The bundles that Jane prefers to her endowment are in the shaded areas *A* and *B*, which lie above her indifference curve, I_j^1.

Similarly, Denise's indifference curve, I_d^1, through her endowment is convex to her origin, 0_d, in the lower left of panel b. This indifference curve, I_d^1, is still convex to 0_d in panel c, but 0_d is in the upper right of the Edgeworth box. (It may help to turn this book around when viewing Denise's indifference curves in an Edgeworth box. Then again, possibly many points will be clearer if the book is held upside down.) The bundles Denise prefers to her endowment are in shaded areas *B* and *C*, which lie on the other side of her indifference curve I_d^1 from her origin 0_d (above I_d^1 if you turn the book upside down).

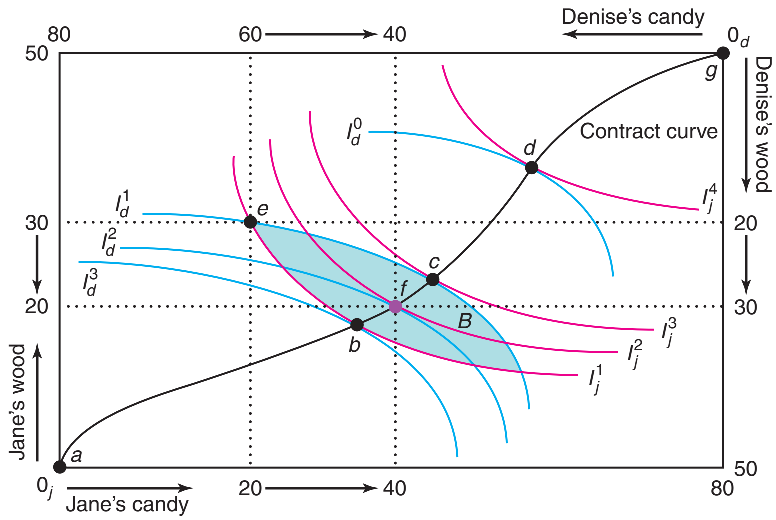
At endowment *e* in panel c, Jane and Denise can both benefit from a trade. Jane prefers bundles in *A* and *B* to *e*, and Denise prefers bundles in *B* and *C* to *e*, so both prefer bundles in area *B* to their endowment at *e*.

Suppose that they trade, reallocating goods from Bundle *e* to *f*. Jane gives up 10 piles of firewood for 20 more candy bars, and Denise gives up 20 candy bars for 10 more piles of wood. As Figure 10.4 illustrates, both gain from such a trade. Jane's indifference curve I_j^2 through allocation *f* lies above her indifference curve I_j^1 through allocation *e*, so she is better off at *f* than at *e*. Similarly, Denise's indifference curve I_d^2 through *f* lies above (if you hold the book upside down) her indifference curve I_d^1 through *e*, so she also benefits from the trade.

Now that they've traded to Bundle *f*, do Jane and Denise want to make further trades? To answer this question, we can repeat our analysis. Jane prefers all bundles above I_j^2, her indifference curve through *f*. Denise prefers all bundles above (when the book is held upside down) I_d^2 to *f*. However, they do not both prefer any other bundle

Figure 10.4 Contract Curve MyLab Economics Video

The contract curve contains all the Pareto-efficient allocations. Any bundle for which Jane's indifference curve is tangent to Denise's indifference curve lies on the contract curve. At such a bundle, because no further trade is possible, and we can't reallocate goods to make one of them better off without harming the other. Starting at an endowment of *e*, Jane and Denise will trade to a bundle on the contract curve in area *B*: bundles between *b* and *c*. The table shows how they would trade to Bundle *f*.



	Endowment, *e*		Trade		New Allocation, *f*	
	Wood	Candy	Wood	Candy	Wood	Candy
Jane	30	20	–10	+20	20	40
Denise	20	60	+10	–20	30	40

because I_j^2 and I_d^2 are tangent at *f*. Neither Jane nor Denise wants to trade from *f* to a bundle such as *e*, which is below both of their indifference curves. Jane would love to trade from *f* to *c*, which is on her higher indifference curve I_j^3, but such a trade would make Denise worse off because this bundle is on a lower indifference curve, I_d^1. Similarly, Denise prefers *b* to *f*, but Jane does not. Thus, any move from *f* harms at least one of them.

The reason no further trade is possible at a bundle like *f* is that Jane's marginal rate of substitution (the slope of her indifference curve), MRS_j, between wood and candy equals Denise's marginal rate of substitution, MRS_d. Jane's MRS_j is $-\frac{1}{2}$: She is willing to trade one pile of wood for two candy bars. Because Denise's indifference curve is tangent to Jane's, Denise's MRS_d must also be $-\frac{1}{2}$. When they both want to trade wood for candy at the same rate, they can't agree on further trades.

In contrast, at a bundle such as *e* where their indifference curves are not tangent, MRS_j does not equal MRS_d. Denise's MRS_d is $-\frac{1}{3}$, and Jane's MRS_j is -2. Denise is willing to give up one pile of wood for three more candy bars or to sacrifice three candy bars for one more pile of wood. If Denise offers Jane three candy bars for one pile of wood, Jane will accept because she is willing to give up two piles of wood for one candy bar. This example illustrates that trades are possible where indifference curves intersect because marginal rates of substitution are unequal.

To summarize, we can make four equivalent statements about allocation *f*:

1. The indifference curves of the two parties are tangent at *f*.
2. The parties' marginal rates of substitution are equal at *f*.
3. No further mutually beneficial trades are possible at *f*.
4. The allocation at *f* is Pareto efficient: One party cannot be made better off without harming the other.

contract curve
the set of all Pareto-efficient bundles

Indifference curves are also tangent at Bundles *b*, *c*, and *d*, so these allocations, like *f*, are Pareto efficient. By connecting all such bundles, we draw the **contract curve**: the set of all Pareto-efficient bundles. The reason for this name is that only at these points are the parties unwilling to engage in further trades or contracts—these allocations are the final contracts. A move from any bundle on the contract curve would harm at least one person.

Solved Problem 10.3

Are allocations *a* and *g* in Figure 10.4 part of the contract curve?

Answer

By showing that no mutually beneficial trades are possible at those points, demonstrate that those bundles are Pareto efficient. The allocation at which Jane has everything, allocation *g*, is on the contract curve because no mutually beneficial trade is possible: Denise has no goods to trade with Jane. Consequently, we cannot make Denise better off without taking goods from Jane. Similarly, when Denise has everything, *a*, we can make Jane better off only by taking wood or candy from Denise and giving it to Jane.

Bargaining Ability

For every allocation off the contract curve, the contract curve has allocations that benefit at least one person. If they start at endowment *e*, Jane and Denise should trade until they reach a point on the contract curve between Bundles *b* and *c* in Figure 10.4. All the allocations in area *B* are beneficial. However, if they trade to any allocation in *B* that is not on the contract curve, further beneficial trades are possible because their indifference curves intersect at that allocation.

Where will they end up on the contract curve between *b* and *c*? That depends on who is better at bargaining. Suppose that Jane is much better at bargaining. Jane knows that the more she gets, the worse off Denise will be and that Denise will not agree to any trade that makes her worse off than she is at *e*. Thus, the best trade Jane can make is one that leaves Denise only as well off as at *e*, which are the bundles on I_d^1. If Jane could pick any point she wanted along I_d^1, she'd choose the bundle on her highest possible indifference curve, which is Bundle *c*, where I_j^3 is just tangent to I_d^1. After this trade, Denise is no better off than before, but Jane is much happier. By similar reasoning, if Denise is sufficiently better at bargaining, the final allocation will be at *b*.

10.3 Competitive Exchange

Most trading throughout the world occurs without one-on-one bargaining between people. When you go to the store to buy a bottle of shampoo, you read its posted price and then decide whether to buy it or not. You've probably never tried to bargain with the store's clerk over the price of shampoo: You're a price taker in the shampoo market.

If we don't know much about how Jane and Denise bargain, all we can say is that they will trade to some allocation on the contract curve. If we know the exact trading process they use, however, we can apply that process to determine the final allocation. In particular, we can examine the competitive trading process to determine the competitive equilibrium in a pure exchange economy.

In Chapter 9, we used a partial-equilibrium approach to show that one measure of welfare, W, is maximized in a competitive market in which many voluntary trades occur. We now use a general-equilibrium model to show that a competitive market has two desirable properties (which hold under fairly weak conditions):

- **The First Theorem of Welfare Economics:** *The competitive equilibrium is efficient.* Competition results in a Pareto-efficient allocation—no one can be made better off without making someone worse off—in all markets.
- **The Second Theorem of Welfare Economics:** *Any efficient allocations can be achieved by competition.* All possible efficient allocations can be obtained by competitive exchange, given an appropriate initial allocation of goods.

Competitive Equilibrium

When two people trade, they are unlikely to view themselves as price takers. However, if the market has a large number of people with tastes and endowments like Jane's and a large number of people with tastes and endowments like Denise's, then each person is a price taker in the markets for the two goods. We can use an Edgeworth box to examine how such price takers would trade.

Because they can trade only two goods, each person needs to consider only the relative price of the two goods when deciding whether to trade. If the price of a pile of wood, p_w, is \$2, and the price of a candy bar, p_c, is \$1, then a candy bar costs half as much as a pile of wood: $p_c/p_w = \frac{1}{2}$. An individual can sell one pile of wood and use that money to buy two candy bars.

At the initial allocation, e, Jane has goods worth \$80 = (\$2 per pile × 30 piles of firewood) + (\$1 per candy bar × 20 candy bars). At these prices, Jane could keep her endowment or trade to an allocation with 40 piles of firewood and no candy, 80 bars of candy and no firewood, or any combination in between, as the price line (budget line) in panel a of Figure 10.5 shows. The price line is all the combinations of goods Jane could get by trading, given her endowment. The price line goes through point e and has a slope of $-p_c/p_w = -\frac{1}{2}$.

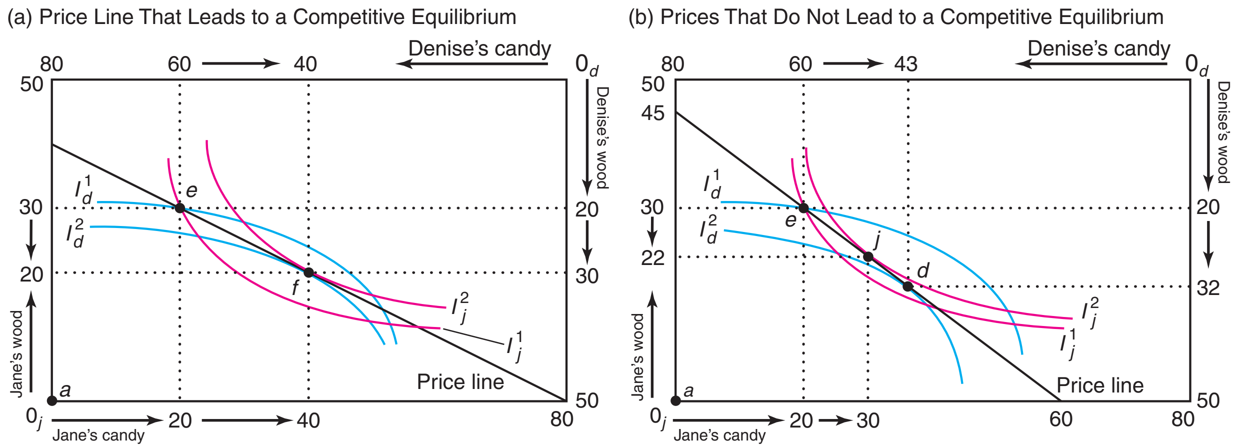
Given the price line, what bundle of goods will Jane choose? She wants to maximize her utility by picking the bundle where one of her indifference curves, I_j^2, is tangent to her budget or price line. Denise wants to maximize her utility by choosing a bundle in the same way.

In a competitive market, prices adjust until the quantity supplied equals the quantity demanded. An auctioneer could help determine the equilibrium. The auctioneer could call out relative prices and ask how much is demanded and how much is offered for sale at those prices. If demand does not equal supply, the auctioneer calls out another relative price. When demand equals supply, the transactions actually occur and the auction stops. At some ports, fishing boats sell their catch to fish wholesalers at a daily auction run in this manner.

Panel a shows that when candy costs half as much as wood, the quantity demanded of each good equals the quantity supplied. Jane (and every person like her) wants to sell 10 piles of firewood and use that money to buy 20 additional candy bars. Similarly, Denise (and everyone like her) wants to sell 20 candy bars and buy 10 piles of wood. Thus, the quantity of wood sold equals the quantity bought, and the quantity of candy demanded equals that supplied. We can see in the figure that the quantities demanded equal the quantities supplied because the optimal bundle for both types of consumers is the same, Bundle f.

Figure 10.5 Competitive Equilibrium

The initial endowment is *e*. (a) Along the price line facing Jane and Denise where $p_w = \$2$ and $p_c = \$1$, they trade to point *f*, where Jane's indifference curve, I_j^2, is tangent to the price line and to Denise's indifference curve, I_d^2. (b) No other price line results in an equilibrium. If $p_w = \$1.33$ and $p_c = \$1$, Denise wants to buy 12 (= 32 − 20) piles of firewood at these prices, but Jane wants to sell only 8 (= 30 − 22) piles. Conversely, Jane wants to buy 10 (= 30 − 20) candy bars, but Denise wants to sell 17 (= 60 − 43). Thus, these prices are not consistent with a competitive equilibrium.



At any other price ratio, the quantity demanded of each good would not equal the quantity supplied. For example, if the price of candy remained constant at $p_c = \$1$ per bar but the price of wood fell to $p_w = \$1.33$ per pile, the price line would be steeper, with a slope of $-p_c/p_w = -1/1.33 = -\frac{3}{4}$ in panel b. At these prices, Jane wants to trade to Bundle *j* and Denise wants to trade to Bundle *d*. Because Jane wants to buy 10 extra candy bars but Denise wants to sell 17 extra candy bars, the quantity supplied does not equal the quantity demanded, so this price ratio does not result in a competitive equilibrium when the endowment is *e*.

The Efficiency of Competition

In a competitive equilibrium, the indifference curves of both types of consumers are tangent at the same bundle on the price line. As a result, the slope (*MRS*) of each person's indifference curve equals the slope of the price line, so the slopes of the indifference curves are equal:

$$MRS_j = -\frac{p_c}{p_w} = MRS_d. \tag{10.1}$$

The marginal rates of substitution are equal across consumers in the competitive equilibrium, so the competitive equilibrium must lie on the contract curve. Thus, we have demonstrated the

First Theorem of Welfare Economics: *Any competitive equilibrium is Pareto efficient.*

The intuition for this result is that people (who face the same prices) make all the voluntary trades they want in a competitive market. Because no additional voluntary trades can occur, we cannot make someone better off without harming someone else. (If an involuntary trade occurs, at least one person is made worse off. A person who steals goods from another person—an involuntary exchange—gains at the expense of the victim.)

Obtaining Any Efficient Allocation Using Competition

Of the many possible Pareto-efficient allocations, the government may want to choose one. Can it achieve that allocation using the competitive market mechanism?

Our previous example illustrates that the competitive equilibrium depends on the endowment: the initial distribution of wealth. For example, if the initial endowment were *a* in panel a of Figure 10.5—where Denise has everything and Jane has nothing—the competitive equilibrium would be *a* because no trades would be possible.

Thus, for competition to lead to a particular allocation—say, *f*—the trading must start at an appropriate endowment. If the consumers' endowment is *f*, a Pareto-efficient point, their indifference curves are tangent at *f*, so no further trades occur. That is, *f* is a competitive equilibrium.

Many other endowments will also result in a competitive equilibrium at *f*. Panel a shows that the resulting competitive equilibrium is *f* if the endowment is *e*. In that figure, a price line goes through both *e* and *f*. If the endowment is any bundle along this price line—not just *e* or *f*—the competitive equilibrium is *f*, because only at *f* are the indifference curves tangent.

To summarize, any Pareto-efficient bundle *x* can be obtained as a competitive equilibrium if the initial endowment is *x*. That allocation can also be obtained as a competitive equilibrium if the endowment lies on a price line through *x*, where the slope of the price line equals the marginal rate of substitution of the indifference curves that are tangent at *x*. Thus, we've demonstrated the

> **Second Theorem of Welfare Economics:** *Any Pareto-efficient equilibrium can be obtained by competition, given an appropriate endowment.*

The first welfare theorem tells us that society can achieve efficiency by allowing competition. The second welfare theorem adds that society can obtain the particular efficient allocation it prefers based on its value judgments about equity by appropriately redistributing endowments.

10.4 Production and Trading

So far our discussion has been based on a pure exchange economy with no production. We now examine an economy in which a fixed amount of a single input can be used to produce two different goods.

Comparative Advantage

Jane and Denise can produce candy or chop firewood using their own labor. They differ, however, in how much of each good they produce from a day's work.

Production Possibility Frontier Jane can produce either 3 candy bars or 6 piles of firewood in a day. By splitting her time between the two activities, she can produce various combinations of the two goods. If a is the fraction of a day she spends making candy and $1 - a$ is the fraction cutting wood, she produces $3a$ candy bars and $6(1 - a)$ piles of wood.

By varying a between 0 and 1, we trace out the line in panel a of Figure 10.6. This line is Jane's production possibility frontier, PPF_j, which shows the maximum combinations of wood and candy that she can produce from a given amount of input(Chapter 7). If Jane works all day using the best available technology (such as a sharp ax), she achieves efficiency in production and produces combinations of goods on PPF_j. If she sits around part of the day or does not use the best technology, she produces an inefficient combination of wood and candy inside PPF_j.

Marginal Rate of Transformation The slope of the production possibility frontier is the *marginal rate of transformation* (MRT).[7] The marginal rate of transformation tells us how much more wood can be produced if the production of candy is reduced by one bar. Because Jane's PPF_j is a straight line with a slope of −2, her MRT is −2 at every allocation.

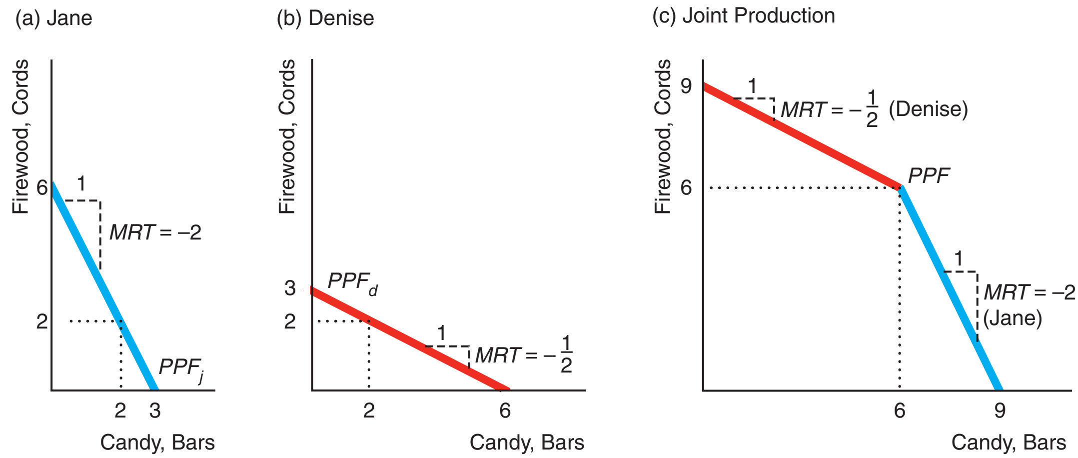
Denise can produce up to 3 piles of wood or 6 candy bars in a day. Panel b shows her production possibility function, PPF_d, with an $MRT = -\frac{1}{2}$. Thus, with a day's work, Denise can produce relatively more candy, and Jane can produce relatively more wood, as reflected by their differing marginal rates of transformation.

comparative advantage the ability to produce a good at a lower opportunity cost than someone else

The marginal rate of transformation shows how much it costs to produce one good in terms of the forgone production of the other good. Someone with the ability to produce a good at a lower opportunity cost than someone else has a **comparative advantage** in producing that good. Denise has a comparative advantage in producing candy (she forgoes less in wood production to produce a given amount of candy), and Jane has a comparative advantage in producing wood.

Figure 10.6 Comparative Advantage and Production Possibility Frontiers

(a) Jane's production possibility frontier, PPF_j, shows that in a day, she can produce 6 piles of firewood or 3 candy bars or any combination of the two. Her marginal rate of transformation (MRT) is −2. (b) Denise's production possibility frontier, PPF_d, has an MRT of $-\frac{1}{2}$. (c) Their joint production possibility frontier, PPF, has a kink at 6 piles of firewood (produced by Jane) and 6 candy bars (produced by Denise) and is concave to the origin.



[7]In the standard consumer model (Chapter 4), the slope of a consumer's budget line is the marginal rate of transformation. That is, for a price-taking consumer who obtains goods by buying them, the budget line plays the same role as the production possibility frontier for someone who produces the two goods.

By combining their outputs, they have the joint production possibility frontier *PPF* in panel c. If Denise and Jane spend all their time producing wood, Denise produces 3 piles and Jane produces 6 piles for a total of 9, which is where the joint *PPF* hits the wood axis. Similarly, if they both produce candy, they can jointly produce 9 bars. If Denise specializes in making candy and Jane specializes in cutting wood, they produce 6 candy bars and 6 piles of wood, a combination that appears at the kink in the *PPF*.

If they choose to produce a relatively large quantity of candy and a relatively small amount of wood, Denise produces only candy and Jane produces some candy and some wood. Jane chops the wood because that's her comparative advantage. The marginal rate of transformation in the lower portion of the *PPF* is Jane's, −2, because only she produces both candy and wood.

Similarly, if they produce little candy, Jane produces only wood and Denise produces some wood and some candy, so the marginal rate of transformation in the higher portion of the *PPF* is Denise's, $-\frac{1}{2}$. In short, the *PPF* has a kink at 6 piles of wood and 6 candy bars and is concave (bowed away from the origin).

Benefits of Trade Because of the difference in their marginal rates of transformation, Jane and Denise can benefit from a trade. Suppose that Jane and Denise like to consume wood and candy in equal proportions. If they do not trade, each produces 2 candy bars and 2 piles of wood in a day. If they agree to trade, Denise, who excels at making candy, spends all day producing 6 candy bars. Similarly, Jane, who has a comparative advantage at chopping, produces 6 piles of wood. If they split this production equally, they can each have 3 piles of wood and 3 candy bars—50% more than if they don't trade.

They do better if they trade because each person uses her comparative advantage. Without trade, if Denise wants an extra pile of wood, she must give up two candy bars. Producing an extra pile of wood costs Jane only half a candy bar in forgone production. Denise is willing to trade up to two candy bars for a pile of wood, and Jane is willing to trade the wood as long as she gets at least half a candy bar. Thus, a mutually beneficial trade is possible.

Solved Problem 10.4

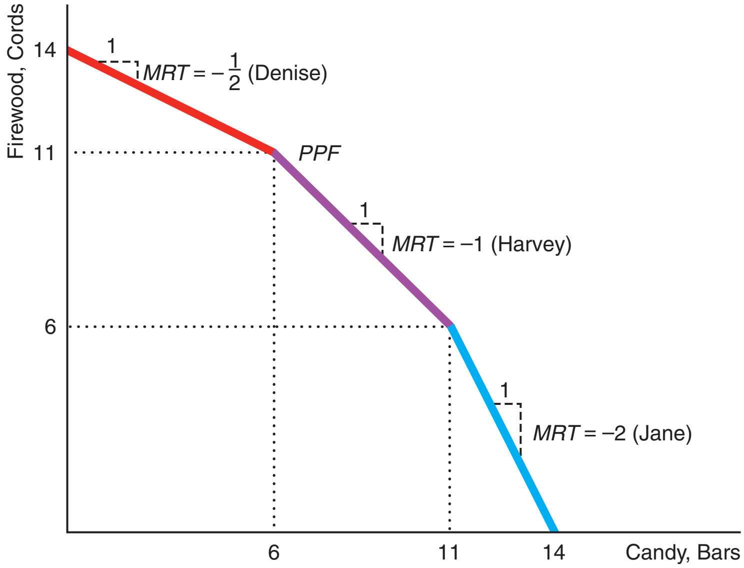
MyLab Economics Solved Problem

How does the joint production possibility frontier for Jane and Denise in panel c of Figure 10.6 change if they can also trade with Harvey, who can produce 5 piles of wood, 5 candy bars, or any linear combination of wood and candy in a day?

Answer

1. *Describe each person's individual production possibility frontier.* Panels a and b of Figure 10.6 show the production possibility frontiers of Jane and Denise. Harvey's production possibility frontier is a straight line that hits the firewood axis at 5 piles and the candy axis at 5 candy bars (not shown in Figure 10.6).

2. *Draw the joint PPF, by starting at the quantity on the horizontal axis that is produced if everyone specializes in candy and then connecting the individual production possibility frontiers in order of comparative advantage in chopping wood.* If all three produce candy, they make 14 candy bars in the figure. Jane has a comparative advantage at chopping wood over Harvey and Denise, and Harvey has a comparative advantage over Denise. Thus, Jane's production possibility frontier is the first one (starting at the lower right), then comes Harvey's, and then Denise's. The resulting *PPF* is concave to the origin. (If we change the order of the individual frontiers, the resulting kinked line lies inside the *PPF*. Thus, the new line cannot be the joint production possibility frontier, which shows the maximum possible production from the available labor inputs.)

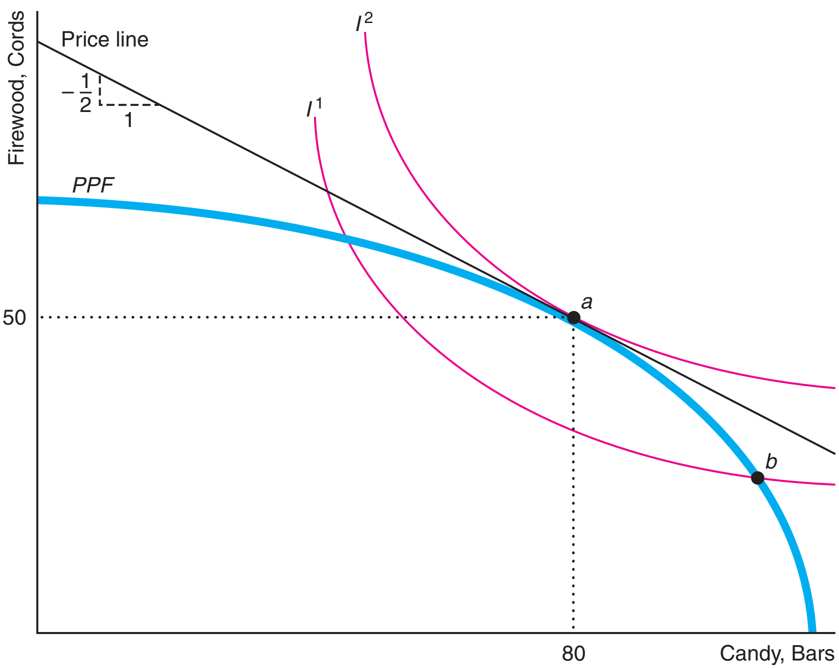


The Number of Producers If the only two ways of producing wood and candy are Denise's and Jane's methods with different marginal rates of transformation, the joint production possibility frontier has a single kink (panel c of Figure 10.6). If another method of production with a different marginal rate of transformation—Harvey's—is added, the joint production possibility frontier has two kinks (as in the figure in Solved Problem 10.4).

If many firms can produce candy and firewood with different marginal rates of transformation, the joint production possibility frontier has even more kinks. As the number of firms becomes very large, the *PPF* becomes a smooth curve that is concave to the origin, as in Figure 10.7.

Figure 10.7 Optimal Product Mix

The optimal product mix, *a*, could be determined by maximizing an individual's utility by picking the allocation for which an indifference curve is tangent to the production possibility frontier. It could also be determined by picking the allocation where the relative competitive price, p_c/p_f, equals the slope of the *PPF*.



Because the *PPF* is concave, the marginal rate of transformation decreases (in absolute value) as we move up the *PPF*. The *PPF* has a flatter slope at *a*, where the $MRT = -\frac{1}{2}$, than at *b*, where the $MRT = -1$. At *a*, giving up a candy bar leads to half a pile more wood production. In contrast, at *b*, where relatively more candy is produced, giving up producing a candy bar frees enough resources that an additional pile of wood can be produced.

The marginal rate of transformation along this smooth *PPF* tells us about the marginal cost of producing one good relative to the marginal cost of producing the other good. The marginal rate of transformation is the negative of the ratio of the marginal cost of producing candy, MC_c, and wood, MC_w:

$$MRT = -\frac{MC_c}{MC_w}. \tag{10.2}$$

Suppose that at point *a* in Figure 10.7, a firm's marginal cost of producing an extra candy bar is \$1 and its marginal cost of producing an additional pile of firewood is \$2. As a result, the firm can produce one extra candy bar or half a pile of wood at a cost of \$1. The marginal rate of transformation is the negative of the ratio of the marginal costs, $-(\$1/\$2) = -\frac{1}{2}$. To produce one more candy bar, the firm must give up producing half a pile of wood.

Efficient Product Mix

Which combination of products along the *PPF* does society choose? If a single person were to decide on the product mix, that person would pick the allocation of wood and candy along the *PPF* that maximized his or her utility. A person with the indifference curves in Figure 10.7 would pick Allocation *a*, which is the point where the *PPF* touches indifference curve I^2.

Because I^2 is tangent to the *PPF* at *a*, that person's marginal rate of substitution (the slope of indifference curve I^2) equals the marginal rate of transformation (the slope of the *PPF*). The marginal rate of substitution, *MRS*, tells us how much a consumer is willing to give up of one good to get another. The marginal rate of transformation, *MRT*, tells us how much of one good we need to give up to produce more of another good.

If the *MRS* doesn't equal the *MRT*, the consumer will be happier with a different product mix. At Allocation *b*, the indifference curve I^1 intersects the *PPF*, so the *MRS* does not equal the *MRT*. At *b*, the consumer is willing to give up one candy bar to get a third of a pile of wood ($MRS = -\frac{1}{3}$), but firms can produce one pile of wood for every candy bar not produced ($MRT = -1$). Thus, at *b*, too little wood is being produced. If the firms increase wood production, the *MRS* will fall (become more negative) and the *MRT* will rise (become less negative) until they are equal at *a*, where $MRS = MRT = -\frac{1}{2}$.

We can extend this reasoning to look at the product mix choice of all consumers simultaneously. Each consumer's marginal rate of substitution must equal the economy's marginal rate of transformation, $MRS = MRT$, if the economy is to produce the optimal mix of goods for each consumer. How can we ensure that this condition holds for all consumers? One way is to use the competitive market.

Competition

Each price-taking consumer picks a bundle of goods so that the consumer's marginal rate of substitution equals the slope of the consumer's price line (the negative of the relative prices):

$$MRS = -\frac{p_c}{p_w}. \tag{10.3}$$

Thus, if all consumers face the same relative prices, in the competitive equilibrium all consumers will buy a bundle where their marginal rates of substitution are equal (Equation 10.1). Because all consumers have the same marginal rates of substitution, no further trades can occur. Thus, the competitive equilibrium achieves *consumption efficiency*: We can't redistribute goods among consumers to make one consumer better off without harming another one. That is, the competitive equilibrium lies on the contract curve.

If candy and wood are sold by competitive firms, each firm sells a quantity of a candy for which its price equals its marginal cost,

$$p_c = MC_c, \tag{10.4}$$

and a quantity of wood for which its price and marginal cost are equal,

$$p_w = MC_w. \tag{10.5}$$

Taking the ratio of Equations 10.4 and 10.5, we find that in competition, $p_c/p_w = MC_c/MC_w$. From Equation 10.2, we know that the marginal rate of transformation equals $-MC_c/MC_w$, so

$$MRT = -\frac{p_c}{p_w}. \tag{10.6}$$

We can illustrate why firms want to produce where Equation 10.6 holds. Suppose that a firm were producing at b in Figure 10.7, where its MRT is -1, and that $p_c = \$1$ and $p_w = \$2$, so $-p_c/p_w = -\frac{1}{2}$. If the firm reduces its output by one candy bar, it loses $1 in candy sales but makes $2 more from selling the extra pile of wood, for a net gain of $1. Thus, at b, where the $MRT < -p_c/p_w$, the firm should reduce its output of candy and increase its output of wood. In contrast, if the firm is producing at a, where the $MRT = -p_c/p_w = -\frac{1}{2}$, it has no incentive to change its behavior: The gain from producing a little more wood exactly offsets the loss from producing a little less candy.

Combining Equations 10.3 and 10.6, we find that in the competitive equilibrium, the MRS equals the relative prices, which equals the MRT:

$$MRS = -\frac{p_c}{p_w} = MRT.$$

Because competition ensures that the MRS equals the MRT, a competitive equilibrium achieves an *efficient product mix*: The rate at which firms can transform one good into another equals the rate at which consumers are willing to substitute between the goods, as reflected by their willingness to pay for the two goods.

By combining the production possibility frontier and an Edgeworth box, we can show the competitive equilibrium in both production and consumption. Suppose that firms produce 50 piles of firewood and 80 candy bars at a in Figure 10.8. The size of the Edgeworth box—the maximum amount of wood and candy available to consumers—is determined by point a on the PPF.

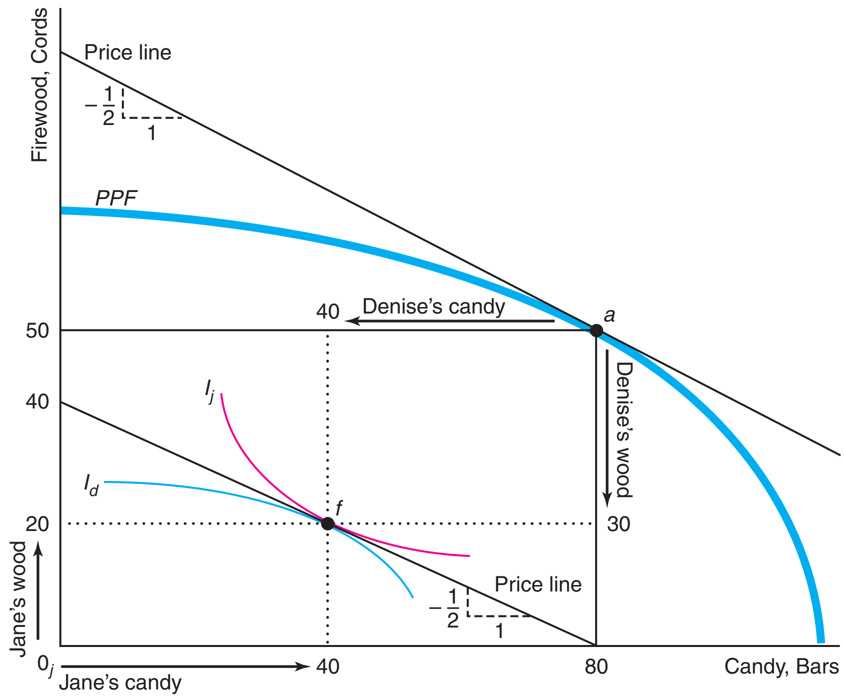
The prices consumers pay must equal the prices producers receive, so the price lines consumers and producers face must have the same slope of $-p_c/p_w$. In equilibrium, the price lines are tangent to each consumer's indifference curve at f and to the PPF at a.

In this competitive equilibrium, supply equals demand in all markets. The consumers buy the mix of goods at f. Consumers like Jane, whose origin, 0_j, is at the lower left, consume 20 piles of firewood and 40 candy bars. Consumers like Denise, whose origin is a at the upper right of the Edgeworth box, consume 30 (= 50 − 20) piles of firewood and 40 (= 80 − 40) candy bars.

The two key results concerning competition still hold in an economy with production. First, a competitive equilibrium is Pareto efficient, achieving efficiency in

Figure 10.8 Competitive Equilibrium

At the competitive equilibrium, the relative prices firms and consumers face are the same (the price lines are parallel), so the $MRS = -p_c/p_w = MRT$.



consumption and in output mix.[8] Second, any particular Pareto-efficient allocation between consumers can be obtained through competition, given that the government chooses an appropriate endowment.

10.5 Efficiency and Equity

How well various members of society live depends on how society deals with efficiency (the size of the pie) and equity (how the pie is divided). The actual outcome depends on choices by individuals and on government actions.

Role of the Government

By altering the efficiency with which goods are produced and distributed and the endowment of resources, governments help determine how much is produced and how goods are allocated. By redistributing endowments or by refusing to do so, governments, at least implicitly, are making value judgments about which members of society should get relatively more of society's goodies.

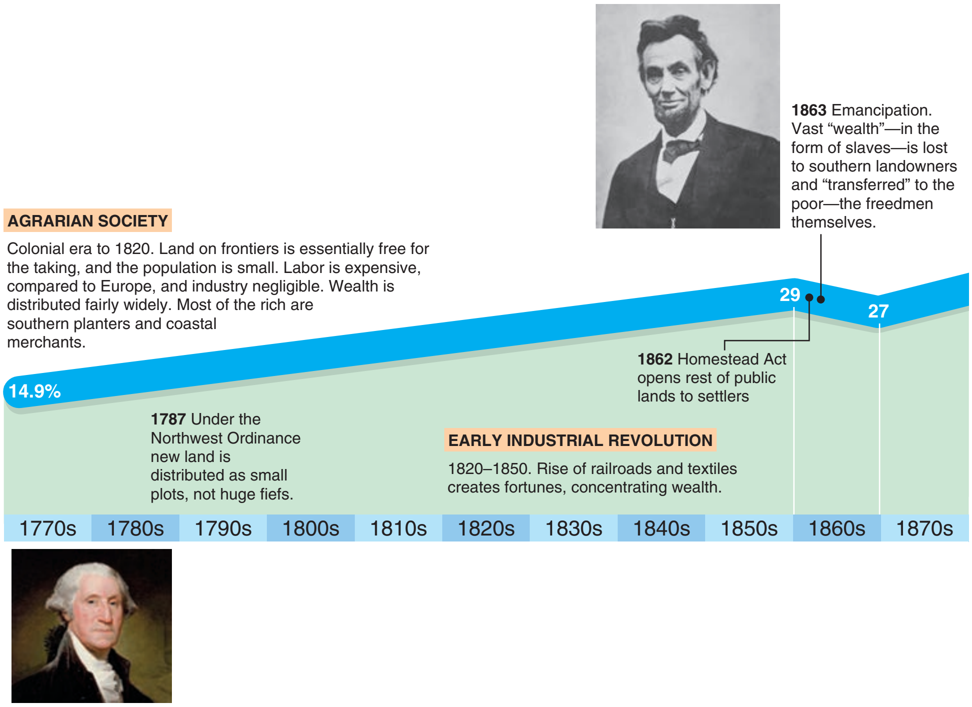
Virtually every government program, tax, or action redistributes wealth. Proceeds from a British lottery, played mostly by lower-income people, support the "rich toffs" who attend the Royal Opera House at Covent Garden. Agricultural price support programs (Chapter 9) redistribute wealth to farmers from other taxpayers. Income taxes (Chapter 5) and food stamp programs (Chapter 4) redistribute income from the rich to the poor.

[8]Although we have not shown it here, competitive firms choose factor combinations so that their marginal rates of technical substitution between inputs equal the negative of the ratios of the relative factor prices (see Chapter 7). That is, competition also results in *efficiency in production*: We could not produce more of one good without producing less of another good.

Application

The 1% Grow Wealthier

Share of Wealth of the Richest 1 Percent



Money is better than poverty, if only for financial reasons. —Woody Allen

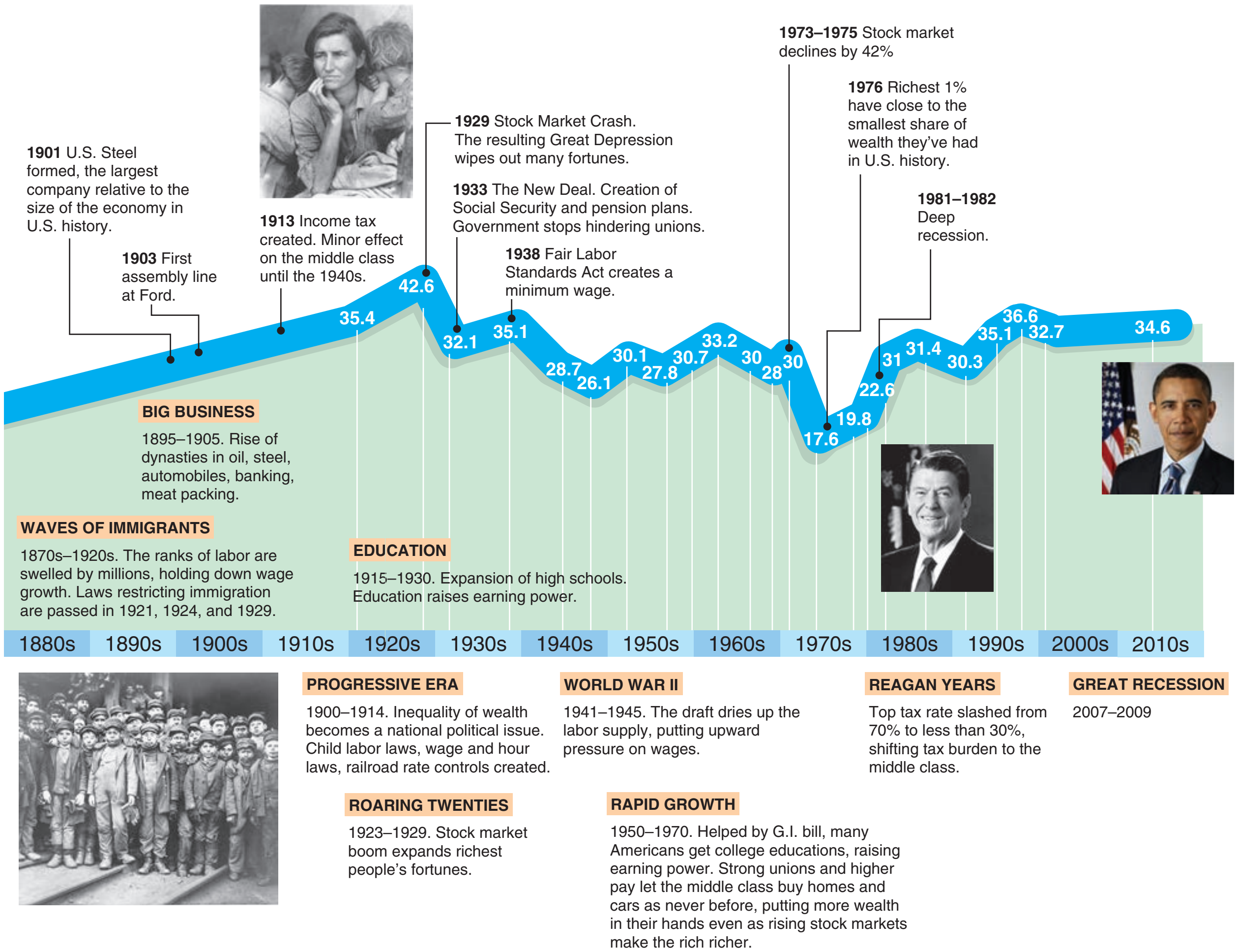
Wealth is unequally distributed, and the wealthiest are becoming wealthier over time. The six heirs to the Walmart fortune have as much wealth as the least wealthy 42% of U.S. families—49 million households.

According to the charity Oxfam, the 85 richest people in the world (such as Bill Gates, Carlos Slim Helu, Warren Buffett, and the Koch brothers)—most of whom live in Europe or the United States—had as much wealth in 2014 as did the poorest half of the world's population. That is, on average, each of these extremely wealthy people has as much wealth as about 42 million of the world's poorest people—about the number of people who live in Argentina. The world's richest man, Bill Gates, had \$76 billion in wealth in 2016, which is the same as the 156 million poorest people. The wealthiest 1% had roughly the same amount of global wealth (48%) as did the poorest 99% (52%). The bottom four-fifths of people had only 5.5% of the wealth.

North America and Europe have nearly two-thirds (65%) of the world's wealth, China and other Asian-Pacific nations have a bit less than a third (29%), while Latin America, India, and Africa combined have only about one-fifteenth (6%).

The United States has less equally distributed wealth than do other developed countries. The top 10% of U.S. households have 74% of U.S. wealth. The corresponding shares are 50% in France, 48% in Canada, 44% in the United Kingdom, and 34% in Japan.

Since its founding, changes in the U.S. economy have altered the share of the nation's wealth held by the richest 1% of Americans (see the figure). An array of social changes—sometimes occurring during or after wars and often codified into new laws—have greatly redistributed wealth. For example, the emancipation of



slaves in 1863 transferred vast wealth—the labor of the former slaves—from rich Southern landowners to the poor freed slaves.

The share of wealth—the total assets owned—held by the richest 1% generally increased until the Great Depression, declined through the mid-1970s, and has increased substantially since then. Thus, the greatest concentrations of wealth occurred in 1929 during the Great Depression and today, following the Great Recession.

A major cause of the recent increased concentration of wealth is that the top income tax rate fell from 70% to less than 30% at the beginning of the Reagan administration, shifting more of the tax burden to the middle class. Between 1989 and 2010, the wealth share of the top 1% of U.S. households rose from 30% to 35%. The share of the wealthiest 10% rose from two-thirds to three-quarters. Correspondingly, the wealth share of the poorest 90% fell from a third to a quarter. The share of the lowest half of households dropped from a tiny 3% to an almost nonexistent 1%.[9]

[9]The U.S. federal government transfers 5% of total national household income from the rich to the poor: 2% using cash assistance such as general welfare programs and 3% using in-kind transfers such as food stamps and school lunch programs. Poor households receive 26% of their income from cash assistance and 18% from in-kind assistance. The U.S. government gives only 0.1% of its gross national product to poor nations. In contrast, Britain gives 0.26% and the Netherlands transfers 0.8%.

Efficiency

Many economists and political leaders make the value judgment that governments *should* use the Pareto principle, preferring reallocations of resources that make someone better off while harming no one else. Consequently, they believe that governments should allow voluntary trades, encourage competition, and otherwise try to prevent problems that reduce efficiency.

We can use the Pareto principle to rank allocations or government policies that alter allocations. The Pareto criterion ranks allocation x over allocation y if some people are better off at x and no one else is harmed. If that condition is met, we say that x is *Pareto superior* to y.

We cannot always use the Pareto principle to compare allocations. If both allocation x and allocation y are Pareto efficient, we cannot use this criterion to rank them. For example, if Denise has all the goods in x and Jane has all the goods in y, then we cannot rank these allocations using the Pareto rule.

To choose between two Pareto-efficient allocations, we have to make a value judgment based on interpersonal comparisons. Society must make interpersonal comparisons to evaluate most government policies.

Suppose that, when a country ends a ban on imports and allows free trade, domestic consumers benefit by many times more than domestic producers suffer. This policy change does not meet the Pareto efficiency criterion that someone may benefit without anyone else suffering. Of course, the government could adopt a more complex policy that meets the Pareto criterion. Because consumers benefit by more than producers suffer, the government could take enough of the free-trade gains from consumers to compensate the producers so that no one is harmed and some people benefit.

The government rarely uses policies that require winners to subsidize losers. If a policy does not require such subsidization, society must make additional value judgments involving interpersonal comparisons to decide whether to adopt that policy.

We have been using a welfare measure, W = consumer surplus + producer surplus, that equally weights benefits and losses to consumers and producers. This measure is making an implicit interpersonal comparison by weighting consumers and producers equally. By this criterion, if a policy results in gains to consumers that outweigh the losses to producers, society should adopt that policy.

Thus, calling for policy changes that lead to Pareto-superior allocations is a weaker rule than calling for policy changes that increase the welfare measure W. We can rank more allocations using the welfare measure than using the Pareto rule. Any policy change that leads to a Pareto-superior allocation must increase W; however, some policy changes that increase W are not Pareto superior, because they produce both winners and losers.

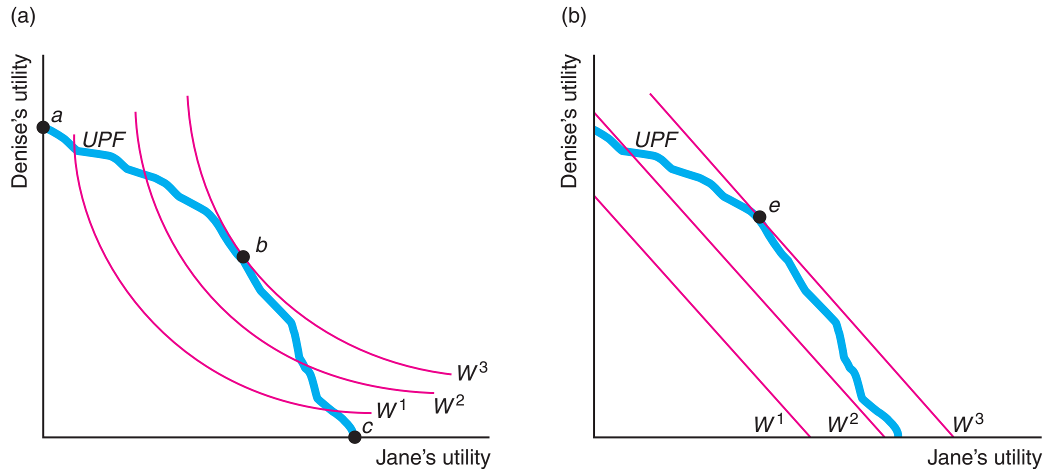
Equity

If we are unwilling to use the Pareto principle or if that criterion does not allow us to rank the relevant allocations, we must make additional value judgments to rank these allocations. We can summarize these value judgments using a *social welfare function* that combines various consumers' utilities to provide a collective ranking of allocations. Loosely speaking, a social welfare function is a utility function for society.

We illustrate the use of a social welfare function using the pure exchange economy in which Jane and Denise trade wood and candy. The contract curve in Figure 10.4 consists of many possible Pareto-efficient allocations. Jane and Denise's utility levels vary along the contract curve. Figure 10.9 shows the utility possibility frontier (*UPF*): the set of utility levels corresponding to the Pareto-efficient allocations along the contract curve. Point *a* in panel a corresponds to the end of the contract curve at which Denise has all the goods, and *c* corresponds to the allocation at which Jane has all the goods.

Figure 10.9 Welfare Maximization

Society maximizes welfare by choosing the allocation for which the highest possible isowelfare curve touches the utility possibility frontier, *UPF*. (a) The isowelfare curves have the shape of a typical indifference curve. (b) The isowelfare lines have a slope of −1, indicating that society views the utilities of the two people as perfect substitutes.



The curves labeled W^1, W^2, and W^3 are *isowelfare curves*. Welfare is the same at every point along one of these curves. In panel a, the isowelfare curves have a similar shape to indifference curves for individuals, so that the utilities of the two people are imperfect substitutes. Society maximizes its welfare at point *b*. In panel b, the isowelfare curves are straight lines with a slope of −1, indicating that society views their utilities as perfect substitutes. Here, society maximizes welfare at point *e*.

Who decides on the welfare function? In most countries, government leaders make decisions about which allocations are most desirable. These officials may believe that transferring money from wealthy people to poor people raises welfare, or vice versa. When government officials choose a particular allocation, they are implicitly or explicitly judging which consumers are relatively deserving and hence should receive more goods than others.

Voting In a democracy, important government policies that determine the allocation of goods are made by voting. Such democratic decision making is often difficult because people fundamentally disagree on how issues should be resolved and which groups of people should be favored.

In Chapter 4, we assumed that consumers can order all bundles of goods in terms of their preferences and that their rank over goods is transitive.[10] Suppose now that consumers have preferences over allocations of goods across consumers. One possibility, as we assumed earlier, is that individuals care only about how many goods they receive—they don't care about how much others have. Another possibility is that because of envy, charity, pity, love, or other interpersonal feelings, individuals do care about how much everyone has.[11]

[10]The transitivity (or *rationality*) assumption is that a consumer's preference over bundles is consistent in the sense that if the consumer weakly prefers Bundle *a* to Bundle *b* and weakly prefers Bundle *b* to Bundle *c*, the consumer weakly prefers Bundle *a* to Bundle *c*.

[11]To an economist, love is nothing more than interdependent utility functions. Thus, it's a mystery how each successive generation of economists is produced.

Let a be a particular allocation of goods that describes how much of each good an individual has. Each person can rank this allocation relative to Allocation b. For instance, individuals know whether they prefer an allocation where everyone has equal amounts of all goods to another allocation where people who work hard—or those of a particular skin color or religion—have relatively more goods than others.

Through voting, individuals express their rankings. One possible voting system requires that before the vote is taken, everyone agrees to be bound by the outcome in the sense that if a majority of people prefer Allocation a to Allocation b, then a is socially preferred to b.

Using majority voting to determine which allocations are preferred by society sounds reasonable, doesn't it? Such a system might work well. For example, if all individuals have the same transitive preferences, the social ordering has the same transitive ranking as that of each individual.

To illustrate this possibility, suppose that three people have different transitive preferences. Individual 1 prefers Allocation a to Allocation b to Allocation c. Table 10.2 shows that the other two individuals have different transitive preferred orderings.

Two out of three of these individuals prefer a to b; two out of three prefer b to c; and two out of three prefer c to a. Thus, voting leads to nontransitive societal preferences, even though the preferences of each individual are transitive. As a result, no output is clearly socially preferred. A majority of people prefers some other allocation to any particular allocation. Compared to Allocation a, a majority prefers c. Similarly, a majority prefers b over c, and a majority prefers a over b.

If people have this type of ranking of allocations, the chosen allocation will depend crucially on the order of pairwise-comparison votes. Suppose that these three people first vote on whether they prefer a or b and then compare the winner to c. The majority prefers a to b in the first vote and c to a in the second vote, so they choose c. If instead they first compare c to a and the winner (c) to b, then they choose b. Thus, the outcome depends on the political skill of various factions in determining the order of voting.

Similar problems arise with other types of voting schemes. Kenneth Arrow (1951), who received a Nobel Prize in Economics in part for his work on social decision making, proved a startling and depressing result about democratic voting. This result is often referred to as Arrow's Impossibility Theorem. Arrow suggested that a socially desirable decision-making system, or social welfare function, should satisfy the following criteria:

- Social preferences should be complete (Chapter 4) and transitive, like individual preferences.
- If everyone prefers Allocation a to Allocation b, a should be socially preferred to b.
- Society's ranking of a and b should depend only on individuals' ordering of these two allocations, not on how they rank other alternatives.
- Dictatorship is not allowed; social preferences must not reflect the preferences of only a single individual.

Table 10.2 Preferences over Allocations of Three People

	Individual 1	Individual 2	Individual 3
First choice	a	b	c
Second choice	b	c	a
Third choice	c	a	b

Although each of these criteria seems reasonable—indeed, innocuous—Arrow proved that it is impossible to find a social decision-making rule that *always* satisfies all of these criteria. His result indicates that *democratic decision making* may fail—not that *democracy* must fail. After all, if everyone agrees on a ranking, these four criteria are satisfied.

If society is willing to give up one of these criteria, a democratic decision-making rule can guarantee that the other three criteria are met. For example, if we give up the third criterion, often referred to as the *independence of irrelevant alternatives*, certain complicated voting schemes in which individuals rank their preferences can meet the other criteria.

Application

How You Vote Matters

The 15 members of a city council must decide whether to build a new road (R), repair the high school (H), or install new street lights (L). Each councilor lists the options in order of preference. Six favor L to H to R; five prefer R to H to L; and four want H over R over L.

One of the proponents of street lights suggests a plurality vote where everyone would cast a single vote for his or her favorite project. Plurality voting would result in six votes for L, five for R, and four for H, so that lights would win.

"Not so fast," responds a council member who favors roads. Given that H was the least favorite first choice, he suggests a run-off between L and R. Since the four members whose first choice was H prefer R to L, roads would win by nine votes to six.

A supporter of schools is horrified by these self-serving approaches to voting. She calls for pairwise comparisons. A majority of 10 would choose H over R, and 9 would prefer H to L. Consequently, although the high school gets the least number of first-place votes, it has the broadest appeal in pairwise comparisons.

Finally, suppose the council uses a voting method developed by Jean-Charles de Borda in 1770 (to elect members to the Academy of Sciences in Paris), where, in an n-person race, a person's first choice gets n votes, the second choice gets $n - 1$, and so forth. Here, H gets 34 votes, R receives 29, and L trails with 27, and so the high school project is backed. Thus, the outcome of an election or other vote may depend on the voting procedures used.

Methods like Borda's are called *instant runoff* voting. This method of voting is used on many issues at educational institutions such as Arizona State University, the College of William and Mary, Harvard, Southern Illinois University at Carbondale, the University of California Los Angeles, University of Michigan, University of Missouri, and Wheaton College. Instant runoffs are used to elect members of the Australian House of Representatives, the President of India, and the President of Ireland. Instant runoff voting is used in elections in many U.S. cities and counties such as Cambridge, Massachusetts; Davis, California; Oakland, California; Minneapolis, Minnesota; Pierce County, Washington; and San Francisco, California. It is also used to elect mayors in London and Wellington, New Zealand.

In the last few years, President Obama, Senator John McCain, consumer advocate Ralph Nader, and others have called for some form of ranked voting. In 2011, in a U.K. national referendum, the voters rejected instant runoffs. An instant runoff vote was used to elect the leader of the Liberal Party of Canada in 2013. In 2016, Maine adopted ranked voting.

Social Welfare Functions How would you rank various allocations if you were asked to vote? Philosophers, economists, newspaper columnists, politicians, radio talk show hosts, and other deep thinkers have suggested various rules that society might use to decide which allocations are better than others. All these rules answer the question of which individuals' preferences should be given more weight in society's decision making. Determining how much weight to give to the preferences of various members of society is usually the key step in determining a social welfare function.

Probably the simplest and most egalitarian rule is that every member of society is given exactly the same bundle of goods. If no further trading is allowed, this rule results in complete equality in the allocation of goods.

Jeremy Bentham (1748–1832) and his followers (including John Stuart Mill), the utilitarian philosophers, suggested that society should maximize the sum of the utilities of all members of society. Their social welfare function, the utilitarian welfare function, is the sum of the utilities of every member of society where the utilities of all people in society are given equal weight.[12] If U_i is the utility of Individual i and n is the number of people, the utilitarian welfare function is

$$W = U_1 + U_2 + \cdots + U_n.$$

This social welfare function may not lead to an egalitarian distribution of goods. Indeed, an allocation is judged superior, all else the same, if people who get the most pleasure from consuming certain goods are given more of those goods.

Panel b of Figure 10.9 shows some isowelfare lines corresponding to the utilitarian welfare function. These lines have a slope of -1 because the utilities of both parties are weighted equally. In the figure, welfare is maximized at e.

A generalization of the utilitarian rule assigns different weights to various individuals' utilities. If the weight assigned to Individual i is a_i, this generalized utilitarian welfare function is

$$W = a_1U_1 + a_2U_2 + \cdots + a_nU_n.$$

Society could give greater weight to adults, hardworking people, or those who meet other criteria. Under South Africa's former apartheid system, the utilities of people with white skin were given more weight than those of people with other skin colors.

John Rawls (1971), a philosopher at Harvard, believed that society should maximize the well-being of the worst-off member of society, who is the person with the lowest level of utility. In Rawls' social welfare function, all the weight should be placed on the utility of the person with the lowest utility level. The Rawlsian welfare function is

$$W = \min\{U_1, U_2, \cdots, U_n\}.$$

Rawls' rule leads to a relatively egalitarian distribution of goods.

One final rule, which is frequently espoused by various members of Congress and by wealthy landowners in less-developed countries, is to maintain the status quo. Exponents of this rule believe that the current allocation is the best possible allocation. They argue against any reallocation of resources from one individual to another. Under this rule, the final allocation is likely to be very unequal. Why else would the wealthy want it?

---

[12]It is difficult to compare utilities across individuals because the scaling of utilities across individuals is arbitrary (Chapters 4 and 9). A rule that avoids this utility comparison is to maximize a welfare measure that equally weights consumer surplus and producer surplus, which are denominated in dollars.

All of these rules or social welfare functions reflect value judgments in which interpersonal comparisons are made. Because each reflects value judgments, we cannot compare them on scientific grounds.

Efficiency Versus Equity

Given a particular social welfare function, *society might prefer an inefficient allocation to an efficient one*. We can show this result by comparing two allocations. In Allocation *a*, you have everything and everyone else has nothing. This allocation is Pareto efficient: It is impossible to make others better off without harming you. In Allocation *b*, everyone has an equal amount of all goods. Allocation *b* is not Pareto efficient: I would be willing to trade all my zucchini for just about anything else. Despite Allocation *b*'s inefficiency, most people probably prefer *b* to *a*.

Although society might prefer an inefficient Allocation *b* to an efficient Allocation *a*, according to most social welfare functions, society would prefer some efficient allocation to *b*. Suppose that Allocation *c* is the competitive equilibrium that would result if people were allowed to trade starting from Endowment *b*, in which everyone has an equal share of all goods. By the utilitarian social welfare functions, Allocation *b* might be socially preferred to Allocation *a*, but Allocation *c* is certainly socially preferred to *b*. After all, if everyone is as well off or better off in Allocation *c* than in *b*, *c* must be better than *b* regardless of weights on individuals' utilities. According to the egalitarian rule, however, *b* is preferred to *c* because only strict equality matters. Thus, by most—but not all—of the well-known social welfare functions, *society prefers an efficient allocation to an inefficient allocation*.

Competitive equilibrium may not be very equitable even though it is Pareto efficient. Consequently, societies that believe in equity may tax the rich to give to the poor. If the money taken from the rich is given directly to the poor, society moves from one Pareto-efficient allocation to another.

Sometimes, however, in an attempt to achieve greater equity, efficiency is reduced. For example, advocates for the poor argue that providing public housing to the destitute leads to an allocation that is superior to the original competitive equilibrium. This reallocation isn't efficient: The poor view themselves as better off receiving an amount of money equal to what the government spends on public housing. They could spend the money on the type of housing they like—rather than the type the government provides—or they could spend some of the money on food or other goods.[13]

Unfortunately, frequently society's goal of efficiency and its goal of an equitable allocation conflict. When the government redistributes money from one group to another, it incurs significant costs from this redistribution. If tax collectors and other government bureaucrats produced goods rather than redistributing them, total output would increase. Similarly, income taxes discourage some people from working as hard as they otherwise would (Chapter 5). Nonetheless, probably few people believe that the status quo is optimal and that the government should engage in no redistribution at all (although some legislators vote for tax laws as though they believe that we should redistribute from the poor to the rich).

[13]Letting the poor decide how to spend their income is efficient by our definition, even if they spend some of it on "sin goods" such as cigarettes, liquor, or illicit drugs. We made a similar argument about food stamps in Chapter 4.

Challenge Solution

Anti-Price Gouging Laws

We can use a multimarket model to analyze the Challenge questions about the effects of a binding price ceiling that applies to some states but not to others. The figure shows what happens if a binding price ceiling is imposed in the covered sector—those states that have anti-price gouging laws—and not in the uncovered sector—the other states.

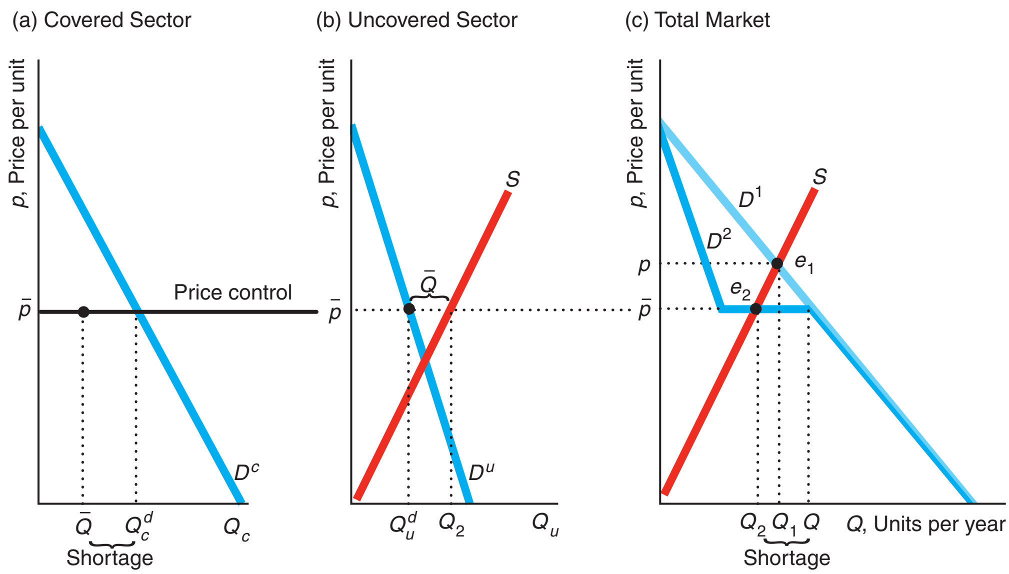
We first consider what happens in the absence of the anti-price gouging laws. The demand curve for the entire market, D^1 in panel c, is the horizontal sum of the demand curve in the covered sector, D^c in panel a, and the demand curve in the uncovered sector, D^u in panel b. In panel c, the national supply curve S intersects the national demand curve D^1 at e_1 where the equilibrium price is p and the quantity is Q_1.

When the covered sector imposes a price ceiling at $\bar{p}$, which is less than p, it chops off the top part of the D^c above $\bar{p}$. Consequently, the new national demand curve, D^2, equals the uncovered sector's demand curve D^u above $\bar{p}$, is horizontal at $\bar{p}$, and is the same as D^1 below $\bar{p}$. The supply curve S intersects the new demand curve in the horizontal section at e_2, where the quantity is Q_2.[14] However, at a price of $\bar{p}$, national demand is Q, so the shortage is $Q - Q_2$.

How the available supply Q_2 is allocated between customers in the covered and uncovered sectors determines in which sector the shortage occurs. If some of the customers in the uncovered sector cannot buy as much as they want at $\bar{p}$, they can offer to pay a slightly higher price to obtain extra supplies. Because of the price control, customers in the covered sector cannot match a higher price. Consequently, customers in the uncovered sector can buy as much as they want, Q_u^d, at $\bar{p}$, as panel b shows.

For convenience, panel b also shows the national supply curve. At $\bar{p}$, the gap between the quantity demanded in the uncovered sector, Q_u^d, and the quantity that

MyLab Economics Video



[14]If $\bar{p}$ were low enough that the supply curve hits D_2 in its downward-sloping section, suppliers would sell only in the uncovered sector. For example, in 2009 when West Virginia imposed anti-price gouging laws after flooding occurred in some parts of the state, Marathon Oil halted sales to independent gasoline retailers there and sold its gasoline in other states. Similarly, some Venezuelan firms avoid price controls by selling in neighboring Colombia (see the Chapter 2 Application "Venezuelan Price Ceilings and Shortages").

firms are willing to sell, Q_2, is $\overline{Q}$. Firms sell this extra amount, $\overline{Q}$, in the covered sector. That quantity is less than the amount demanded, Q_c^d, so the shortage in the covered sector is $Q_c^d - \overline{Q}$ $(= Q - Q_2)$.

In conclusion, the anti-price gouging law lowers the price in both sectors to $\overline{p}$, which is less than the price p that would otherwise be charged. The consumers in the uncovered states do not suffer from a shortage, in contrast to consumers in the covered sector. Thus, anti-gouging laws benefit residents of neighboring jurisdictions who can buy as much as they want at a lower price. Residents of jurisdictions with anti-gouging laws who can buy the good at a lower price benefit, but those who cannot buy the good suffer.

Summary

1. **General Equilibrium.** A general-equilibrium analysis takes account of the direct effects of a shock in a market and the spillover effects in other markets. In contrast, a partial-equilibrium analysis (such as we used in earlier chapters) looks only at one market and ignores the spillover effects in other markets. The partial-equilibrium and general-equilibrium effects can differ.
2. **Trading Between Two People.** If people make all the trades they want, the resulting equilibrium will be Pareto efficient: By moving from this equilibrium, we cannot make one person better off without harming another person. At a Pareto-efficient equilibrium, the marginal rates of substitution between people are equal because their indifference curves are tangent.
3. **Competitive Exchange.** Competition, in which all traders are price takers, leads to an allocation in which the ratio of relative prices equals the marginal rates of substitution of each person. Thus, *every competitive equilibrium is Pareto efficient*. Moreover, *any Pareto-efficient equilibrium can be obtained by competition, given an appropriate endowment*.
4. **Production and Trading.** When one person can produce more of one good and another person can produce more of another good using the same inputs, trading can result in greater combined production.
5. **Efficiency and Equity.** The Pareto efficiency criterion reflects a value judgment that a change from one allocation to another is desirable if it makes someone better off without harming anyone else. This criterion does not allow all allocations to be ranked, because some people may be better off with one allocation and others may be better off with another. Majority voting may not result in a consensus nor produce a transitive ordering of allocations. Economists, philosophers, and others have proposed many welfare functions. Society may use a welfare function to choose among Pareto-efficient (or other) allocations.

Questions

Select questions are available on MyLab Economics*; * = answer appears at the back of this book;* **A** *= algebra problem.*

1. General Equilibrium

1.1 The demand functions for the only two goods in the economy are $Q_1 = 10 - 2p_1 + p_2$ and $Q_2 = 10 - 2p_2 + p_1$. Five units of each good are available for sale. Solve for the equilibrium: p_1, p_2, Q_1, and Q_2. What is the general equilibrium? (*Hint*: See Solved Problem 10.1.) **A**

1.2 The demand functions for each of two goods depend on the prices of the goods, p_1 and p_2: $Q_1 = 15 - 3p_1 + p_2$ and $Q_2 = 6 - 2p_2 + p_1$. However, each supply curve depends on only its own price: $Q_1 = 2 + p_1$ and $Q_2 = 1 + p_2$. Solve for the equilibrium: p_1, p_2, Q_1, and Q_2. (*Hint*: See Solved Problem 10.1.) **A**

1.3 People who qualify for affordable housing in Sydney, Australia, have their rent capped at 30% of their gross annual income. People who live in private rental properties in the city are not subject to rent control. Sydney's goal is to designate 7.5% of all housing in the local area as affordable housing by 2030 to allow people who qualify for it to live close to work and maintain employment. If the total demand for rental housing does not change when rent controls are introduced for only some properties in that market, what happens to the equilibrium price and to the equilibrium quantity in the total rental housing market, the part of the market subject to rent control and the part not subject to rent control? (*Hint*: See Solved Problem 10.2.)

*1.4 What is the effect of a subsidy of s per hour on labor in only one sector of the economy on the equilibrium wage, total employment, and employment in the covered and uncovered sectors? (*Hint*: See Solved Problem 10.2.)

1.5 Governments use wage subsidies to maintain employment in times of economic recession, to encourage firms to hire and train specific groups of workers such as low-income workers, women, and students and to stimulate employment in certain regions or sectors of the economy. What effect would a wage subsidy provided to a particular sector of the economy have on the wage that workers receive and on employment in the total labor market, the covered sector, and the uncovered sector? (*Hint*: See Solved Problem 10.2.)

1.6 Competitive firms located in Africa sell their output only in Europe and the United States (which do not produce the good themselves). The industry's supply curve is upward sloping. Europe puts a tariff of t per unit on the good but the United States does not. What is the effect of the tariff on total quantity of the good sold, the quantity sold in Europe and in the United States, and equilibrium price(s)? (*Hint*: See Solved Problem 10.2.)

1.7 A competitive industry with an upward-sloping supply curve sells Q_h of its product in its home country and Q_f in a foreign country, so the total quantity it sells is $Q = Q_h + Q_f$. No one else produces this product. Shipping is costless. Determine the equilibrium price and quantity in each country. Now the foreign government imposes a binding quota, $\underline{Q}$ ($< Q_f$ at the original price). What happens to prices and quantities in both the home and the foreign markets? (*Hint*: See Solved Problem 10.2.)

1.8 The demand curve in Sector 1 of the labor market is $L_1 = a - bw$. The demand curve in Sector 2 is $L_2 = c - dw$. The supply curve of labor for the entire market is $L = e + fw$. In equilibrium, $L_1 + L_2 = L$.

a. Solve for the equilibrium with no minimum wage.

b. Solve for the equilibrium at which the minimum wage is $\underline{w}$ in Sector 1 ("the covered sector") only. (*Hint*: See Solved Problem 10.2.)

c. Solve for the equilibrium at which the minimum wage $\underline{w}$ applies to the entire labor market.

1.9 Municipalities in Belgium levy surtaxes on the personal income tax payable by their residents. Rates vary from 0% to 9%, with the average rate being about 7%. Suppose the rate of surtax on wage income is initially 5% in two neighboring municipalities; one municipality then increases the rate of its surtax to 9%. What effect would this tax increase have on the equilibrium wage and on employment in the total labor market, the higher tax municipality, and the lower tax municipality? (*Hint*: See Solved Problem 10.2.)

1.10 Suppose that a given quantity of milk is used to produce both cheese and protein powder. Cheese is made using curd from the milk, while the leftover whey is used to make protein powder, used by many people as part of their regular exercise routine. If the number of people exercising regularly and using protein powder (but not cheese) increases, how might this be reflected in the prices of protein powder and cheese?

2. Trading Between Two People

2.1 Initially, Michael has 10 candy bars and 5 cookies, and Tony has 5 candy bars and 10 cookies. After trading, Michael has 12 candy bars and 3 cookies. In an Edgeworth box, label the initial Allocation A and the new Allocation B. Draw some indifference curves that are consistent with this trade being optimal for both Michael and Tony.

2.2 Two people in a pure exchange economy have identical utility functions. Will they ever want to trade?

2.3 Both Omar and Zaid have an initial endowment of the same two goods. Omar has relatively more of Good 1 and his indifference curves are convex to the origin. In contrast, Zaid's indifference curves are concave to the origin. Can both Omar and Zaid benefit from voluntary trade? If so, will their trade result in a Pareto efficient allocation in which each has a positive quantity of both goods? (*Hint*: Consider the relative slopes of the indifference curves.)

*2.4 In a pure exchange economy with two goods, G and H, the two traders have Cobb-Douglas utility functions. Amos' utility is $U_a = (G_a)^{\alpha}(H_{\alpha})^{1-\alpha}$ and Elise's is $U_e = (G_e)^{\beta}(H_e)^{1-\beta}$. What are their marginal rates of substitution? Between them, Amos and Elise own 100 units of G and 50 units of H. Thus, if Amos has G_a and H_a, Elise has $G_e = 100 - G_a$ and $H_e = 50 - H_a$. Solve for their contract curve.

2.5 Adrienne and Sarah consume pizza, Z, and cola, C. Adrienne's utility function is $U_A = Z_A C_A$, and Sarah's is $Z_D^{0.5}C_D^{0.5}$. Adrienne's marginal utility of pizza is $MU_A^Z = C_A$. Similarly, $MU_C^A = Z_A$, $MU_Z^D = \frac{1}{2}Z_D^{-0.5}C_D^{0.5}$, and $MU_C^D = \frac{1}{2}Z_D^{0.5}C_D^{-0.5}$. Their endowments are $Z_A = 10$, $C_A = 20$, $Z_D = 20$, $C_D = 10$.

a. What are the marginal rates of substitution for each person?

b. What is the formula for the contract curve? Draw an Edgeworth box and indicate the contract curve. **A**

2.6 Explain why point *e* in Figure 10.4 is not on the contract curve. (*Hint*: See Solved Problem 10.3.)

3. Competitive Exchange

3.1 In an Edgeworth box, illustrate that a Pareto-efficient equilibrium, point *a*, can be obtained by competition, given an appropriate endowment. Do so by identifying an initial endowment point, *b*, located somewhere other than at point *a*, such that the competitive equilibrium (resulting from competitive exchange) is *a*. Explain.

4. Production and Trading

*4.1 In panel c of Figure 10.6, the joint production possibility frontier is concave to the origin. When the two individual production possibility frontiers are combined, however, the resulting *PPF* could have been drawn so that it was convex to the origin. How do we know which of these two ways of drawing the *PPF* to use?

*4.2 Pat and Chris can spend their nonleisure time working either in the marketplace or at home (preparing dinner, taking care of children, doing repairs). In the marketplace, Pat earns a higher wage, $w_p = \$20$, than Chris, $w_c = \$10$. How would marriage affect their individual and combined budget constraints (Chapters 4 and 5) and their labor-leisure choice (Section 5.5, "Deriving Labor Supply Curves"). In your discussion, take into account the theory of comparative advantage.

4.3 If Jane and Denise have identical, linear production possibility frontiers, can they benefit by trading? Why? (*Hint*: See Solved Problem 10.4.)

4.4 Modify Solved Problem 10.4 to show that the *PPF* more closely approximates a quarter of a circle if there are five people rather than three. One of the two new people, Bill, can produce five piles of wood, or four candy bars, or any linear combination. The other, Helen, can produce four piles of wood, or five candy bars, or any linear combination.

4.5 Several countries produce many of the same goods. Suppose that there are two countries, each of which produces the same two goods. Country 1 has 300 workers and Country 2 has 900 workers. The following table shows how many workers are necessary in each country to produce one unit of each good, and how much of each good each country produces when there is no trade:

	Workers Per Unit of		No Trade: Units of Production of	
	Good 1	Good 2	Good 1	Good 2
Country 1	10	50	15	3
Country 2	15	30	20	20

a. What is the maximum amount of each good that Country 1 can produce when there is no trade? How many units of each good can Country 2 produce?

b. Which country has a comparative advantage in producing Good 1 and Good 2?

c. Draw the production possibility frontier for each country when there is no trade and show where the two produce on it.

d. Draw the production possibility frontier when the two countries trade with each other.

e. Show that both countries can benefit from trade. **A**

4.6 Suppose that Britain can produce 10 units of cloth or 5 units of food per day (or any linear combination) with available resources and Greece can produce 2 units of food per day or 1 unit of cloth (or any combination). Britain has an *absolute advantage* over Greece in producing both goods. Does it still make sense for these countries to trade?

5. Efficiency and Equity

5.1 A society consists of two people with utilities U_1 and U_2, and the social welfare function is $W = a_1U_1 + a_2U_2$. Draw a utility possibility frontier similar to the ones in Figure 10.9. When social welfare is maximized, show that as a_1/a_2 increases, Person 1 benefits and Person 2 is harmed. **A**

5.2 Six people are getting together for dinner but are having trouble deciding where to go. Three of them prefer French to Indian to Thai food, two favor Thai to Indian to French food, and one prefers Indian to Thai to French food. If they decide to vote on the restaurant they will dine at, what would the result be under plurality voting, pairwise comparisons, and instant runoff voting? Assume that, under instant runoff voting, every person's first choice gets three votes, every person's second choice gets two votes, and every person's third choice gets one vote. (*Hint*: See Application "How You Vote Matters.")

5.3 Society uses various rules to decide which allocations are better than others. What rules does the generalized utilitarian welfare function use? If the generalized utilitarian welfare function were to be applied to two groups in society—the rich and the poor—instead of to individuals, how might weights be assigned so as to achieve a result similar to the Rawlsian social welfare function? (*Hint*: What distinguishes poor from rich people?)

6. Challenge

6.1 Modify the figure in the Challenge Solution to show how much would be sold in both sectors in the absence of anti-price gouging laws. Discuss how these quantities differ from those that result from implementing such laws.

6.2 The market for peaches is competitive. The market has two types of demanders: consumers who eat fresh peaches and canners. If the government places a binding price ceiling only on peaches sold directly to consumers, what happens to prices and quantities of peaches sold for each use?

6.3 Initially, a good sells in a competitive market at a single price, but a local government then restricts the quantity of the good that its citizens can buy. What happens to the equilibrium price of the good and to the equilibrium quantity in that part of the market subject to the quota and the part not subject to a quota? (*Hint*: See Solved Problem 10.2.)

6.4 A competitive industry with an upward-sloping supply curve sells Q_h of its product in its home country and Q_f in a foreign country, so the total quantity it sells is $Q = Q_h + Q_f$. No one else produces this product. The shipping cost is zero. Determine the equilibrium price and quantity in each country. Now the foreign government imposes a binding quota, $\bar{Q}$ ($< Q_f$ at the original price). What happens to prices and quantities in both the home and the foreign market?

Monopoly

11

Monopoly: one parrot.

Challenge

Brand-Name and Generic Drugs

A firm that creates a new drug may receive a patent that gives it the right to be the sole producer of that drug for up to 20 years. As a result, the firm can charge a price much greater than its marginal cost of production. For example, one of the world's best-selling drugs, the heart medication Plavix, sold for about $7 per pill, though it costs about 3¢ per pill to produce. A new drug to treat hepatitis C, Harvoni, sells for $1,350 a pill or $113,400 for a 12-week course of treatment. In 2015, Martin Shkreli, then the head of Turing Pharmaceuticals raised the price of Daraprim, used to treat infections that are common in HIV/AIDS and cancer patients, from $13.50 to $750 per pill. In the first half of 2016, Pfizer raised the average prices of its pharmaceuticals by nearly 20%.

Every year, many pharmaceuticals lose their patent protection, as Plavix has. In 2016, the patents for Crestor (cholesterol), Glumetz (diabetes), Nuvigil (attention disorder/weight loss), and other high-revenue drugs expired. In 2017, patent protection for Tamiflu (viral infections), Acthar Gel (endocrine disorders), and others end.

Generally, when a patent for a highly profitable drug expires, firms enter the market and sell generic (equivalent) versions of the brand-name drug. Generics' share of all U.S. prescriptions rose from about 18% in 1984 to 89% in 2015. In 2001, in just the first two weeks after the entry of generics, Prozac lost 73% of its market share.

The U.S. Congress, when it originally passed a law permitting generic drugs to quickly enter a market after a patent expires, expected that patent expiration would subsequently lead to sharp declines in drug prices.[1] If consumers view the generic product and the brand-name product as perfect substitutes, both goods will sell for the same price, and entry by many firms will drive the price down to the competitive level. Even if consumers view the goods as imperfect substitutes, one might expect the price of the brand-name drug to fall.

However, the prices of many brand-name drugs have increased after their patents expired and generics entered the market. The generic drugs are relatively inexpensive, but the brand-name drugs often continue to enjoy a significant market share and sell for high prices. Regan

[1]Under the 1984 Hatch-Waxman Act, the U.S. government allows a firm to sell a generic product after a brand-name drug's patent expires if the generic-drug firm can prove that its product delivers the same amount of active ingredient or drug to the body in the same way as the brand-name product. Sometimes the same firm manufactures both a brand-name drug and an identical generic drug, so the two have identical ingredients. Generics produced by other firms usually have a different appearance and name than the original product and may have different nonactive ingredients, but they have the same active ingredients.

(2008), who studied the effects of generic entry on post-patent price competition for 18 prescription drugs, found an average 2% increase in brand-name prices. Studies based on older data have found up to a 7% average increase. Why do some brand-name prices rise after the entry of generic drugs?

monopoly
the only supplier of a good that has no close substitute

A **monopoly** is the only supplier of a good that has no close substitute. Monopolies have been common since ancient times. In the fifth century B.C., the Greek philosopher Thales gained control of most of the olive presses during a year of exceptionally productive harvests. Similarly, the ancient Egyptian pharaohs controlled the sale of food. In England, until Parliament limited the practice in 1624, kings granted monopoly rights called royal charters or patents to court favorites. Today, nearly every country grants a *patent*—an exclusive right to sell that lasts for a limited period—to an inventor of a new product, process, substance, or design. Until 1999, the U.S. government gave one company the right to be the sole registrar of Internet domain names. When first introduced, Apple's iPod had a virtual monopoly in the hard-disk, music player market, and Apple's iPad had a near monopoly in the tablet market.

A monopoly can *set* its price—it is not a price taker like a competitive firm. A monopoly's output is the market output, and the demand curve a monopoly faces is the market demand curve. Because the market demand curve is downward sloping, the monopoly (unlike a competitive firm) doesn't lose all its sales if it raises its price. Consequently, the monopoly sets its price above marginal cost to maximize its profit. Consumers buy less at this high monopoly price than they would at the competitive price, which equals marginal cost.

In this chapter, we examine six main topics

1. **Monopoly Profit Maximization.** Like all firms, a monopoly maximizes its profit by setting its price or output so that its marginal revenue equals its marginal cost.
2. **Market Power.** How much the monopoly's price is above its marginal cost depends on the shape of the demand curve it faces.
3. **Market Failure Due to Monopoly Pricing.** By setting its price above marginal cost, a monopoly creates a deadweight loss.
4. **Causes of Monopoly.** Two important causes of monopoly are cost factors and government actions that restrict entry, such as patents.
5. **Government Actions That Reduce Market Power.** The welfare loss of a monopoly can be reduced or eliminated if the government regulates the price the monopoly charges or allows other firms to enter the market.
6. **Networks, Dynamics, and Behavioral Economics.** If its current sales affect a monopoly's future demand curve, a monopoly that maximizes its long-run profit may choose not to maximize its short-run profit.

11.1 Monopoly Profit Maximization

All firms, including competitive firms and monopolies, maximize their profits by setting *marginal revenue equal to marginal cost* (Chapter 8). We already know how to derive the marginal cost curve of a monopoly from its cost curve (Chapter 7). We now derive the monopoly's marginal revenue curve and then use the marginal revenue and marginal cost curves to examine the monopoly's profit-maximizing behavior.

Marginal Revenue

A firm's marginal revenue curve depends on its demand curve. We will show that a monopoly's marginal revenue curve lies below its demand curve at any positive quantity because its demand curve is downward sloping.

Marginal Revenue and Price A firm's demand curve shows the price, p, it receives for selling a given quantity, q. The price is the *average revenue* the firm receives, so a firm's revenue is $R = pq$.

A firm's *marginal revenue*, MR, is the change in its revenue from selling one more unit. A firm that earns ΔR more revenue when it sells Δq extra units of output has a marginal revenue (Chapter 8) of

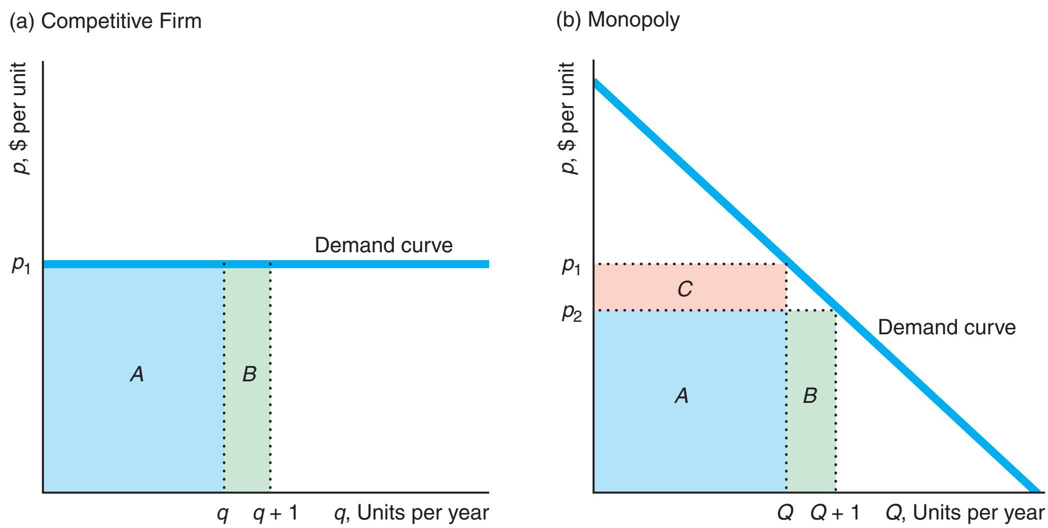
$$MR = \Delta R/\Delta q.$$

If the firm sells exactly one more unit, $\Delta q = 1$, its marginal revenue is $MR = \Delta R$.

Although the marginal revenue curve is horizontal for a competitive firm, it is downward sloping for a monopoly. The competitive firm in panel a of Figure 11.1 faces a horizontal demand curve at the market price, p_1. Because its demand curve is horizontal, the competitive firm can sell another unit of output without dropping its price. As a result, the marginal revenue it receives from selling the last unit of output is the market price.

Figure 11.1 Average and Marginal Revenue MyLab Economics Video

The demand curve shows the average revenue or price per unit of output sold. (a) The competitive firm's marginal revenue, area B, equals the market price, p_1. (b) The monopoly's marginal revenue is less than the price p_2 by area C (the revenue lost due to a lower price on the Q units originally sold).



	Initial Revenue, R_1	Revenue with One More Unit, R_2	Marginal Revenue, $R_2 - R_1$
Competition	A	$A + B$	$B = p_1$
Monopoly	$A + C$	$A + B$	$B - C = p_2 - C$

Initially, the competitive firm sells q units of output at the market price of p_1, so its revenue, R_1, is area A, which is a rectangle that is $p_1 \times q$. If the firm sells one more unit, its revenue is $R_2 = A + B$, where area B is $p_1 \times 1 = p_1$. The competitive firm's marginal revenue equals the market price:

$$\Delta R = R_2 - R_1 = (A + B) - A = B = p_1.$$

A monopoly faces a downward-sloping market demand curve, as in panel b of Figure 11.1. (We've called the number of units of output a firm sells q and the output of all the firms in a market, or market output, Q. Because a monopoly is the only firm in the market, q and Q do not differ, so we use Q to describe both the firm's and the market's output.) The monopoly, which is initially selling Q units at p_1, can sell one extra unit only if the price falls to p_2.

The monopoly's initial revenue, $p_1 \times Q$, is $R_1 = A + C$. When it sells the extra unit, its revenue, $p_2 \times (Q + 1)$, is $R_2 = A + B$. Thus, its marginal revenue is

$$\Delta R = R_2 - R_1 = (A + B) - (A + C) = B - C.$$

The monopoly sells the extra unit of output at the new price, p_2, so its extra revenue is $B = p_2 \times 1 = p_2$. The monopoly loses the difference between the new price and the original price, $\Delta p = (p_2 - p_1)$, on the Q units it originally sold: $C = \Delta p \times Q$. Thus, the monopoly's marginal revenue, $B - C = p_2 - C$, is less than the price it charges by an amount equal to area C.

The competitive firm in panel a does not lose an area C from selling an extra unit because its demand curve is horizontal. It is the downward slope of the monopoly's demand curve that causes its marginal revenue to be less than its price.

Marginal Revenue Curve Thus, *the monopoly's marginal revenue curve lies below the demand curve* at every positive quantity. In general, the relationship between the marginal revenue and demand curves depends on the shape of the demand curve.

For all *linear* demand curves, the relationship between the marginal revenue and demand curve is the same. The marginal revenue curve is a straight line that starts at the same point on the vertical (price) axis as the demand curve but has twice the slope of the demand curve, so the marginal revenue curve hits the horizontal (quantity) axis at half the quantity as the demand curve (see Appendix 11A). In Figure 11.2, the demand curve has a slope of -1 and hits the horizontal axis at 24 units, while the marginal revenue curve has a slope of -2 and hits the horizontal axis at 12 units.

Deriving the Marginal Revenue Curve To derive the monopoly's marginal revenue curve, we write an equation summarizing the relationship between price and marginal revenue that panel b of Figure 11.1 illustrates. (Because we want this equation to hold at all prices, we drop the subscripts from the prices.) For a monopoly to increase its output by ΔQ, the monopoly lowers its price per unit by $\Delta p/\Delta Q$, which is the slope of the demand curve. By lowering its price, the monopoly loses $(\Delta p/\Delta Q) \times Q$ on the units it originally sold at the higher price (area C), but it earns an additional p on the extra output it now sells (area B). Thus, the monopoly's marginal revenue is[2]

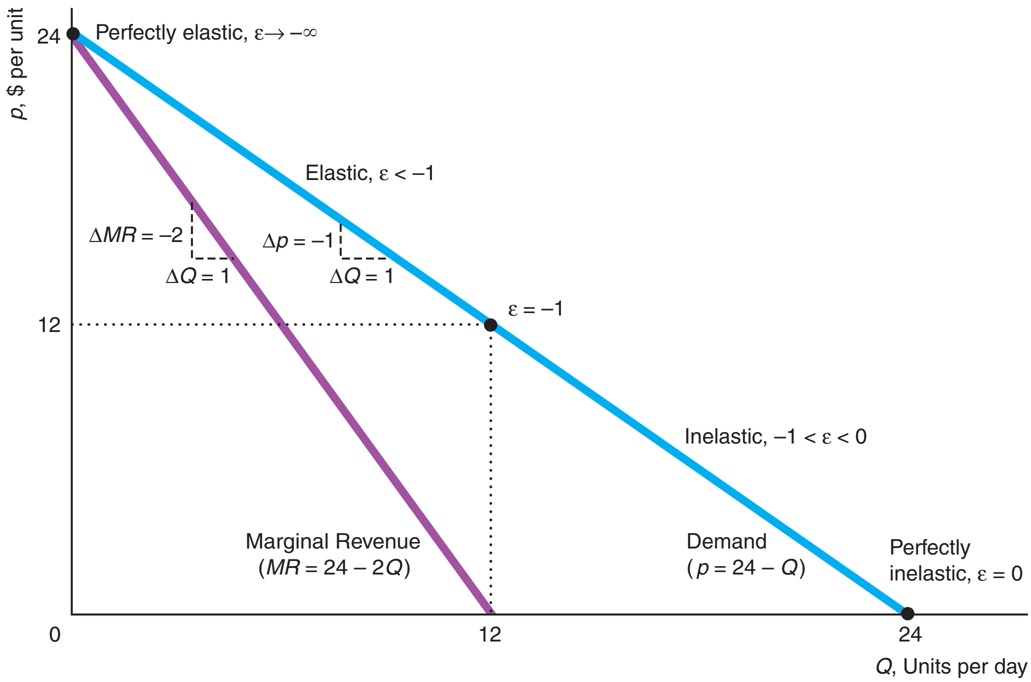
$$MR = p + \frac{\Delta p}{\Delta Q} Q. \tag{11.1}$$

Because the slope of the monopoly's inverse demand curve, $\Delta p/\Delta Q$, is negative, the last term in Equation 11.1, $(\Delta p/\Delta Q)Q$, is negative. Equation 11.1 confirms that the price is greater than the marginal revenue, which equals p plus a negative term.

[2]Revenue is $R(Q) = p(Q)Q$, where $p(Q)$, the inverse demand function, shows how price changes as quantity increases along the demand curve. Differentiating, we find that the marginal revenue is $MR = \mathrm{d}R(Q)/\mathrm{d}Q = p(Q) + [\mathrm{d}p(Q)/\mathrm{d}Q]Q$.

Figure 11.2 Elasticity of Demand and Total, Average, and Marginal Revenue

The demand curve (or average revenue curve), $p = 24 - Q$, lies above the marginal revenue curve, $MR = 24 - 2Q$. Where the marginal revenue equals zero, $Q = 12$, the elasticity of demand is $= -1$.



Solved Problem 11.1

Derive the marginal revenue curve when the monopoly faces the linear inverse demand function,

$$p = 24 - Q, \tag{11.2}$$

in Figure 11.2. How does the slope of the marginal revenue curve compare to the slope of the inverse demand curve?

Answer

1. *Use the demand curve to calculate how much the price consumers are willing to pay falls if quantity increases by one unit.* According to the inverse demand function, Equation 11.2, the price consumers are willing to pay falls by 1 if quantity increases by one unit, so the slope of the inverse demand curve is $\Delta p/\Delta Q = -1$ (Chapter 2).[3]

[3]In general, if the linear inverse demand curve is $p = a - bQ$ and the quantity increases from Q to $Q + \Delta Q$, then the new price is $p^* = a - b(Q + \Delta Q) = a - bQ - b\Delta Q = p - b\Delta Q$, so $\Delta p = p^* - p = -b\Delta Q$. By dividing both sides of this expression by ΔQ, we find that the slope of the demand curve is $\Delta p/\Delta Q = -b$. Here, $b = 1$, so $\Delta p/\Delta Q = -1$. Equivalently, we can use calculus to determine that the slope of the general linear demand curve is $dp/dQ = -b$.

2. *Use Equations 11.1 and 11.2 and the slope of the inverse demand curve to derive the marginal revenue function.* We obtain the marginal revenue function for this monopoly by substituting into Equation 11.1 the slope of the inverse demand function, $\Delta p/\Delta Q = -1$, and replacing p with $24 - Q$ (using Equation 11.2):

$$MR = p + \frac{\Delta p}{\Delta Q}Q = (24 - Q) + (-1)Q = 24 - 2Q. \tag{11.3}$$

   The MR curve in Figure 11.2 is a plot of Equation 11.3.

3. *Use Equation 11.3 to determine the slope of the marginal revenue curve.* Using the same type of calculation as in Step 1, we can use Equation 11.3 to show that the slope of this marginal revenue curve is $\Delta MR/\Delta Q = -2$, so the marginal revenue curve is twice as steeply sloped as is the demand curve.

Marginal Revenue and Price Elasticity of Demand The marginal revenue at any given quantity depends on the demand curve's height (the price) and shape. The shape of the demand curve at a particular quantity is described by the price elasticity of demand (Chapter 3), $\quad = (\Delta Q/Q)/(\Delta p/p) < 0$, which is the percentage by which quantity demanded falls as the price increases by 1%.

At a given quantity, the marginal revenue equals the price times a term involving the elasticity of demand:[4]

$$MR = p\left(1 + \frac{1}{\quad}\right). \tag{11.4}$$

According to Equation 11.4, marginal revenue is closer to price as demand becomes more elastic. Where the demand curve hits the price axis ($Q = 0$), the demand curve is perfectly elastic, so the marginal revenue equals price: $MR = p$.[5] Where the demand elasticity is unitary, $\quad = -1$, marginal revenue is zero: $MR = p[1 + 1/(-1)] = 0$. Marginal revenue is negative where the demand curve is inelastic, $-1 < \quad \le 0$.

With the demand function in Equation 11.2, $\Delta Q/\Delta p = -1$, so the elasticity of demand is $\quad = (\Delta Q/\Delta p)(p/Q) = -p/Q$. Table 11.1 shows the relationship among quantity, price, marginal revenue, and elasticity of demand for this linear example. As Q approaches 24, $\quad$ approaches 0, and marginal revenue is negative. As Q approaches zero, the demand becomes increasingly elastic, and marginal revenue approaches the price.

---

[4]By multiplying the last term in Equation 11.1 by $p/p(=1)$ and using algebra, we can rewrite the expression as

$$MR = p + p\frac{\Delta p}{\Delta Q}\frac{Q}{p} = p\left[1 + \frac{1}{(\Delta Q/\Delta p)(p/Q)}\right].$$

The last term in this expression is $1/\quad$, because $\quad = (\Delta Q/\Delta p)(p/Q)$.

[5]As $\quad$ approaches $-\infty$ (perfectly elastic demand), the $1/\quad$ term approaches zero, so $MR = p(1 + 1/\quad)$ approaches p.

Table 11.1 Quantity, Price, Marginal Revenue, and Elasticity for the Linear Inverse Demand Curve $p = 24 - Q$

Quantity, Q	Price, p	Marginal Revenue, MR	Elasticity of Demand, $\varepsilon = -p/Q$	
0	24	24	$-\infty$	more elastic ↑
1	23	22	−23	
2	22	20	−11	
3	21	18	−7	
4	20	16	−5	
5	19	14	−3.8	
6	18	12	−3	
7	17	10	−2.43	
8	16	8	−2	
9	15	6	−1.67	
10	14	4	−1.4	
11	13	2	−1.18	
12	**12**	**0**	**−1**	
13	11	−2	−0.85	
...	...	...	...	
23	1	−22	−0.043	
24	0	−24	0	less elastic ↓

Choosing Price or Quantity

Any firm maximizes its profit by operating where its marginal revenue equals its marginal cost. Unlike a competitive firm, a monopoly can adjust its price, so it has a choice of setting its price *or* its quantity to maximize its profit. (A competitive firm sets its quantity to maximize profit because it cannot affect market price.)

The monopoly is constrained by the market demand curve. Because the demand curve slopes down, the monopoly faces a trade-off between a higher price and a lower quantity or a lower price and a higher quantity. The monopoly chooses the point on the demand curve that maximizes its profit. Unfortunately for the monopoly, it cannot set both its quantity and its price. If it could do so, the monopoly would choose an extremely high price and an extremely high output level—above its demand curve—and would become exceedingly wealthy.

If the monopoly sets its price, the demand curve determines how much output it sells. If the monopoly picks an output level, the demand curve determines the price. Because the monopoly wants to operate at the price and output at which its profit is maximized, it chooses the same profit-maximizing solution whether it sets the price or output. In the rest of this chapter, we assume that the monopoly sets quantity.

Graphical Approach

All firms, including monopolies, use a two-step analysis to determine the output level that maximizes their profit (Chapter 8). First, the firm determines the output, Q^*, at which it makes the highest possible profit—the output at which its marginal revenue equals its marginal cost. Second, the firm decides whether to produce Q^* or shut down.

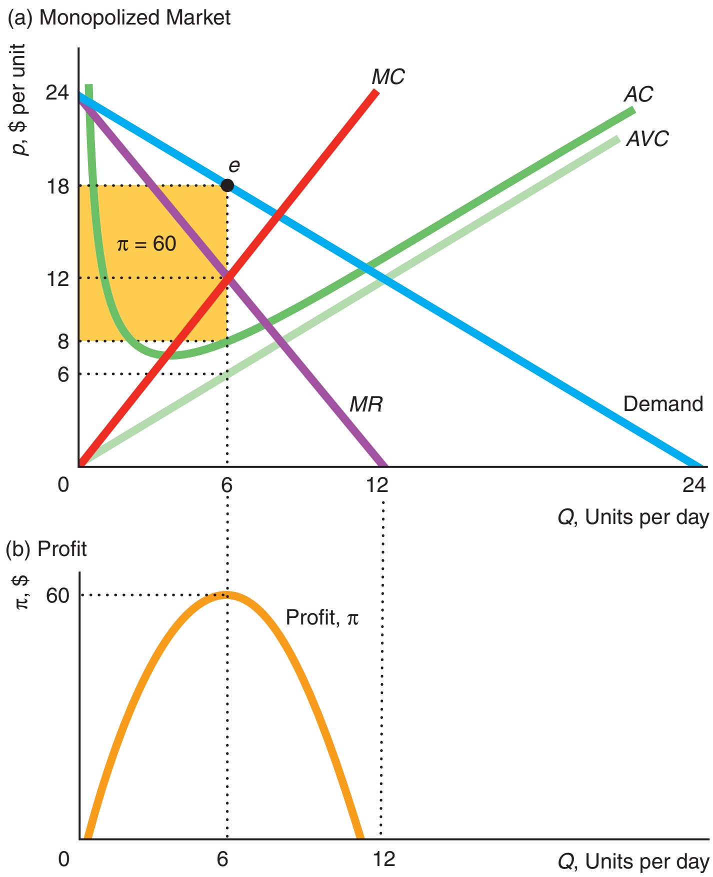
Profit-Maximizing Output To illustrate how a monopoly chooses its output to maximize its profit, we continue to use the same linear demand and marginal revenue curves but add a linear marginal cost curve in panel a of Figure 11.3. Panel b shows the corresponding profit curve. The profit curve reaches its maximum at 6 units of output, where marginal profit—the slope of the profit curve—is zero. Because *marginal profit is marginal revenue minus marginal cost* (Chapter 8), marginal profit is zero where marginal revenue equals marginal cost. In panel a, marginal revenue equals marginal cost at 6 units. The price on the demand curve at that quantity is \$18. Thus, the monopoly maximizes its profit at point *e*, where it sells 6 units per day at a price of \$18 per unit.

Why does the monopoly maximize its profit by producing where its marginal revenue equals its marginal cost? At smaller quantities, the monopoly's marginal revenue is greater than its marginal cost, so its marginal profit is positive—the profit curve is upward sloping. By increasing its output, the monopoly raises its profit. Similarly, at quantities greater than 6 units, the monopoly's marginal cost is greater than its marginal revenue, so its marginal profit is negative, and the monopoly can increase its profit by reducing its output.

As Figure 11.2 illustrates, the marginal revenue curve is positive where the elasticity of demand is elastic, it is zero at the quantity where the demand curve has a unitary elasticity, and it is negative at larger quantities where the demand curve is inelastic. Because the marginal cost curve is never negative, the marginal revenue

Figure 11.3 Maximizing Profit MyLab Economics Video

(a) At $Q = 6$, where marginal revenue, MR, equals marginal cost, MC, profit is maximized. The rectangle shows that the profit is \$60, where the height of the rectangle is the average profit per unit, $p - AC = \$18 - \$8 = \$10$, and the length is the number of units, 6. (b) Profit is maximized at $Q = 6$, where marginal revenue equals marginal cost.



curve can only intersect the marginal cost curve where the marginal revenue curve is positive, in the range where the demand curve is elastic. That is, *a monopoly's profit is maximized in the elastic portion of the demand curve*. (In our example, profit is maximized at $Q = 6$, where the elasticity of demand is -3.) *A profit-maximizing monopoly never operates in the inelastic portion of its demand curve.*

Shutdown Decision A monopoly shuts down to avoid making a loss in the short run if its price is below its average variable cost at its profit-maximizing (or loss-minimizing) quantity (Chapter 8). In the long run, the monopoly shuts down if the price is less than its average cost.

In the short-run example in Figure 11.3, the average variable cost, $AVC = \$6$, is less than the price, $p = \$18$, at the profit-maximizing output, $Q = 6$, so the firm chooses to produce. Price is also above average cost at $Q = 6$, so the monopoly makes a positive profit.[6] At the profit-maximizing quantity of 6 units, the price is $p(6) = \$18$ and the average cost is $AC(6) = \$8$. As a result, the profit, $\pi = \$60$, isthe golden rectangle with a height equal to the average profit per unit, $p(6) - AC(6) = \$18 - \$8 = \$10$, and a width of 6 units.

Mathematical Approach

We can also solve for the profit-maximizing quantity mathematically. We already know the demand and marginal revenue functions for this monopoly. We need to determine its marginal cost curve. The monopoly's cost is a function of its output, $C(Q)$. In Figure 11.3, we assume that the monopoly faces a short-run cost function of

$$C(Q) = Q^2 + 12, \tag{11.5}$$

where Q^2 is the monopoly's variable cost as a function of output and 12 is its fixed cost (Chapter 7). Given this cost function, Equation 11.5, the monopoly's marginal cost function is[7]

$$MC = 2Q. \tag{11.6}$$

This marginal cost curve is a straight line through the origin with a slope of 2 in panel a. The average variable cost is $AVC = Q^2/Q = Q$, so it is a straight line through the origin with a slope of 1. The average cost is $AC = C/Q = (Q^2 + 12)/Q = Q + 12/Q$, which is U-shaped.

We determine the profit-maximizing output by equating the marginal revenue (Equation 11.3) and marginal cost (Equation 11.6) functions:

$$MR = 24 - 2Q = 2Q = MC.$$

Solving for Q, we find that $Q = 6$. Substituting $Q = 6$ into the inverse demand function (Equation 11.2), we learn that the profit-maximizing price is

$$p = 24 - Q = 24 - 6 = \$18.$$

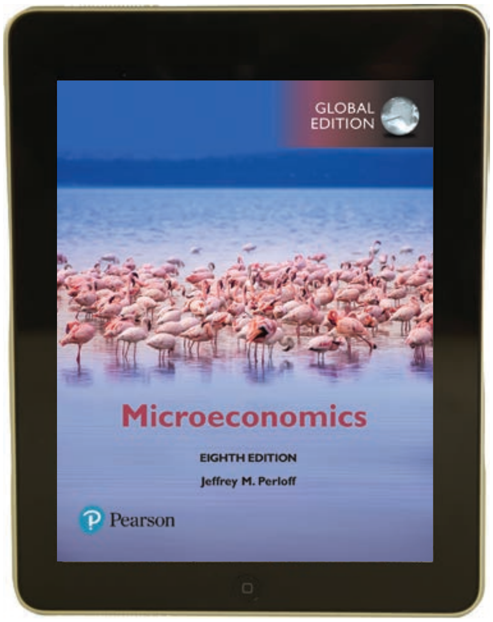
At that quantity, the average variable cost is $AVC = \$6$, which is less than the price, so the firm does not shut down. The average cost is $AC = 6 + 12/6 = \$8$, which is less than the price, so the firm makes a profit.

[6]Because profit is $\pi = p(Q)Q - C(Q)$, average profit is $\pi/Q = p(Q) - C(Q)/Q = p(Q) - AC$. Thus, average profit (and hence profit) is positive only if price is above average cost.

[7]By differentiating Equation 11.5 with respect to output, we find that the marginal cost is $MC = dC(Q)/dQ = 2Q$.

Application

Apple's iPad



Apple started selling the iPad on April 3, 2010. The iPad was not the first tablet.[8] But it was the most elegant one, and the first one that large numbers of consumers wanted to own. The iPad was a pioneer in a multi-touch, finger-sensitive touchscreen (rather than a pressure-triggered stylus) and a virtual onscreen keyboard. Most importantly, the iPad offered an intuitive interface, which allows the user easy access to Apple's iTunes, ebooks, and various application programs.

People loved the original iPad. Even at $499 for the basic model, Apple had a virtual monopoly in its first year. According to the research firm IDC, Apple's share of the 2010 tablet market was 87%. Moreover, most consumers didn't view the competing tablets as close substitutes. Apple sold 25 million iPads worldwide in its first full year, 2010–2011. According to one estimate, the basic iPad's marginal cost was only $220.

Unfortunately for Apple, its monopoly was short lived. Within a year of the iPad's introduction, over a hundred iPad want-to-be tablets were available. Apple's share of the tablet market fell to 26% by early 2016.

Solved Problem 11.2

MyLab Economics Solved Problem

When the iPad was introduced, Apple's constant marginal cost of producing its top-of-the-line iPad was about $220, its fixed cost was $2,000 million (= $2 billion), and we estimate that its inverse demand function was $p = 770 - 11Q$, where Q is the millions of iPads purchased.[9] What were Apple's average variable cost and average cost functions? What was its marginal revenue function? What were its profit-maximizing price and quantity? What was its profit? Show Apple's profit-maximizing solution in a figure.

Answer

1. *Derive the average cost function using the information about Apple's marginal and fixed costs.* Given that Apple's marginal cost was constant, its average variable cost, AVC, equaled its marginal cost, $220. Its average fixed cost, AFC, was its fixed cost divided by the quantity produced, $2{,}000/Q$. Thus, its average cost was $AC = AVC + AFC = 220 + 2{,}000/Q$.

2. *Derive Apple's marginal revenue function using the information about its demand function.* Given that Apple's inverse demand function was linear, $p = 770 - 11Q$, its marginal revenue function has the same intercept and twice the slope: $MR = 770 - 22Q$.[10]

[8]Indeed, the iPad wasn't Apple's first tablet. Apple sold another tablet, the Newton, from 1993–1998.

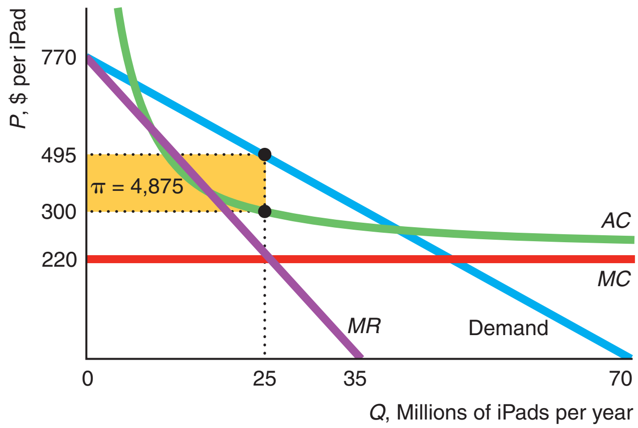
[9]See the Sources for the "Apple's iPad" Application for details on these estimates.

[10]We can derive the marginal revenue curve using calculus. Multiply the inverse demand function by Q to obtain Apple's revenue function, $R = 770Q - 11Q^2$. The marginal revenue function is the derivative of the revenue function with respect to quantity: $MR = dR/dQ = 770 - 22Q$.

3. *Derive Apple's profit-maximizing quantity and price by equating the marginal revenue and marginal cost functions and solving.* Apple maximized its profit where

$$MR = 770 - 22Q = 220 = MC.$$

Solving this equation for the profit-maximizing output, we find that $Q = 25$ million iPads, as the figure illustrates. By substituting this quantity into the inverse demand equation, we determine that the profit-maximizing price was $p = \$495$ per unit as the figure shows.



4. *Calculate Apple's profit using the profit-maximizing price and quantity and the average cost.* At $Q = 25$, the firm's average cost was $AC = 220 + 2{,}000/25 = \$300$. The firm's profit was $\pi = (p - AC)Q = (495 - 300)25 = \$4{,}875$ million. The figure shows that the profit is a rectangle with a height of $(p - AC) = \$195$ and a length of $Q = 25$ million.

Effects of a Shift of the Demand Curve

Shifts in the demand curve or marginal cost curve affect the monopoly optimum and can have a wider variety of effects in a monopolized market than in a competitive market. In a competitive market, the effect of a shift in demand on a competitive firm's output depends only on the shape of the marginal cost curve (Chapter 8). In contrast, the effect of a shift in demand on a monopoly's output depends on the shapes of both the marginal cost curve and the demand curve.

As we saw in Chapter 8, a competitive firm's marginal cost curve tells us everything we need to know about the amount that firm will supply at any given market price. The competitive firm's supply curve is its upward-sloping marginal cost curve (above its minimum average variable cost). A competitive firm's supply behavior does not depend on the shape of the market demand curve because it always faces a horizontal demand curve at the market price. Thus, if you know a competitive firm's marginal cost curve, you can predict how much that firm will produce at any given market price.

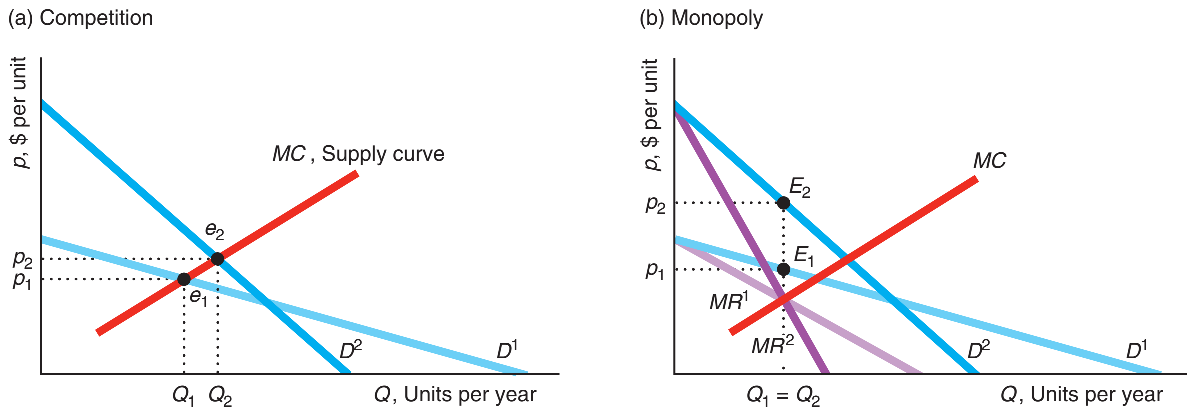
In contrast, a monopoly's output decision depends on the shapes of its marginal cost curve and its demand curve. Unlike a competitive firm, *a monopoly does not have a supply curve*. Knowing the monopoly's marginal cost curve is not enough for us to predict how much a monopoly will sell at any given price.

Figure 11.4 illustrates that the relationship between price and quantity is unique in a competitive market but not in a monopoly market. If the market is competitive, the initial equilibrium is e_1 in panel a, where the original demand curve D^1 intersects the supply curve, MC, which is the sum of the marginal cost curves of a large number of competitive firms. When the demand curve shifts to D^2, the new competitive equilibrium, e_2, has a higher price and quantity. A shift of the demand curve maps out competitive equilibria along the marginal cost curve, so every equilibrium quantity has a single corresponding equilibrium price.

Now consider the monopoly example in panel b. As the demand curve shifts from D^1 to D^2, the monopoly optimum shifts from E_1 to E_2, so the price rises but the quantity stays constant, $Q_1 = Q_2$. Thus, *a given quantity can correspond to more than one monopoly-optimal price*. A shift in the demand curve may cause the monopoly-optimal price to stay constant and the quantity to change or both price and quantity to change.

Figure 11.4 Effects of a Shift of the Demand Curve MyLab Economics Video

(a) A shift of the demand curve from D^1 to D^2 causes the competitive equilibrium to move from e_1 to e_2 along the supply curve (the horizontal sum of the marginal cost curves of all the competitive firms). Because the competitive equilibrium lies on the supply curve, each quantity corresponds to only one possible equilibrium price. (b) With a monopoly, this same shift of demand causes the monopoly optimum to change from E_1 to E_2. The monopoly quantity stays the same, but the monopoly price rises. Thus, a shift in demand does not map out a unique relationship between price and quantity in a monopolized market: The same quantity, $Q_1 = Q_2$, is associated with two different prices, p_1 and p_2.



11.2 Market Power

market power
the ability of a firm to charge a price above marginal cost and earn a positive profit

A monopoly has **market power**: the ability of a firm to charge a price above marginal cost and earn a positive profit. We now examine the factors that determine how much above its marginal cost a monopoly sets its price.

Many people falsely believe that the biggest monopolies have the most power over prices:

Common Confusion: The larger the monopoly, the more it can markup its price over its cost.

Size doesn't matter. Rather, a monopoly marks up its price more over its marginal cost, the less sensitive consumers are to price: the less elastic is the demand curve. For example, the drugs mentioned in the Challenge at the beginning of the chapter that have extremely high prices are not widely sold, but they are crucial for a small segment of the population.

Market Power and the Shape of the Demand Curve

The degree to which the monopoly raises its price above its marginal cost depends on the shape of the demand curve at the profit-maximizing quantity. If the monopoly faces a highly elastic—nearly flat—demand curve at the profit-maximizing quantity, it loses substantial sales if it raises its price by even a small amount. Conversely, if the demand curve is not very elastic (relatively steep) at that quantity, the monopoly loses fewer sales from raising its price by the same amount.

We can derive the relationship between market power and the elasticity of demand at the profit-maximizing quantity using the expression for marginal revenue in Equation 11.4 and the firm's profit-maximizing condition that marginal revenue equals marginal cost:

$$MR = p\left(1 + \frac{1}{\;}\right) = MC. \tag{11.7}$$

By rearranging terms, we can rewrite Equation 11.7 as

$$\frac{p}{MC} = \frac{1}{1 + (1/\;)}. \tag{11.8}$$

Equation 11.8 says that the ratio of the price to marginal cost depends *only* on the elasticity of demand at the profit-maximizing quantity.

In our linear demand example in panel a of Figure 11.3, the elasticity of demand is $\;= -3$ at the monopoly optimum, point e, where $Q = 6$. As a result, the ratio of price to marginal cost is $p/MC = 1/[1 + 1/(-3)] = 1.5$, or $p = 1.5MC$. The profit-maximizing price, \$18, in panel a is 1.5 times the marginal cost of 12.

Table 11.2 illustrates how the ratio of price to marginal cost varies with the elasticity of demand. When the elasticity is -1.01, only slightly elastic, the monopoly's profit-maximizing price is 101 times larger than its marginal cost: $p/MC = 1/[1 + 1/(-1.01)] \approx 101$. As the elasticity of demand approaches negative infinity (becomes perfectly elastic), the ratio of price to marginal cost shrinks to $p/MC = 1$.[11]

Table 11.2 Elasticity of Demand, Price, and Marginal Cost

	Elasticity of Demand,	Price/Marginal Cost Ratio, $p/MC = 1/[1 + (1/\varepsilon)]$	Lerner Index, $(p - MC)/p = -1/\varepsilon$
↑ less elastic	-1.01	101	0.99
	-1.1	11	0.91
	-2	2	0.5
	-3	1.5	0.33
	-5	1.25	0.2
	-10	1.11	0.1
	-100	1.01	0.01
more elastic ↓	$-\infty$	1	0

[11]As the elasticity approaches negative infinity, $1/\;$ approaches zero, so $1/(1 + 1/\;)$ approaches $1/1 = 1$.

This table illustrates that not all monopolies can set high prices. A monopoly that faces a horizontal, perfectly elastic demand curve sets its price equal to its marginal cost—just like a price-taking, competitive firm. If this monopoly were to raise its price, it would lose all its sales, so it maximizes its profit by setting its price equal to its marginal cost.

The more elastic the demand curve at the optimum, the less a monopoly can raise its price without losing sales. All else the same, the more close substitutes for the monopoly's good, the more elastic is the demand curve at the optimum. For example, the publisher Pearson has the monopoly right to produce and sell this textbook. However, many other publishers have the rights to produce and sell similar microeconomics texts (though you wouldn't like them as much). The demand curve that Pearson faces is much more elastic than it would be if no substitutes were available. If you think this textbook is expensive, imagine the cost if no substitutes were published!

Application

Cable Cars and Profit Maximization

Since San Francisco's cable car system started operating in 1873, it has been one of the city's main tourist attractions. In mid-2005, the cash-strapped Municipal Railway raised the one-way fare by two-thirds from \$3 to \$5. Not surprisingly, the number of riders dropped substantially, and many residents called for a rate reduction.

The rate increase prompted many locals to switch to buses or other forms of transportation, but most tourists have a relatively inelastic demand curve for cable car rides. Frank Bernstein of Arizona, who visited San Francisco with his wife, two children, and mother-in-law, said that they would not visit San Francisco without riding a cable car: "That's what you do when you're here." But the \$50 cost for his family to ride a cable car from the Powell Street turnaround to Fisherman's Wharf and back "is a lot of money for our family. We'll do it once, but we won't do it again."

If the city ran the cable car system like a profit-maximizing monopoly, the decision to raise fares would be clearer. The 67% rate hike resulted in a 23% increase in revenue to \$9,045,792 in the 2005–2006 fiscal year. For a reduction in rides (output) to raise revenue, the city must have been operating in the inelastic portion of its demand curve (> -1) where $MR = p(1 + 1/\ ) < 0$ prior to the fare increase. With fewer riders, costs stay constant or fall (if the city chooses to run fewer than its traditional 40 cars). Thus, its profit must increase.

However, the city was not interested in maximizing its profit on the cable cars. Mayor Gavin Newsom said that having fewer riders "was my biggest fear when we raised the fare. I think we're right at the cusp of losing visitors who come to San Francisco and want to enjoy a ride on a cable car." The mayor believed that enjoyable and inexpensive cable car rides attracted tourists to the city, thereby benefiting many local businesses.[12] Mayor Newsom observed, "Cable cars are so fundamental to the lifeblood of the city, and they represent so much more than the revenue they bring in."

The mayor decided to continue to run the cable cars at a price below the profit-maximizing level. The fare stayed at \$5 for six years while Mayor Newsom was in office. After he left, the fare rose to \$6 in 2011 and to \$7 in 2015, where it remained through 2016.

[12]That is, the mayor believed that cable cars provide a positive externality; see Chapter 17.

Lerner Index

Lerner Index
the ratio of the difference between price and marginal cost to the price: $(p - MC)/p$

Another way to show how the elasticity of demand affects a monopoly's price relative to its marginal cost is to look at the firm's **Lerner Index** (or *price markup*): the ratio of the difference between price and marginal cost to the price: $(p - MC)/p$. This measure is zero for a competitive firm because a competitive firm cannot raise its price above its marginal cost. The greater the difference between price and marginal cost, the larger the Lerner Index and the greater the monopoly's ability to set price above marginal cost.

If the firm is maximizing its profit, we can express the Lerner Index in terms of the elasticity of demand by rearranging Equation 11.8:

$$\frac{p - MC}{p} = -\frac{1}{\;}. \tag{11.9}$$

Because $MC \geq 0$ and $p \geq MC$, $0 \leq p - MC \leq p$, so the Lerner Index ranges from 0 to 1 for a profit-maximizing firm.[13] Equation 11.9 confirms that a competitive firm has a Lerner Index of zero because its demand curve is perfectly elastic.[14] As Table 11.2 illustrates, the Lerner Index for a monopoly increases as the demand becomes less elastic at the monopoly optimum. If $\; = -5$, the monopoly's markup (Lerner Index) is $1/5 = 0.2$; if $\; = -2$, the markup is $1/2 = 0.5$; and if $\; = -1.01$, the markup is 0.99. Monopolies that face demand curves that are only slightly elastic set prices that are multiples of their marginal cost and have Lerner Indexes close to 1.

Solved Problem 11.3

In 2016, Microsoft's Surface Pro 4 laptop and tablet sold for \$735. Its marginal cost was about \$470.[15] What was its Lerner Index? If it was operating at the short-run profit-maximizing level, what was the elasticity of demand for the Surface Pro?

Answer

1. *Determine the Lerner Index by substituting into the Lerner definition.* The Surface Pro's Lerner Index is

$$\frac{p - MC}{p} = \frac{735 - 470}{735} \approx 0.361.$$

2. *Use Equation 11.9 to infer the elasticity.* According to that equation, a profit-maximizing monopoly operates where $(p - MC)/p = -1/\;$. Combining that equation with the Lerner Index from the previous step, we learn that $0.361 = -1/\;$, or $\; \approx -2.77$.

Sources of Market Power

When will a monopoly face a relatively elastic demand curve and hence have little market power? Ultimately, the elasticity of demand of the market demand curve depends on consumers' tastes and options. The more consumers want a good—the more willing they are to pay "virtually anything" for it—the less elastic is the demand curve.

[13]For the Lerner Index to be above 1, would have to be a negative fraction, indicating that the demand curve was inelastic at the monopoly optimum. However, a profit-maximizing monopoly never operates in the inelastic portion of its demand curve.

[14]As the elasticity of demand approaches negative infinity, the Lerner Index, $-1/\;$, approaches zero.

[15]Andrew Rassweiler, "Teardown: Microsoft – Surface Pro 4," **electronics360.globalspec.com**, based on an IHS Technology analysis.

All else the same, the demand curve a firm (not necessarily a monopoly) faces becomes more elastic as (1) *better substitutes* for the firm's product are introduced, (2) *more firms* enter the market selling the same product, or (3) firms that provide the same service *locate closer* to this firm. The demand curves for Xerox, the U.S. Postal Service, and McDonald's have become more elastic in recent decades for these three reasons.

When Xerox started selling its plain-paper copier, no other firm sold a close substitute. Other companies' machines produced copies on special slimy paper that yellowed quickly. As other firms developed plain-paper copiers, the demand curve that Xerox faced became more elastic.

In the past, the U.S. Postal Service (USPS) had a monopoly in overnight delivery services. Now FedEx, United Parcel Service, and many other firms compete with the USPS in providing overnight deliveries. Because of these increases in competition, the USPS's share of business and personal correspondence fell from 77% in 1988 to 59% in 1996. Its total mail volume fell 40% from 2006 to 2010. Its share of the ground shipping market fell to 16% by 2012 (FedEx has about a third and UPS about half of the market).[16] Compared to when it was a monopoly, the USPS's demand curves for first-class mail and package delivery have shifted downward and become more elastic.

As you drive down a highway, you may notice that McDonald's restaurants are located miles apart. The purpose of this spacing is to reduce the likelihood that two McDonald's outlets will compete for the same customer. Although McDonald's can prevent its own restaurants from competing with each other, it cannot prevent Wendy's or Burger King from locating near its restaurants. As other fast-food restaurants open near a McDonald's, that restaurant faces a more elastic demand. What happens as a profit-maximizing monopoly faces more elastic demand? It has to lower its price.

market failure
inefficient production or consumption, often because a price exceeds marginal cost

11.3 Market Failure Due to Monopoly Pricing

Of course you could get it done for less if I weren't the only plumber in town.

Unlike perfect competition, which achieves *economic efficiency* by maximizing welfare, W (= consumer surplus + producer surplus = $CS + PS$), a profit-maximizing monopoly is economically inefficient because it wastes potential surplus, resulting in a deadweight loss, DWL. The inefficiency of monopoly pricing is an example of a **market failure**: inefficient production or consumption, often because a price exceeds marginal cost (Chapter 9). The market failure from a monopoly occurs because its price is greater than its marginal cost. This economic inefficiency creates a rationale for governments to intervene.

We illustrate this loss using our continuing example. If the monopoly were to act like a competitive market and operate where its inverse demand curve, Equation 11.2, intersects its marginal cost (supply) curve, Equation 11.6,

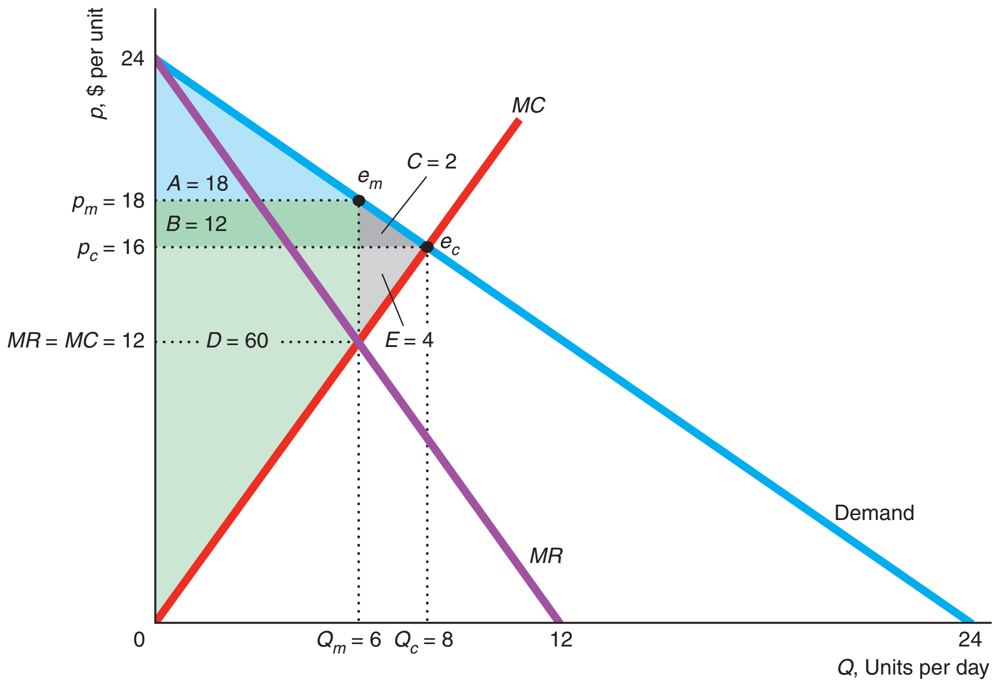
$$p = 24 - Q = 2Q = MC,$$

it would sell $Q_c = 8$ units of output at a price of 16, as in Figure 11.5. At this competitive price, consumer surplus is area $A + B + C$ and producer surplus is $D + E$.

[16]Peter Passell, "Battered by Its Rivals," *New York Times*, May 15, 1997, C1; Grace Wyler, "11 Things You Should Know about the U.S. Postal Service Before It Goes Bankrupt," *Business Insider*, May 31, 2011; "The U.S. Postal Service Nears Collapse," *Bloomberg Businessweek*, May 26, 2011; **http://www.economicfreedom.org/2012/12/12/stamping-out-waste.**

Figure 11.5 Deadweight Loss of Monopoly MyLab Economics Video

A competitive market would produce $Q_c = 8$ at $p_c = \$16$, where the demand curve intersects the marginal cost (supply) curve. A monopoly produces only $Q_m = 6$ at $p_m = \$18$, where the marginal revenue curve intersects the marginal cost curve. Under monopoly, consumer surplus is A, producer surplus is $B + D$, and the lost welfare or deadweight loss of monopoly is $-C - E$.



	Competition	Monopoly	Change
Consumer Surplus, *CS*	$A + B + C$	A	$-B - C = \Delta CS$
Producer Surplus, *PS*	$D + E$	$B + D$	$B - E = \Delta PS$
Welfare, $W = CS + PS$	$A + B + C + D + E$	$A + B + D$	$-C - E = \Delta W = DWL$

If the firm acts like a monopoly and operates where its marginal revenue equals its marginal cost, only 6 units are sold at the monopoly price of \$18, and consumer surplus is only A. Part of the lost consumer surplus, B, goes to the monopoly, but the rest, C, is lost.

By charging the monopoly price of \$18 instead of the competitive price of \$16, the monopoly receives \$2 more per unit and earns an extra profit of area $B = \$12$ on the $Q_m = 6$ units it sells. The monopoly loses area E, however, because it sells less than the competitive output. Consequently, the monopoly's producer surplus increases by $B - E$ over the competitive level. We know that its producer surplus increases, $B - E > 0$, because the monopoly had the option of producing at the competitive level and chose not to do so.

Welfare under monopoly is lower than with competition. The deadweight loss of monopoly is $-C - E$, which is the consumer surplus and producer surplus that is lost because less than the competitive output is produced. As in the tax analysis of a competitive market in Chapter 9, the deadweight loss is due to the gap between price and marginal cost at the monopoly output. At $Q_m = 6$, the price, \$18, is above the marginal cost, \$12, so consumers are willing to pay more for the last unit of output than it costs to produce it.

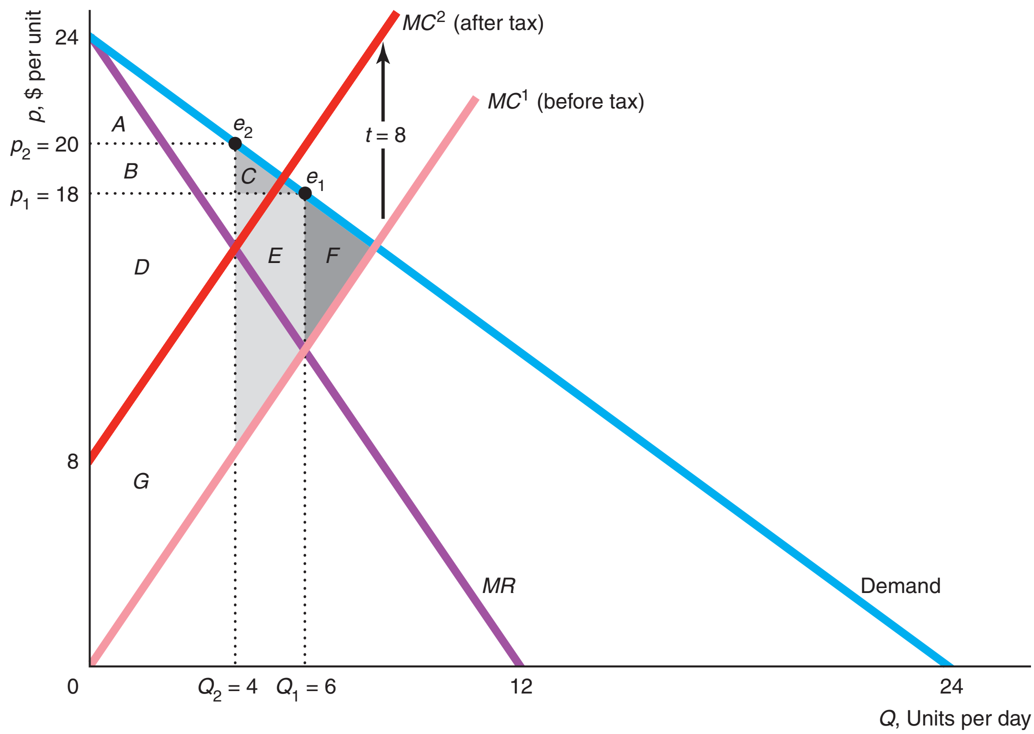
Solved Problem 11.4

MyLab Economics Solved Problem

In the linear example in panel a of Figure 11.3, how does charging the monopoly a specific tax of $t = \$8$ per unit affect the monopoly optimum and the welfare of consumers, the monopoly, and society (where society's welfare includes the tax revenue)? What is the incidence of the tax on consumers?

Answer

1. *Determine how imposing the tax affects the monopoly optimum.* In the figure, the intersection of the marginal revenue curve, MR, and the before-tax marginal cost curve, MC^1, determines the monopoly optimum quantity, $Q_1 = 6$. At the before-tax optimum, e_1, the price is $p_1 = \$18$. The specific tax causes the monopoly's before-tax marginal cost curve, $MC^1 = 2Q$, to shift upward by 8 to $MC^2 = MC^1 + 8 = 2Q + 8$. After the tax is applied, the monopoly operates where $MR = 24 - 2Q = 2Q + 8 = MC^2$. In the after-tax monopoly optimum, e_2, the quantity is $Q_2 = 4$ and the price is $p_2 = \$20$. Thus, output falls by $\Delta Q = 6 - 4 = 2$ units and the price increases by $\Delta p = \$20 - \$18 = \$2$.



	Monopoly Before Tax	Monopoly After Tax	Change
Consumer Surplus, CS	$A + B + C$	A	$-B - C = \Delta CS$
Producer Surplus, PS	$D + E + G$	$B + D$	$B - E - G = \Delta PS$
Tax Revenues, $T = tQ$	0	G	$G = \Delta T$
Welfare, $W = CS + PS + T$	$A + B + C + D + E + G$	$A + B + D + G$	$-C - E = \Delta W$
Deadweight Loss, DWL	$-F$	$-C - E - F$	$-C - E = DWL$

2. *Calculate the change in the various welfare measures*. The figure shows how the welfare measures change. Area G is the tax revenue collected by the government, $tQ = \$32$, because its height is the distance between the two marginal cost curves, $t = \$8$, and its width is the output the monopoly produces after the tax is imposed, $Q_2 = 4$. The tax reduces consumer and producer surplus and increases the deadweight loss. We know that producer surplus falls because (a) the monopoly could have produced this reduced output level in the absence of the tax but did not because it was not the profit-maximizing output, so its before-tax profit falls, and (b) the monopoly must now pay taxes. The before-tax deadweight loss from monopoly is $-F$. The after-tax deadweight loss is $-C - E - F$, so the increase in deadweight loss due to the tax is $-C - E$. The table below the graph shows that consumer surplus changes by $-B - C$ and that producer surplus changes by $B - E - G$.

3. *Calculate the incidence of the tax*. Because the tax goes from 0 to 8, the change in the tax is $\Delta t = 8$. The incidence of the tax (Chapter 3) on consumers is $\Delta p/\Delta t = 2/8 = \frac{1}{4}$. (The monopoly absorbs \$6 of the tax and passes on only \$2.)[17]

Comment: A tax increases the deadweight loss in a monopolized market.

11.4 Causes of Monopoly

Why are some markets monopolized? The two most important reasons involve costs and government actions.[18]

Cost-Based Monopoly

Certain cost structures may facilitate the creation of a monopoly. One possibility is that a firm may have substantially lower costs than potential rivals. A second possibility is that the firms in an industry have cost functions such that one firm can produce any given output at a lower cost than two or more firms can.

Cost Advantages If a low-cost firm profitably sells at a price so low that other potential competitors with higher costs would lose money, no other firms enter the market. Thus, the low-cost firm is a monopoly.

A firm can have a cost advantage over potential rivals for several reasons. It may have a superior technology or a better way of organizing production.[19] For example,

[17]In contrast to a competitive market, when a monopoly is taxed, the incidence of the tax on consumers can exceed 100%, as Appendix 11B demonstrates. "Welfare Effects of Ad Valorem Versus Specific Taxes" in MyLab Economics, Chapter 11, proves that a government raises more tax revenue with an ad valorem tax applied to a monopoly than with a specific tax when the tax rates are set so that the after-tax output is the same with either tax.

[18]In later chapters, we discuss three other means by which monopolies are created. One method is the merger of several firms into a single firm (Chapter 13). This merger creates a monopoly if new firms fail to enter the market. A second method is for firms to coordinate their activities and set their prices as a monopoly would (Chapter 13). Firms that act collectively in this way are called a *cartel*. A third method is for a monopoly to use strategies that discourage other firms from entering the market (Chapter 14).

[19]When a firm develops a better production method that provides it with a cost advantage, it is important for the firm to either keep the information secret or obtain a patent, so that the government protects it from having its innovation imitated. Thus, both secrecy and patents facilitate cost-based monopolies.

Henry Ford's methods of organizing production using assembly lines and standardization allowed him to produce cars at substantially lower cost than rival firms until they copied his organizational techniques.

Another example is that the firm controls an *essential facility*: a scarce resource that a rival needs to use to survive. A firm that owns the only quarry in a region is the only firm that can profitably sell gravel to local construction firms.

natural monopoly
one firm can produce the total output of the market at lower cost than several firms could

Natural Monopoly A market has a **natural monopoly** if one firm can produce the total output of the market at lower cost than several firms could. A firm can be a natural monopoly even if it does not have a cost advantage over rivals because average cost is lower if only one firm operates. Given that $C(q)$ is the cost for any firm to produce q, the condition for a natural monopoly is

$$C(Q) < C(q_1) + C(q_2) + \cdots + C(q_n), \tag{11.10}$$

where $Q = q_1 + q_2 + \ldots + q_n$ is the sum of the output of any $n \geq 2$ firms.

If a firm has economies of scale (Chapter 7) at all levels of output, its average cost curve falls as output increases for any observed level of output. If all potential firms have the same strictly declining average cost curve, this market is a natural monopoly, as we now illustrate.[20]

A company that supplies water to homes incurs a high fixed cost, F, to build a plant and connect houses to the plant. The firm's marginal cost, m, of supplying water is constant, so its marginal cost curve is horizontal and its average cost, $AC = m + F/Q$, declines as output rises. (An example is the iPad in Solved Problem 11.2.)

Figure 11.6 shows such marginal and average cost curves where $m = 10$ and $F = \$60$. If the market output is 12 units per day, one firm produces that output at an average cost of \$15, or a total cost of \$180 (= \$15 × 12). If two firms each produce 6 units, the average cost is \$20 and the cost of producing the market output is \$240 (= \$20 × 12), which is greater than the cost with a single firm.

If the two firms divided total production in any other way, their cost of production would still exceed the cost of a single firm (as Solved Problem 11.5 shows). The reason is that the marginal cost per unit is the same no matter how many firms produce, but each additional firm adds a fixed cost, which raises the cost of producing a given quantity. If only one firm provides water, the cost of building a second plant and a second set of pipes is avoided.

In an industry with a natural monopoly cost structure, having only one firm produce is the lowest cost way to produce any given output level.[21] Believing that they are natural monopolies, governments frequently grant monopoly rights to *public utilities* to provide essential goods or services such as water, gas, electric power, or mail delivery.

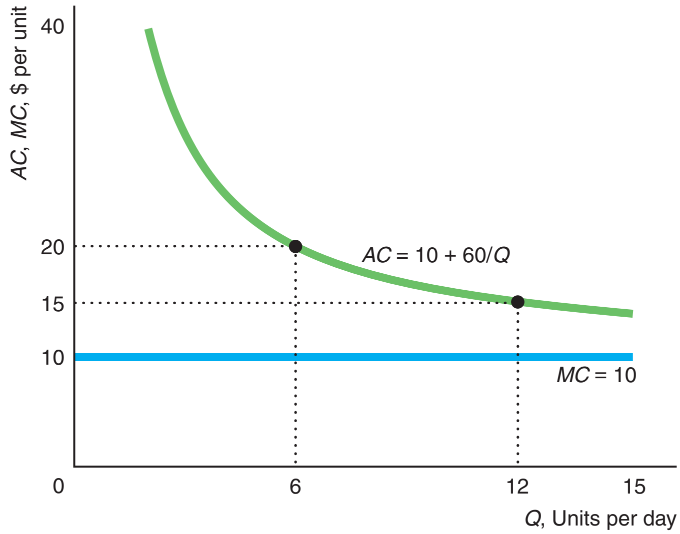
---

[20]A firm may be a natural monopoly even if its cost curve does not fall at all levels of output. If a U-shaped average cost curve reaches its minimum at 100 units of output, it may be less costly for only one firm to produce an output of slightly more than 100 units (such as 101 or 102) even though average cost is rising at that output. Thus, a cost function with economies of scale everywhere is a sufficient but not a necessary condition for a natural monopoly.

[21]However, society's welfare may be greater with more than one firm in the industry producing at higher cost, because competition drives down the price from the monopoly level. A solution that allows society to maximize welfare is to have only one firm produce, but the government regulates that firm to charge a price equal to marginal cost (as we discuss later in this chapter).

Figure 11.6 Natural Monopoly

This natural monopoly has a strictly declining average cost.



Solved Problem 11.5

A firm that delivers Q units of water to households has a total cost of $C(Q) = mQ + F$. If any entrant would have the same cost, does this market have a natural monopoly?

Answer

Determine whether costs rise if two firms produce a given quantity. Let q_1 be the output of Firm 1 and q_2 be the output of Firm 2. The combined cost of these two firms producing $Q = q_1 + q_2$ is

$$C(q_1) + C(q_2) = (mq_1 + F) + (mq_2 + F) = m(q_1 + q_2) + 2F = mQ + 2F.$$

If a single firm produces Q, its cost is $C(Q) = mQ + F$. Thus, the cost of producing any given Q is greater with two firms than with one firm (the condition in Equation 11.10), so this market has a natural monopoly.

Government Creation of a Monopoly

By preventing other firms from entering a market, governments create monopolies. We start by discussing common barriers to entry the governments erect and then concentrate on patents.

Barriers to Entry Sometimes governments own and manage monopolies, forbidding other firms from entering. In the United States, as in most countries, the postal service is a government monopoly. Several countries, such as China, maintain a tobacco monopoly. Many local governments own and operate public utility monopolies that provide garbage collection, electricity, water, gas, phone services, and other utilities.

Governments around the world have privatized many state-owned monopolies in the past several decades. By selling its monopolies to private firms, a government can capture the value of future monopoly earnings today. However, for political or other reasons, governments frequently sell at a lower price that does not capture all future profits.

In other markets, governments give or auction monopoly rights—a license to operate—to a private firm. For example, many cities let a single, private firm provide cable television services. Many government-owned airports auction off the rights for a single firm to provide a particular service, such as selling luggage. By auctioning a monopoly to a private firm, a government can capture the future value of monopoly earnings.[22]

patent
an exclusive right granted to the inventor to sell a new and useful product, process, substance, or design for a fixed period

Patents If a firm cannot prevent imitation by keeping its discovery secret, it may obtain government protection to prevent other firms from duplicating its discovery and entering the market. Virtually all countries provide such protection through a **patent**: an exclusive right granted to the inventor to sell a new and useful product, process, substance, or design for a fixed period. The lifetime of a patent varies across countries, although it is now 20 years in the United States and in most other countries.

This right allows the patent holder to be the exclusive seller or user of the new invention.[23] Patents often give rise to monopoly, but not always. For example, although a patent may grant a firm the exclusive right to use a particular process in producing a product, other firms may be able to produce the same product using different processes.

A firm with a patent monopoly sets a high price that results in deadweight loss. Why, then, do governments grant patent monopolies? The main reason is to encourage inventive activity—less innovation would occur if successful inventors did not receive a patent monopoly. The costs of developing a new drug or new computer chip are often hundreds of millions or even billions of dollars. If anyone could copy a new drug or chip and compete with the inventor, few individuals or firms would undertake costly research. Thus, the government is explicitly trading off the long-run benefits of additional inventions against the shorter-term harms of monopoly pricing during the period of patent protection.

An alternative to using patents to spur research is for the government to provide research grants or offer prizes for important inventions. However, doing so is costly, so most governments rely primarily on patents.

Application

Botox Patent Monopoly

Ophthalmologist Dr. Alan Scott turned the deadly poison botulinum toxin into a miracle drug to treat two eye conditions: strabismus, a condition in which the eyes are not properly aligned, and blepharospasm, an uncontrollable closure of the eyes. Strabismus affects about 4% of children and blepharospasm left about 25,000 Americans functionally blind before Scott's discovery. His patented drug, Botox, is sold by Allergan, Inc.

Dr. Scott has been amused to see several of the unintended beneficiaries of his research at the annual Academy Awards. Even before it was explicitly approved for cosmetic use, many doctors were injecting Botox into the facial muscles of actors, models, and others to smooth out their wrinkles. (The drug paralyzes the muscles,

[22] Alternatively, a government could auction the rights to the firm that offers to charge the lowest price, so as to maximize welfare. Oakland, California, tried to do that once with cable television service.

[23] Owners of patents may sell or license the right to other firms to use a patented process or produce a patented product.

so those injected with it also lose their ability to frown or smile—and, some would say, act.) The treatment is only temporary, lasting up to 120 days, so repeated injections are necessary.

Allergan has a near-monopoly in the treatment of wrinkles, although plastic surgery and collagen, Restylane, hyaluronic acids, and other filler injections provide limited competition. However, 54% of Botox sales are now for other uses, including as a treatment for chronic migraine and overactive bladder.

Allergan had Botox sales of \$800 million in 2004 and about \$2.8 billion in 2015. Indeed, Botox's value may increase. As Allergan finds new uses for Botox, its sales continue to increase. According to one forecast, Botox's global sales will reach \$3 billion by the end of 2018.

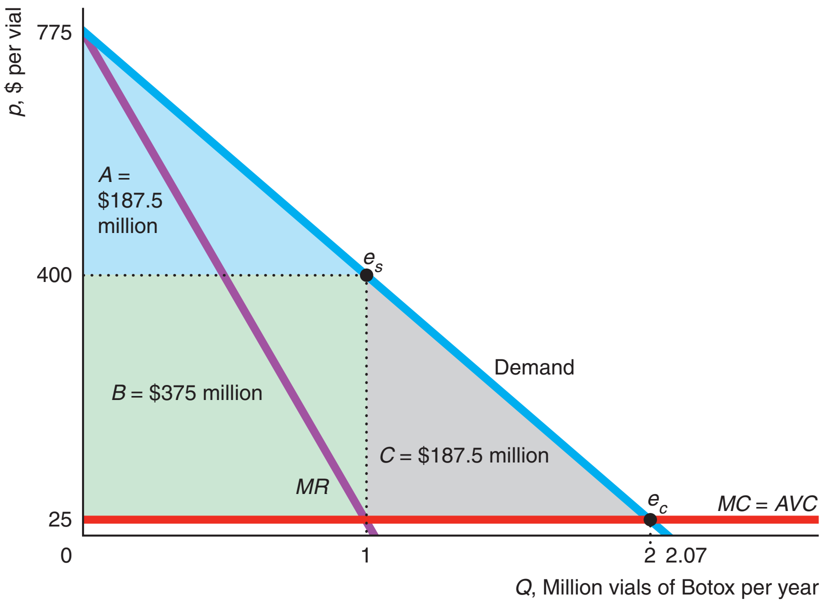
Dr. Scott can produce a vial of Botox in his lab for about \$25. Allergan sells the potion to doctors for about \$400. Assuming that the firm is setting its price to maximize its short-run profit, we can rearrange Equation 11.9 to determine the elasticity of demand for Botox:

$$= -\frac{p}{p - MC} = -\frac{400}{400 - 25} \approx -1.067.$$

Thus, the demand that Allergan faces is only slightly elastic: A 1% increase in price causes quantity to fall by slightly more than 1%.

If we assume that the demand curve is linear and given that the elasticity of demand is −1.067 at the monopoly optimum, e_m (1 million vials sold at \$400 each, producing revenue of \$400 million in 2002), then Allergan's inverse demand function was[24]

$$p = 775 - 375Q.$$



[24]The graph shows an inverse linear demand curve: $p = a - bQ$. The elasticity of demand for such a linear demand curve is $= -(1/b)(p/Q)$. Using Equation 11.9 and the data for the Application, the elasticity of demand at the optimum is $-400/375 = -(1/b)(400/1)$, where Q is measured in millions of vials. Thus, $b = 375$. Solving $p = 400 = a - (375 \times 1)$, we find that $a = 775$.

This demand curve (see the graph) has a slope of −375 and hits the price axis at \$775 and the quantity axis at about 2.07 million vials per year. The corresponding marginal revenue curve,

$$MR = 775 - 750Q,$$

strikes the price axis at \$775 and has twice the slope, −750, of the demand curve.

The intersection of the marginal revenue and marginal cost curves,

$$MR = 775 - 750Q = 25 = MC,$$

determines the monopoly optimum at the profit-maximizing quantity of 1 million vials per year and at a price of \$400 per vial.

Were the company to sell Botox at a price equal to its marginal cost of \$25 (as a competitive industry would), consumer surplus would equal area $A + B + C$ The height of triangle $A + B + C$ is $\$750 = \$775 - \$25$, and its length is 2 million vials, so its area is \$750 ($= \frac{1}{2} \times 750 \times 2$) million. At the higher monopoly price of \$400, the consumer surplus is $A = \$187.5$ million. Compared to the competitive solution, e_c, buyers lose consumer surplus of $B + C = \$562.5$ million per year. Part of this loss, $B = \$375$ million per year, is transferred from consumers to Allergan. The rest, $C = \$187.5$ million per year, is the deadweight loss from monopoly pricing. Allergan's profit is its producer surplus, B, minus its fixed costs.

11.5 Government Actions That Reduce Market Power

Some governments act to reduce or eliminate monopolies' market power. Many governments directly regulate monopolies, especially those created by the government, such as public utilities. Most Western countries have designed laws to prevent a firm from driving other firms out of the market so as to monopolize it. A government may destroy a monopoly by breaking it up into smaller, independent firms (as the government did with Alcoa, the former aluminum monopoly).

Regulating Monopolies

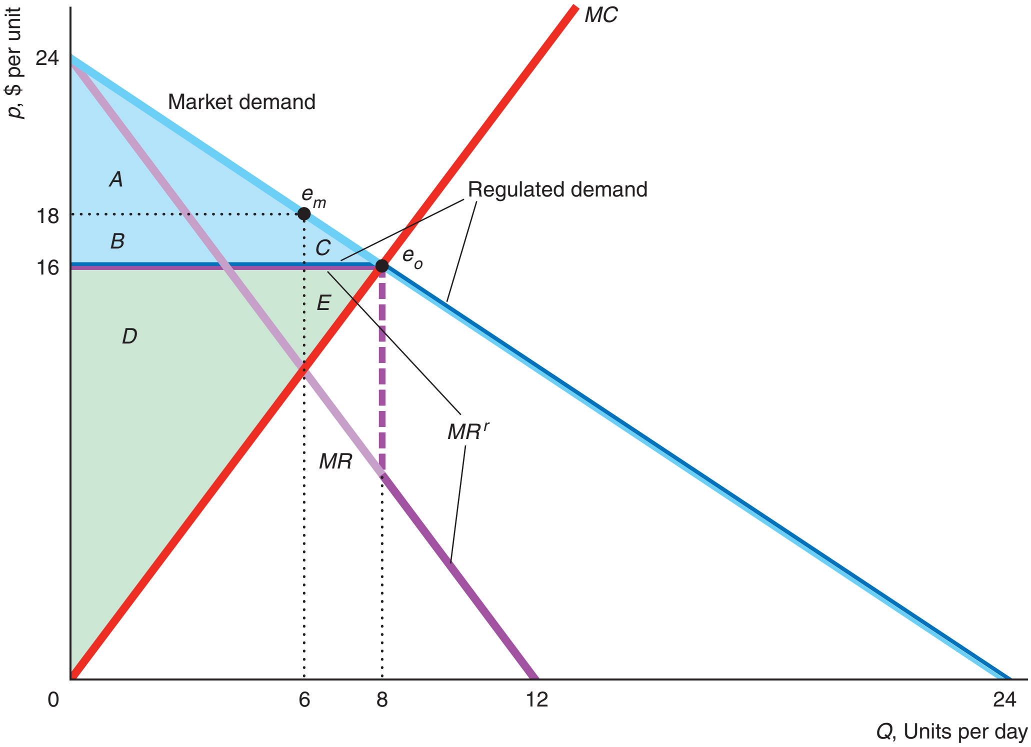
Governments limit monopolies' market power in various ways. For example, most utilities are subject to direct regulation. Today, the most commonly used approach to regulating monopoly pricing is to impose a price ceiling, called a *price cap*. Price cap regulation is used for telecommunications monopolies in 33 U.S. states and in many countries, including Australia, Canada, Denmark, France, Germany, Mexico, Sweden, and the United Kingdom (Sappington and Weisman, 2010).

Optimal Price Regulation A government can eliminate the deadweight loss of monopoly by imposing a price cap equal to the price that would prevail in a competitive market. We use our earlier linear example to illustrate this type of regulation in Figure 11.7.

If the government doesn't regulate the profit-maximizing monopoly, the monopoly optimum is e_m, at which 6 units are sold at the monopoly price of \$18. Suppose that the government sets a ceiling price of \$16, the price at which the marginal cost curve intersects the market demand curve. Because the monopoly cannot charge more than \$16 per unit, the monopoly's regulated demand curve is horizontal at \$16 (up to 8 units) and is the same as the market demand curve at lower prices.

Figure 11.7 Optimal Price Regulation MyLab Economics Video

If the government sets a price ceiling at 16, where the monopoly's marginal cost curve hits the demand curve, the new demand curve that the monopoly faces has a kink at 8 units, and the corresponding marginal revenue curve, MR^r, "jumps" at that quantity. The regulated monopoly sets its output where $MR^r = MC$, selling the same quantity, 8 units, at the same price, 16, as a competitive industry would. The regulation eliminates the monopoly deadweight loss, $C + E$. Consumer surplus, $A + B + C$, and producer surplus, $D + E$, are the same as under competition.



	Monopoly Without Regulation	Monopoly with Optimal Regulation	Change
Consumer Surplus, CS	A	$A + B + C$	$B + C = \Delta CS$
Producer Surplus, PS	$B + D$	$D + E$	$E - B = \Delta PS$
Welfare, $W = CS + PS$	$A + B + D$	$A + B + C + D + E$	$C + E = \Delta W$
Deadweight Loss, DWL	$-C - E$	0	$C + E = \Delta DWL$

The marginal revenue curve MR^r, which corresponds to the regulated demand curve, is horizontal where the regulated demand curve is horizontal (up to 8 units) and equals the marginal revenue curve MR, which corresponds to the market demand curve at larger quantities.

The regulated monopoly sets its output at 8 units, where MR^r equals its marginal cost, MC, and charges the maximum permitted price, \$16. The regulated firm still makes a profit, because its average cost is less than \$16 at 8 units. The optimally regulated monopoly optimum, e_o, is the same as the competitive equilibrium, where marginal cost (supply) equals the market demand curve.[25] Thus, setting a price ceiling where the *MC* curve and market demand curve intersect eliminates the deadweight loss of monopoly.

How do we know that this regulation is optimal? The answer is that this regulated outcome is the same as would occur if this market were competitive, where welfare is maximized (Chapter 9). As the table accompanying Figure 11.7 shows, the deadweight loss of monopoly, $C + E$, is eliminated by this optimal regulation.

Problems in Regulating Governments often fail to regulate monopolies optimally for at least three reasons. First, due to limited information about the demand and marginal cost curves, governments may set a price ceiling above or below the competitive level.

Second, regulation may be ineffective when regulators are *captured*: influenced by the firms they regulate. Typically, this influence is more subtle than an outright bribe. Many American regulators worked in the industry before they became regulators and hence are sympathetic to those firms. For many other regulators, the reverse is true: They aspire to obtain good jobs in the industry eventually, so they do not want to offend potential employers. And some regulators, relying on industry experts for their information, may be misled or at least heavily influenced by the industry. For example, the California Public Utilities Commission urged telephone and cable companies to negotiate among themselves about how they wanted to open local phone markets to competition. Arguing that these influences are inherent, some economists contend that price and other types of regulation are unlikely to result in efficiency.

Third, because regulators generally cannot subsidize the monopoly, they may be unable to set the price as low as they want because the firm may shut down. In a natural monopoly where the average cost curve is strictly above the marginal cost curve, if the regulator sets the price equal to the marginal cost so as to eliminate deadweight loss, the firm cannot afford to operate. If the regulators cannot subsidize the firm, they must raise the price to a level where the firm at least breaks even.

Nonoptimal Price Regulation If the government sets the price ceiling at a nonoptimal level, a deadweight loss results. Suppose that the government sets the regulated price below the optimal level (such as 16 in Figure 11.7). If it sets the price below the firm's minimum average cost, the firm shuts down, so the deadweight loss equals the sum of the consumer surplus and the producer surplus under optimal regulation, $A + B + C + D + E$.

Many consumers want the government to set regulated monopoly price as low as possible:

> **Common Confusion:** Consumers benefit the lower the government sets the regulated price that a monopoly may charge (without causing the firm to shut down).

A very low regulated price may help some consumers, but hurt others. If the government sets the price ceiling below the optimally regulated price but high enough that the firm does not shut down, consumers who are lucky enough to buy the good benefit because they can buy it at a lower price than they could with optimal regulation. As we show in Solved Problem 11.6, society suffers a deadweight loss because less output is sold than with optimal regulation.

---

[25]The monopoly produces at e_o only if the regulated price is greater than its average variable cost. Here the regulated price, \$16, exceeds the average variable cost at 8 units of \$8. Indeed, the firm makes a profit because the average cost at 8 units is \$9.50.

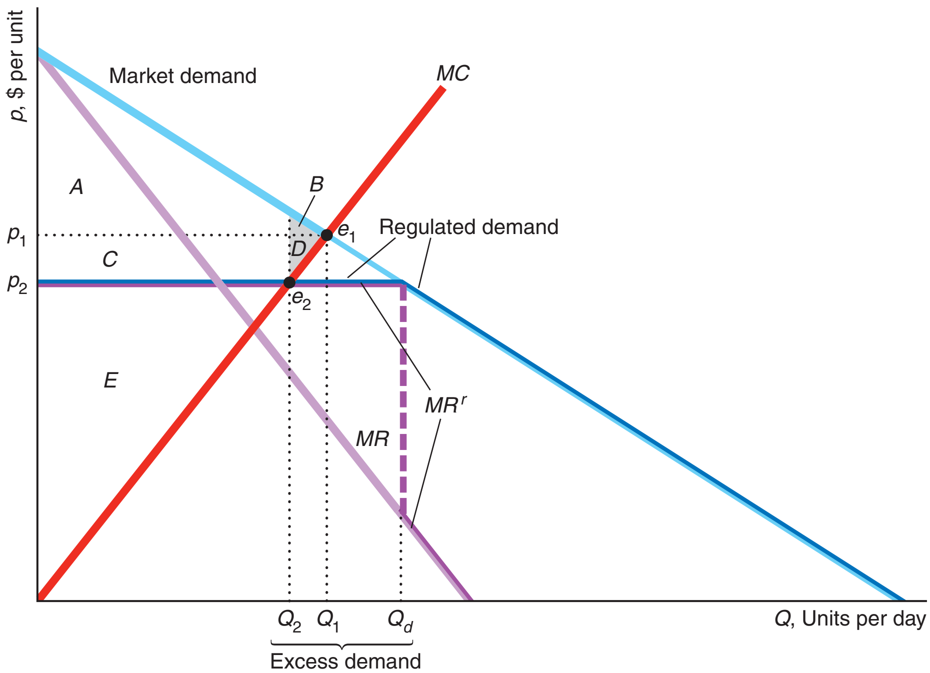
Solved Problem 11.6

MyLab Economics Solved Problem

Suppose that the government sets a price, p_2, that is below the socially optimal level, p_1, but above the monopoly's minimum average cost. How do the price, the quantity sold, the quantity demanded, and welfare under this regulation compare to those under optimal regulation?

Answer

1. *Describe the optimally regulated outcome.* With optimal regulation, e_1, the price is set at p_1, where the market demand curve intersects the monopoly's marginal cost curve on the accompanying graph. The optimally regulated monopoly sells Q_1 units.



	Monopoly with Optimal Regulation	Monopoly with a Low Regulated Price	Change
Consumer Surplus, CS	$A + B$	$A + C$	$C - B = \Delta CS$
Producer Surplus, PS	$C + D + E$	E	$-C - D = \Delta PS$
Welfare, $W = CS + PS$	$A + B + C + D + E$	$A + C + E$	$-B - D = \Delta W = DWL$

2. *Describe the outcome when the government regulates the price at p_2.* Where the market demand is above p_2, the regulated demand curve for the monopoly is horizontal at p_2(up to Q_d). The corresponding marginal revenue curve, MR^r, is kinked. It is horizontal where the regulated demand curve is horizontal. The MR^r curve is the same as the marginal revenue curve corresponding to the market demand curve, MR, where the regulated demand curve is downward sloping. The monopoly maximizes its profit by selling Q_2 units at p_2. The new regulated monopoly optimum is e_2, where MR^r intersects MC. The firm does not shut down when regulated as long as its average variable cost at Q_2 is less than p_2.

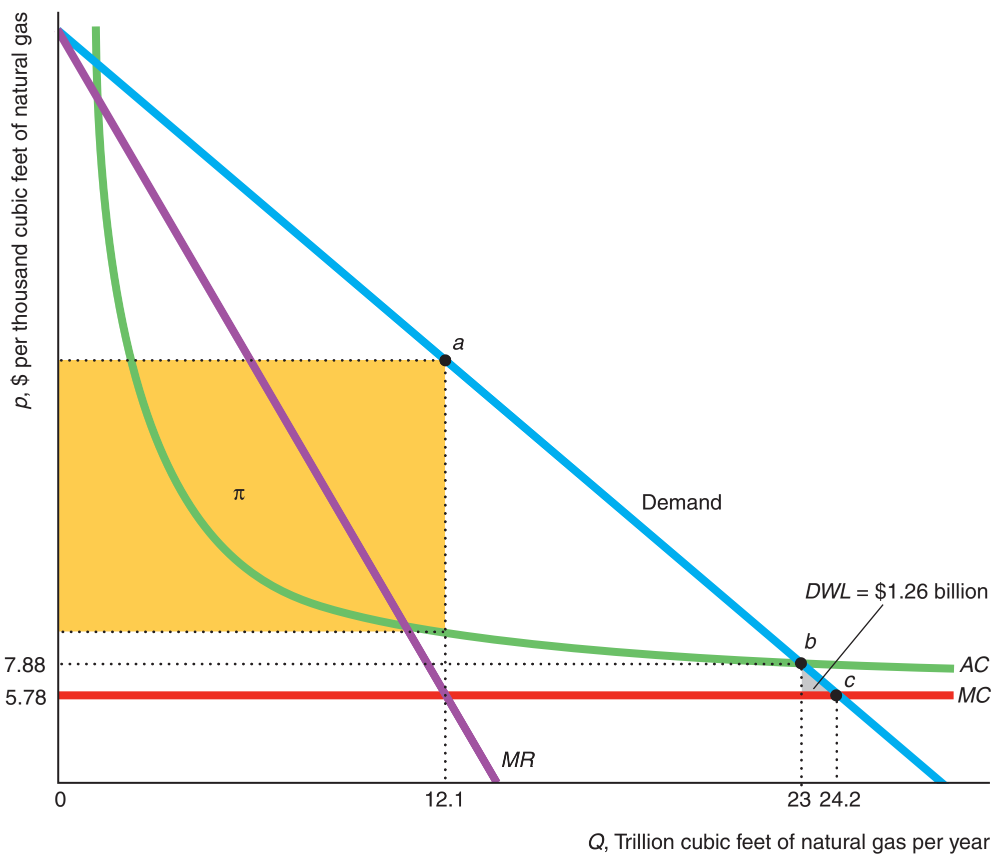
3. *Compare the outcomes*. The quantity that the monopoly sells falls from Q_1 to Q_2 when the government lowers its price ceiling from p_1 to p_2. At that low price consumers want to buy Q_d, resulting in excess demand of $Q_d - Q_2$. Compared to optimal regulation, welfare is lower by at least $B + D$.

Comment: The welfare loss is greater if unlucky consumers waste time trying to buy the good unsuccessfully or if goods are not allocated optimally among consumers. A consumer who values the good at only p_2 may be lucky enough to buy it, while a consumer who values the good at p_1 or more may not be able to obtain it (Chapter 9).

Application

Natural Gas Regulation

Because U.S. natural gas monopolies usually have significant economies of scale and regulators generally cannot subsidize them, the regulated price is set above marginal cost, creating a deadweight loss. The figure is based on the estimates of Davis and Muehlegger (2010).[26] If unregulated, this monopoly would sell 12.1 trillion cubic feet of natural gas per year, which is determined by the intersection of its marginal revenue and marginal cost curves. It would charge the corresponding price on the demand curve at point *a*. Its profit is the rectangle labeled π, with a length equal to the quantity, 12.1 trillion cubic feet, and a height equal to the difference between the price at *a* and the corresponding average cost.



[26]We use their most conservative estimate: the one that produces the smallest deadweight loss. We approximate their demand curve with a linear one that has the same price elasticity of demand of −0.2 at point *b*. This figure represents the aggregation of state-level monopolies to the national level.

To eliminate deadweight loss, the government should set the price ceiling equal to the marginal cost of $5.78 per thousand cubic feet of natural gas so that the monopoly behaves like a price taker. The price ceiling or marginal cost curve hits the demand curve at *c*, where the quantity is 24.2 trillion cubic feet per year—double the unregulated quantity. At that quantity, the regulated utility would lose money. The average cost at that quantity is $7.78, which is 10¢ less than the average cost of $7.88 at a quantity of 23 trillion cubic feet. The regulated price, $5.78, is less than the average cost of $7.78 at that quantity, so it would lose $2 on each thousand cubic feet it sells, or $48.4 billion in total. Thus, it would be willing to sell this quantity at this price only if the government subsidizes it.

Typically, it is politically infeasible for a government regulatory agency to subsidize a monopoly. On average, the natural gas regulatory agencies set the price at $7.88 per thousand cubic feet, where the demand curve intersects the average cost curve and the monopoly breaks even, point *b*. The monopoly sells 23 trillion cubic feet per year. The corresponding price, $7.88, is 36% above marginal cost, $5.78. Consequently, society incurs a deadweight loss of $1.26 billion annually, which is the gray triangle in the figure. This deadweight loss is much smaller than it would be if the monopoly were unregulated.

Increasing Competition

Encouraging competition is an alternative to regulation as a means of reducing the harms of monopoly. When a government has created a monopoly by preventing entry, it can quickly reduce the monopoly's market power by allowing other firms to enter. As new firms enter the market, the former monopoly must lower its price to compete, so welfare rises. Many governments are actively encouraging entry into telephone, electricity, and other utility markets that were formerly monopolized.

Similarly, a government may end a ban on imports so that a domestic monopoly faces competition from foreign firms. If the market includes many foreign firms with the same costs as the domestic firm, the former monopoly becomes just one of many competitive firms. As the market becomes competitive, consumers pay the competitive price, and the deadweight loss of monopoly is eliminated.

Governments around the world are increasing competition in formerly monopolized markets. For example, many U.S. and European governments are forcing former telephone and energy monopolies to compete.

Similarly, under pressure from the World Trade Organization, many countries are reducing or eliminating barriers that protected domestic monopolies. The entry of foreign competitive firms into a market can create a new, more competitive market structure.

Application

Movie Studios Attacked by 3D Printers!

Disney, Marvel, Lucas, and a variety of other companies make a fortune from selling figurines and other plastic toys based on their movies, such as Disney's *Frozen* and *Finding Dory*. These firms hold monopoly rights to produce their toys under copyright laws, which give the creators of original works such as comics and movies the exclusive right to its use and distribution for a limited time.

However, 3D printers are undermining their monopolies. Fans upload high-quality designs to the Web, which anyone with a 3D printer can use to produce pirated versions of these toys. The online marketplace for 3D designs and objects include comic-book heroes, cartoon characters, Angelina Jolie's headdress in the movie *Maleficent*, Homer Simpson, and even Walt Disney's head.

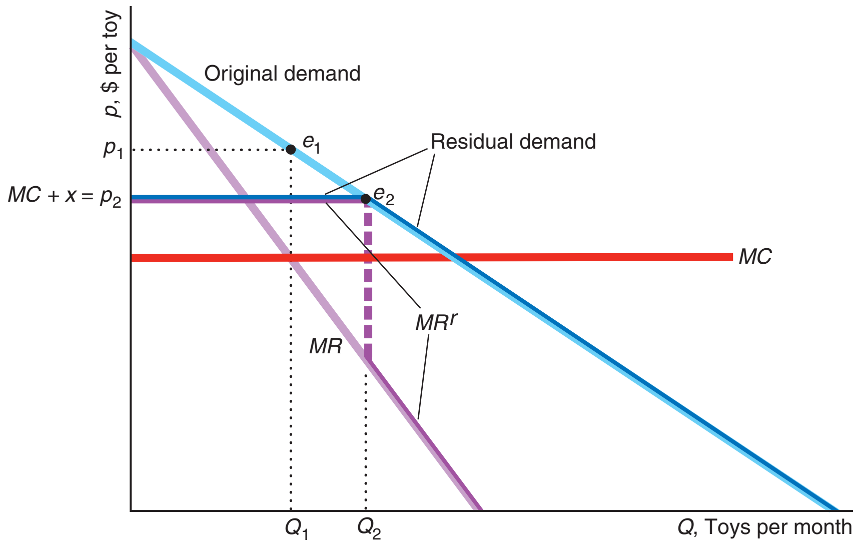
Thus, a movie firm is no longer a monopoly. It is a *dominant firm* that faces a *competitive fringe* made up of small, price-taking firms. If the movie firms can't block such pirating using the legal system, they may have to drop the prices for their toys to compete. The competitive fringe limits the movie firm's price much as government regulation would. The movie firm's main advantage is that its mass-production marginal costs are lower than the 3D-printer marginal costs of hobbyists and pirates, but 3D-printing costs are falling.

Solved Problem 11.7

How does the presence of pirated, 3D toys affect the price that Disney charges for a *Frozen* or *Inside Out* toy? Assume that Disney has a constant marginal cost MC. It faces a large number of identical, higher-cost rivals—the competitive fringe—which acts like (competitive) price takers so that their collective supply curve is horizontal at $p_2 = MC + x$.

Answer

1. *Show how Disney prices a toy figurine if it is a monopoly by equating its marginal revenue and marginal cost.* The figure shows Disney's original (market) demand curve for its toy as a light-blue line. The light-purple line is the corresponding marginal revenue curve. Its profit-maximizing outcome was e_1 when Disney set its quantity, Q_1, where its MR curve hit its MC curve, and the corresponding price was p_1.



2. *Show how the competitive supply curve alters the demand curve facing Disney.* The competitive supply curve acts like a government price ceiling. Now, Disney cannot charge more than $p_2 = MC + x$. Thus, its dark-blue residual demand curve is flat at $MC + x$ and the same as the original downward-sloping demand curve at lower prices. (That is, the residual demand curve for the toy is similar to that of the regulated monopoly in Figure 11.7.)

3. *Determine Disney's new optimal outcome by equating its new marginal revenue with its marginal cost.* Disney acts like a monopoly with respect to its residual demand curve (rather than to its original demand curve). Corresponding to Disney's residual demand curve in the figure is a dark-purple, kinked marginal revenue curve, MR^r, that crosses Disney's marginal cost line at Q_2.[27] Disney maximizes its profit by selling Q_2 units for p_2 at e_2. That is, Disney sells more toys at a lower price than before the other firms entered the market. Once Disney lowers its price, the fringe sells virtually nothing.

11.6 Networks, Dynamics, and Behavioral Economics

We have examined how a monopoly behaves in the current period, ignoring the future. For many markets, such an analysis is appropriate. However, in some markets, decisions today affect demand or cost in a future period, creating a need for a *dynamic* analysis, in which firms explicitly consider relationships between periods.

In such markets, the monopoly may maximize its long-run profit by making a decision today that does not maximize its short-run profit. For example, frequently a firm introduces a new product—such as a new type of candy bar—by initially charging a low price or giving away free samples to generate word-of-mouth publicity or to let customers learn about its quality in hopes of getting their future business. We now consider an important reason why consumers' demand in the future may depend on a monopoly's actions in the present.

Network Externalities

network externality
one person's demand for a good depends on the consumption of the good by others

The number of customers a firm has today may affect the demand curve it faces in the future. A good has a **network externality** if one person's demand depends on the consumption of a good by others.[28] If a good has a *positive* network externality, its value to a consumer grows as the number of units sold increases.

When a firm introduces a new good with a network externality, it faces a chicken-and-egg problem: It can't get Max to buy the good unless Sofia will buy it, but it can't get Sofia to buy it unless Max will. The firm wants its customers to coordinate or to make their purchase decisions simultaneously.

The telephone provides a classic example of a positive network externality. When the phone was introduced, potential adopters had no reason to get phone service unless their family and friends did. Why buy a phone if there's no one to call? For Bell's phone network to succeed, it had to achieve a *critical mass* of users—enough adopters that others wanted to join. Had it failed to achieve this critical mass, demand would have withered and the network would have died. Similarly, the market for fax machines grew very slowly until a critical mass was achieved where many firms had them.

---

[27] If MC crossed MR^r in the downward-sloping section, Disney would be a monopoly because its monopoly price would be less than $MC + x$.

[28] In Chapter 18, we discuss the more general case of an *externality*, which occurs when a person's well-being or a firm's production capability is directly affected by the actions of other consumers or firms rather than indirectly through changes in prices. The following discussion on network externalities is based on Leibenstein (1950), Rohlfs (1974), Katz and Shapiro (1994), Economides (1996), Shapiro and Varian (1999), and Rohlfs (2001).

Direct Size Effect Many industries exhibit positive network externalities where the customer gets a *direct* benefit from a larger network. The larger an automated teller machine (ATM) network such as the Plus network, the greater the odds that you will find an ATM when you want one, so the more likely it is that you will want to use that network. The more people who use a particular computer program, the more attractive it is to someone who wants to exchange files with other users.

Indirect Effects In some markets, positive network externalities are indirect and stem from complementary goods that are offered when a product has a critical mass of users. The more applications (apps) available for a smart phone, the more people want to buy that smart phone. However, many of these extra apps will be written only if a critical mass of customers buys the smart phone. Similarly, the more people who drive electric cars, the more likely it is that firms will provide charging stations; and the more charging stations, the more likely it is that someone will want to drive an electric car. Likewise, once a critical mass of customers had broadband Internet access, more services provided downloadable music and movies and high-definition Web pages. Once those popular apps appeared, more people signed up for broadband service.

Network Externalities and Behavioral Economics

The direct effect of network externalities depends on the size of the network, because customers want to interact with each other. However, sometimes consumers' behavior depends on beliefs or tastes that can be explained by psychological and sociological theories, which economists study in *behavioral economics* (Chapter 4).

One such explanation for a direct network externality effect is based on consumer attitudes toward other consumers. Harvey Leibenstein (1950) suggested that consumers sometimes want a good because "everyone else has it." A fad or other popularity-based explanation for a positive network externality is called a **bandwagon effect**: A person places greater value on a good as more and more other people possess it.[29] The success of the iPad today may be partially due to its early popularity. Ugg boots seem to be another example of a bandwagon effect.

bandwagon effect
a person places greater value on a good as more and more other people possess it

snob effect
a person places greater value on a good as fewer and fewer other people possess it

The opposite, a negative network externality is called a **snob effect**: A person places greater value on a good as fewer and fewer other people possess it. Some people prefer an original painting by an unknown artist to a lithograph by a star because no one else can possess that painting. (As Yogi Berra said, "Nobody goes there anymore; it's too crowded.")

Network Externalities as an Explanation for Monopolies

Because of the need for a critical mass of customers in a market with a positive network externality, we frequently see only one large firm surviving. Visa's ad campaign tells consumers that Visa cards are accepted "everywhere you want to be," including places that "don't take American Express." One could view its ad campaign as an attempt to convince consumers that its card has a critical mass and therefore that everyone should switch to it.

The Windows operating system largely dominates the market—not because it is technically superior to Apple's operating system or Linux—but because it has a critical mass of users. Consequently, a developer can earn more producing software that works with Windows than with other operating systems, and the larger number

[29] *Jargon alert*: Some economists use *bandwagon effect* to refer to any positive network externality—not just those that are based on popularity.

of software programs makes Windows increasingly attractive to users. Similarly, Engström and Forsell (2015) found that a 10 percentile increase in the displayed number of downloads of Android apps on Google Play increases downloads by about 20%.

But having obtained a monopoly, a firm does not necessarily keep it. History is filled with examples where one product knocks off another: "The king is dead; long live the king." Google replaced Yahoo! as the dominant search engine. Explorer displaced Netscape as the big-dog browser, and then was replaced by Chrome. Levi Strauss is no longer the fashion leader among the jeans set.

Application

eBay's Critical Mass

Many people argue that natural monopolies emerge after brief periods of Internet competition. A typical Web business requires a large up-front fixed cost—primarily for development and promotion—but has a relatively low marginal cost. Thus, Internet start-ups typically have downward sloping average cost per user curves. Which of the actual or potential firms with decreasing average costs will dominate and become a natural monopoly?[30]

In the early years of the Internet, eBay's online auction site, which started in 1995, faced competition from a variety of other Internet sites including Yahoo! Auctions that the then mighty Yahoo! created in 1998. At the time, many commentators correctly predicted that whichever auction site first achieved a critical mass of users would drive the other sites out of business. Indeed, most of these alternative sites died or faded into obscurity. For example, Yahoo! Auctions closed its U.S. and Canada sections of the site in 2007 (although its Hong Kong, Taiwanese, and Japanese sites continue to operate).

Apparently the convenience of having one site where virtually all buyers and sellers congregate—which lowers buyers' search cost—and that creates valuable reputations by having a feedback system (Brown and Morgan, 2006), more than compensates sellers for the lack of competition in sellers' fees. Brown and Morgan (2009) found that, prior to the demise of the Yahoo! Auction site, the same type of items attracted an average of two additional bidders on eBay and, consequently, the prices on eBay were consistently 20% to 70% percent higher than Yahoo! Auction prices.

Today, we see a battle to gain a critical mass between ridesharing companies, such as Uber and Lyft. This competition drove Sidecar from the market in 2016. Vacation rental sites such as Airbnb, VRBO, and Tripping are waging a similar battle.

A Two-Period Monopoly Model

A monopoly may be able to solve the chicken-and-egg problem of getting a critical mass for its product by initially selling the product at a low introductory price. By doing so, the firm maximizes its long-run profit but not its short-run profit.

Suppose that a monopoly sells its good—say, root-beer-scented jeans—for only two periods (after that, the demand goes to zero as a new craze hits the market). If the monopoly sells less than a critical quantity of output, Q, in the first period,

[30]If Internet sites provide differentiated products (see Chapter 14), then several sites may coexist even though average costs are strictly decreasing. In 2007, commentators were predicting the emergence of natural monopolies in social networks such as MySpace. However, whether a single social network can dominate for long is debatable as the sites have different attributes. MySpace lost dominance to Facebook. In turn, Facebook may be losing ground to similar sites such as Google+; newer models, such as Twitter, Flickr, and Instagram; or to sites that cater to specialized audiences such as Classmates and LinkedIn.

its second-period demand curve lies close to the price axis. However, if the good is a success in the first period—at least Q units are sold—the second-period demand curve shifts substantially to the right.

If the monopoly maximizes its short-run profit in the first period, it charges p^* and sells Q^* units, which is fewer than Q. To sell Q units, it would have to lower its first-period price to $\underline{p} < p^*$, which would reduce its first-period profit from π^* to $\underline{\pi}$.

In the second period, the monopoly maximizes its profit given its second-period demand curve. If the monopoly sold only Q^* units in the first period, it earns a relatively low second-period profit of π_l. However, if it sells Q units in the first period, it makes a relatively high second-period profit, π_h.

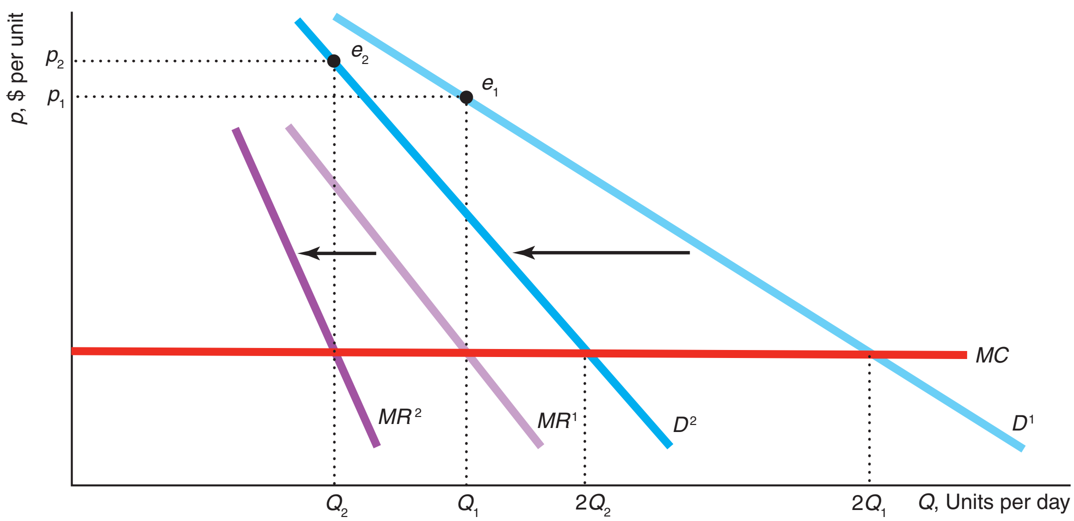
Should the monopoly charge a low introductory price in the first period? Its objective is to maximize its long-run profit: the sum of its profit in the two periods.[31] If the firm has a critical mass in the second period, its extra profit is $\pi_h - \pi_l$. To obtain this critical mass by charging a low introductory price in the first period, it lowers its first period profit by $\pi^* - \underline{\pi}$. Thus, the firm chooses to charge a low introductory period in the first period if its first period loss is less than its extra profit in the second period. This policy must be profitable for some firms: A 2013 Google search for "introductory price" found 13.2 million Web pages.

Challenge Solution

Brand-Name and Generic Drugs

When generic drugs enter the market after the patent on a brand-name drug expires, the demand curve facing the brand-name firm shifts to the left. Why do many brand-name drug companies raise their prices after generic rivals enter the market? The reason is that the demand curve not only shifts to the left but it rotates so that it is less elastic at the original price.

The price the brand-name firm sets depends on the elasticity of demand. When the firm has a patent monopoly, it faces the linear demand curve D^1 in the figure. Its monopoly optimum, e_1, is determined by the intersection of the corresponding marginal revenue curve MR^1 and the marginal cost curve. (Because it is twice as steeply sloped as the demand curve, MR^1 intersects the MC curve at Q_1, while the demand curve D^1 intersects the MC curve at $2Q_1$.) The monopoly sells the Q_1 units at a price of p_1.



[31]In Chapter 16, we discuss why firms place lower value on profit in the future than profit today. However, for simplicity in this analysis, we assume that the monopoly places equal value on profit in either period.

After the generic drugs enter the market, the linear demand curve facing the original patent holder shifts left to D^2 and becomes steeper and less elastic at the original price. The firm now maximizes its profit at e_2, where the quantity, Q_2, is smaller than Q_1 because D^2 lies to the left of D^1. However, the new price, p_2, is higher than the initial price, p_1, because the D^2 demand curve is less elastic at the new optimum quantity Q_2 than is the D^1 curve at Q_1.

Why might the demand curve rotate and become less elastic at the initial price? One explanation is that the brand-name firm has two types of consumers with different elasticities of demand who differ in their willingness to switch to a generic. One group of consumers is relatively price-sensitive and switches to the lower-priced generics. However, the brand-name drug remains the monopoly supplier to the remaining brand-loyal customers whose demand is less elastic than that of the price-sensitive consumers. These loyal customers prefer the brand-name drug because they are more comfortable with a familiar product, worry that new products may be substandard, or fear that differences in the inactive ingredients might affect them.

Older customers are less likely to switch brands than younger people. A survey by the American Association of Retired Persons found that people aged 65 and older were 15% less likely than people aged 45 to 64 to request generic versions of a drug from their doctor or pharmacist. Similarly, patients with generous insurance plans may be more likely to pay for expensive drugs (if their insurer permits) than customers with more limited insurance policies.

Summary

1. **Monopoly Profit Maximization.** Like any firm, a monopoly—a single seller—maximizes its profit by setting its output so that its marginal revenue equals its marginal cost. The monopoly makes a positive profit if its average cost is less than the price at the profit-maximizing output.

2. **Market Power.** Market power is the ability of a firm to charge a price above marginal cost and earn a positive profit. The more elastic the demand the monopoly faces at its profit-maximizing quantity, the closer is its price to its marginal cost and the closer is the Lerner Index or price markup, $(p - MC)/p$, to zero, the competitive level.

3. **Market Failure Due to Monopoly Pricing.** Because a monopoly's price is above its marginal cost, too little output is produced, and society suffers a deadweight loss. The monopoly makes higher profit than it would if it acted as a price taker. Consumers are worse off, buying less output at a higher price.

4. **Causes of Monopoly.** A firm may be a monopoly if it has lower operating costs than rivals, such as from superior knowledge or control of a key input. A market may also have a *natural monopoly* if one firm can produce the market output at lower average cost than can a larger number of firms (even if all firms have the same cost function). Many, if not most, monopolies are created by governments, which prevent other firms from entering the markets. One important barrier to entry is a patent, which gives the inventor of a new product or process the exclusive right to sell the product or to use the process for 20 years in most countries.

5. **Government Actions That Reduce Market Power.** A government can eliminate the welfare harm of a monopoly by forcing the firm to set its price at the competitive level. If the government sets the price at a different level or otherwise regulates nonoptimally, welfare at the regulated monopoly optimum is lower than in the competitive equilibrium. A government can eliminate or reduce the harms of monopoly by allowing or facilitating entry.

6. **Networks, Dynamics, and Behavioral Economics.** If a good has a positive network externality so that its value to a consumer grows as the number of units sold increases, then current sales affect a monopoly's future demand curve. A monopoly may maximize its long-run profit—its profit over time—by setting a low introductory price in the first period that it sells the good and then later raising the price as the product's popularity ensures large future sales at a higher price. Consequently, the monopoly is not maximizing its short-run profit in the first period but it is maximizing the sum of its profits over all periods. Behavioral economics provides an explanation for some network externalities, such as bandwagon effects and snob effects.

Questions

Select questions are available on MyLab Economics; ** = answer appears at the back of this book;* **A** *= algebra problem;* **C** *= calculus problem.*

1. Monopoly Profit Maximization

1.1 If the linear inverse demand function is $p = 100 - 2Q$, what is the marginal revenue function? Draw the demand and marginal revenue curves. (*Hint*: See Solved Problem 11.1) **A**

1.2 Does a monopoly's profit differ if it chooses price or quantity (assuming it chooses them optimally)? Why can't a monopoly choose both price and quantity?

*1.3 If the inverse demand function is $p = 500 - 10Q$, what is the elasticity of demand and revenue at $Q = 10$?

1.4 According to a January 5, 2017, article in *The Guardian*, consumer spending in Britain on subscription streaming services for movies and TV shows grew by 23% between 2015 and 2016 to £1.3 billion and, for the first time, exceeded sales of DVDs and Blu-ray discs, which fell by 17% to £0.9 billion in 2016. If the market for physical video discs were a monopoly, how far would its demand curve have to fall before the firm exited the industry? Use a diagram to explain your answer.

1.5 Show why a monopoly may operate in the upward- or downward-sloping section of its long-run average cost curve but a competitive firm will operate only in the upward-sloping section.

1.6 The inverse demand curve a monopoly faces is $p = 100 - Q$. The firm's cost curve is $C(Q) = 10 + 5Q$ (so $MC = 5$). What is the profit-maximizing solution? How does your answer change if $C(Q) = 100 + 5Q$? (*Hint*: See Solved Problem 11.2.) **A**

1.7 If a monopolist produces its product in two factories, it must determine how to allocate production among those locations to minimize its total costs of production and thereby maximize its profits. It will do this by producing more at the factory with the lower marginal cost of production until the marginal cost of the last unit produced in each factory is equal: $MC_1 = MC_2$. If the marginal cost of producing at Factory 1 is $MC_1 = 2Q_1$, the marginal cost of producing at Factory 2 is $MC_2 = 4Q_2$, and the firm's demand curve is $p = 70 - Q$, where $Q = Q_1 + Q_2$, how much will the monopolist produce in each of its two factories, and what price will it set for its product? **A**

1.8 Use a figure to show that after a shift in the demand curve, a monopoly's price may remain constant but its output may rise.

1.9 With a seating capacity of 99,354 at present, Camp Nou is the largest association football stadium in Europe and the second largest in the world. It is the home stadium for, and owned by, Futbol Club Barcelona (FC Barcelona). Its seating capacity has changed over the years from 93,000 when it first opened in 1957 to over 115,000 in the early 1980s. In the 1990s, standing areas were converted to seats and the capacity reduced to its present level. Given that FC Barcelona sells out very few games per season, what economic reason might this monopolist have had to reduce its seating capacity?

1.10 The inverse demand curve a monopoly faces is $p = 10Q^{-0.5}$.

a. What is the firm's marginal revenue curve?

b. The firm's cost curve is $C(Q) = 5Q$. What is the profit-maximizing solution? (*Hint*: See Solved Problem 11.2.) **C**

1.11 If a monopolist faces the linear demand curve $p = 12 - Q$, what is its level of output, revenue, and the elasticity of demand when $MR = 0$? If a monopolist were to decrease its output from 4 to 2 units, what would happen to its revenue, and what is its marginal revenue and the elasticity of demand at those two levels of output? What does a comparison of your results tell you about where the monopoly operates on its demand curve? (*Hint*: Marginal cost can never be negative.)

1.12 The demand curve $Q = A/p$, where A is a positive constant, has the property that the elasticity of demand is $= -1$ everywhere.

a. Use math to show that the revenue is the same at any given point on the constant elasticity of demand curve. **C**

b. Show that, for any point on the constant elasticity of demand curve, the corresponding marginal revenue is zero. **C**

1.13 The demand curve $Q = A/p$, where A is a positive constant, has the property that the elasticity of demand is $= -1$. If a monopoly faces this demand curve, where would it set its price or quantity if it has a positive marginal cost? Explain. Is this situation plausible?

2. Market Power

2.1 Under what circumstances does a monopoly set its price equal to its marginal cost? (*Hint*: consider the elasticity of demand. Also consider the issue raised in the Application "Cable Cars and Profit Maximization.")

2.2 Regardless of the sport, the owners of professional sports teams are local monopolists. Many people argue that increasing attendance (and revenue) at sporting events should be a key objective of any sports club. Can a monopolist maximize both profit and revenue? (*Hint*: Compare the point where revenue is at a maximum along a linear demand curve with the portion of the demand curve along which a monopolist operates.)

2.3 Why is the ratio of the monopoly's price to its marginal cost, p/MC, smaller, the less elastic the demand curve at the optimum quantity? Can the demand curve be inelastic at that quantity?

2.4 When will a monopoly set its price equal to its marginal cost?

2.5 Draw an example of a monopoly with a linear demand curve and a constant marginal cost curve.

a. Show the profit-maximizing price and output, p^* and Q^*, and identify the areas of consumer surplus, producer surplus, and deadweight loss. Also show the quantity, Q_c, that would be produced if the monopoly were to act like a price taker.

b. Now suppose that the demand curve is a smooth, concave to the origin curve , and is tangent to the original demand curve at the point (Q^*, p^*). Explain why the monopoly optimum will be the same as with the linear demand curve. Show how much output the firm would produce if it acted like a price taker. Show how the welfare areas change.

c. Repeat the exercises in part b with a demand curve that is a smooth, *convex* to the origin curve that is tangent to the original demand curve at the point (Q^*, p^*).

2.6 If the inverse demand curve for rides at the Yamit Water Park located in the city of Holon, near Tel Aviv, Israel, is $p = 200 - Q/50$, where p is the price of admission in shekels and Q is the number of visitors per day, what is the revenue-maximizing price? If Holon estimates that local businesses benefit by IS50 per day for each tourist that visits the water park, what price should it charge for admission to maximize both water park revenues and spin-off benefits for local businesses? **C**

2.7 Using the information in Solved Problem 11.2, calculate the elasticity of demand for the iPad at the profit-maximizing solution using its inverse demand curve. Also calculate the inverse demand curve using the Lerner Index method in Solved Problem 11.3.

2.8 In 2015, Apple introduced the Apple Watch. According to HIS, the cost of producing the 38mm Apple Watch Sport was \$84. The price was \$349. What was Apple's price/marginal cost ratio? What was its Lerner Index? If Apple is a short-run profit-maximizing monopoly, what elasticity of demand did Apple believe it faced? (*Hint*: See Solved Problem 11.3.) **A**

2.9 A 2011 report by the management consulting firm O'Rourke Group Partners indicated that a generic \$14 polo shirt sold in Canada and made in Bangladesh actually costs a retailer only \$5.67 (*Maclean's*, 2013; Ecouterre, n.d.). Assuming a constant marginal cost, calculate the retailer's price/marginal cost ratio, its Lerner Index, and the elasticity of demand it believes it faces (assuming that it is trying to maximize its short-run profit). (*Hint*: See Solved Problem 11.3.) **A**

*2.10 Introduced in the nineteenth century, bicycles are the principal means of transportation in many parts of the world. Suppose there is a 60% markup on the marginal cost of a bicycle tire that sells for ₹1,400. What is the firm's Lerner Index? If the firm is a profit-maximizing monopoly, what is the elasticity of demand it faces for its tires? (*Hint*: See Solved Problem 11.3.) **A**

2.11 When Apple introduced its iPod, its constant marginal cost of producing its top-of-the-line iPod was \$200 (iSuppli), its fixed cost was approximately \$736 million, and I estimate that its inverse demand function was $p = 600 - 25Q$, where Q is units measured in millions. What was Apple's average cost function? Assuming that Apple was maximizing short-run monopoly profit, what was its marginal revenue function? What were its profit-maximizing price and quantity, profit, and Lerner Index? What was the elasticity of demand at the profit-maximizing level? Show Apple's profit-maximizing solution in a figure. What was the deadweight loss from monopoly pricing? (*Hint*: See Solved Problems 11.2 and 11.3.) **A**

3. Market Failure Due to Monopoly Pricing

3.1 A monopoly has a constant marginal cost of production of \$1 per unit and a fixed cost of \$10. Draw the firm's MC, AVC, and AC curves. Add a downward-sloping demand curve, and show the profit-maximizing quantity and price. Indicate the profit as an area on your diagram. Show the deadweight loss. (*Hint*: See Solved Problem 11.4.)

3.2 What is the effect of a franchise (lump-sum) tax on a monopoly? (*Hint*: Consider the possibility that the firm may shut down.) (*Hint*: See Solved Problem 11.4.)

*3.3 Suppose that a monopoly faces the short-run cost function $C = 1{,}000 + 2Q^2$ and the inverse demand curve $p = 500 - 0.5Q$. How does charging the monopoly a per-unit tax of $t = 100$ affect the monopoly optimum and its profits? If a profits tax were instead imposed, would it affect the monopoly optimum? What would be the rate of a profits tax that raises the same amount of tax revenue as the specific tax? (*Hint*: See Solved Problem 11.4.)

3.4 If the inverse demand curve is $p = 120 - Q$ and the marginal cost is constant at 10, how does charging the monopoly a specific tax of $t = 10$ per unit affect the monopoly optimum and the welfare of consumers, the monopoly, and society (where society's welfare includes the tax revenue)? What is the incidence of the tax on consumers? (*Hint*: See Solved Problem 11.4.) **A**

*3.5 Show mathematically that a monopoly may raise the price to consumers by more than the specific tax t imposed on it. (*Hints*: Consider a monopoly facing a constant-elasticity demand curve, $Q = Ap$, and a constant marginal cost, m. See Solved Problem 11.4.) **C**

3.6 Show that a monopoly will not necessarily lower its price by the same percentage as its constant marginal cost drops. (*Hint*: See Solved Problem 11.4.)

3.7 If the inverse demand function facing a monopoly is $p(Q)$ and its cost function is $C(Q)$, show the effect of a specific tax, t, on its profit-maximizing output. How does imposing t affect its profit? (*Hint*: See Solved Problem 11.4.) **C**

3.8 A monopoly with a constant marginal cost m has a profit-maximizing price of p_1. It faces a constant elasticity demand curve with elasticity . After the government applies a specific tax of \$1, its price is p_2. What is the price change $p_2 - p_1$ in terms of ? How much does the price rise if the demand elasticity is -2? (*Hint*: Use Equation 11.9.) **C**

3.9 Consider the inverse demand curve $p = 210 - 3Q$ and the cost function $C = 100 + 2Q^2$. If the market were competitive, calculate the incidence of a specific tax, $t = 7$, that would fall on consumers. Calculate the incidence of the same tax if the market were instead a monopoly. **A**

3.10 In a figure, show the effect of an ad valorem tax (see Chapter 2) on a monopoly optimum, consumer surplus, producer surplus, welfare, and deadweight loss.

4. Causes of Monopoly

*4.1 Can a firm be a natural monopoly if it has a U-shaped average cost curve? Why or why not? (*Hint*: See Solved Problem 11.5.)

4.2 Can a firm operating in the upward-sloping portion of its average cost curve be a natural monopoly? Explain. (*Hint*: See Solved Problem 11.5.)

4.3 In the Application "Botox Patent Monopoly," consumer surplus, triangle A, equals the deadweight loss, triangle C. Show that this equality is a result of the linear demand and constant marginal cost assumptions. **A**

4.4 Based on the information in the Botox Application, what would happen to the optimum price and quantity if the government had collected a specific tax of \$75 per vial of Botox? What welfare effects would such a tax have? **A**

4.5 Once the copyright runs out on a book or musical composition, the work can legally be posted on the Internet for anyone to download. U.S. copyright law protects the monopoly for 95 years after the original publication. But in Australia and Europe, the copyright holds for only 50 years. Thus, an Australian website can post *Gone With the Wind*, a 1936 novel, or Elvis Presley's 1954 single "That's All Right," while a U.S. site cannot. Obviously, this legal nicety won't stop American fans from downloading from Australian or European sites. Discuss how limiting the length of a copyright would affect the pricing used by the publisher of a novel.

5. Government Actions That Reduce Market Power

5.1 Electric utilities are often natural monopolies. Due to the extensive and costly infrastructure needed to generate, transmit, and deliver electricity, one firm can produce the total output of the market at a lower cost than could several firms. Regulatory authorities such as the Central Electricity Regulatory Commission in India and the Australian Energy Regulator oversee the operation of electric utilities, and many limit their market power by setting electricity rates for residential, commercial, and industrial customers. If the marginal cost of production for a monopoly is upward sloping and the linear market demand curve is downward sloping, draw a diagram indicating the socially optimal price. If the regulatory authority sets a price ceiling that is above the socially optimal price but below the monopoly's profit-maximizing price, how do the price, quantity, and welfare compare to those under optimal regulation? (*Hint*: See Solved Problem 11.6.)

5.2 Describe the effects on output and welfare if the government regulates a monopoly so that it may not charge a price above $\bar{p}$, which lies between the unregulated monopoly price and the optimally regulated price (determined by the intersection of the firm's marginal cost and the market demand curve). (*Hint*: See Solved Problem 11.6.)

5.3 Based on the information in the Application "Botox Patent Monopoly," what would happen to the optimum price and quantity if the government had set a price ceiling of $200 per vial of Botox? What welfare effects would such a restriction have? (*Hint*: See Solved Problem 11.6.) **A**

5.4 A monopoly drug company produces a lifesaving medicine at a constant cost of $10 per dose. The demand for this medicine is perfectly inelastic at prices less than or equal to the $100 (per day) income of the 100 patients who need to take this drug daily. At a higher price, nothing is bought. Show the optimum price and quantity and the consumer and producer surplus in a graph. Now the government imposes a price ceiling of $30. Show how the optimum, consumer surplus, and producer surplus change. What is the deadweight loss, if any, from this price control?

5.5 International trade benefits countries by increasing total production and the quantity of goods and services available to consumers. However, not everyone wins. While trade creates new jobs for firms that export products to foreign markets, companies that are less efficient may shut down and jobs may be lost. Use a diagram to show how import competition from a lower wage country could drive a monopoly out of business. (*Hint*: See Solved Problem 11.7.)

5.6 Malaysia's monopoly auto manufacturer produces the Proton, which is protected from imports by a specific tariff, t, on imported goods. The monopoly's profit-maximizing price is p^*. The world price of the good (comparable autos) is p_w, which is less than p^*. Because the price of imported goods with the tariff is $p_w + t$, no foreign goods are imported. Under WTO pressure the government removes the tariff so that the supply of foreign goods to the country's consumers is horizontal at p_w. Show how much the former monopoly produces and what price it charges. Show who gains and who loses from removing the tariff. (*Hint*: Look at Solved Problem 11.7.)

6. Networks, Dynamics, and Behavioral Economics

6.1 A monopoly chocolate manufacturer faces two types of consumers. The larger group, the hoi polloi, loves desserts and has a relatively flat, linear demand curve for chocolate. The smaller group, the snobs, is interested in buying chocolate only if the hoi polloi do not buy it. Given that the hoi polloi do not buy the chocolate, the snobs have a relatively steep, linear demand curve. Show the monopoly's possible outcomes—high price, low quantity, low price, high quantity—and explain the condition under which the monopoly chooses to cater to the snobs rather than to the hoi polloi.

*6.2 A firm is considering selling a new good at an introductory price that is less than the monopoly price. By doing so, it hopes to create a critical mass of users and benefit from an increased future demand generated by a positive network externality for the product. The marginal cost of production is constant at $MC = 4$ and equal to the average cost. The inverse demand curve for the product is $p = 20 - 4Q$.

a. If the firm were to charge the monopoly price, what would its total profits be over two periods?

b. If the firm tried to take advantage of the positive network externality by instead charging the competitive price in the first period and the monopoly price in the second period, by how much would the demand curve have to rotate outward (that is, its slope has to change) before the firm's total profits over the two periods exceeded your answer to part a? **A**

7. Challenge

7.1 Under what circumstances will a drug company charge more for its drug after its patent expires?

7.2 Does the Challenge Solution change if the entry of the generic causes a parallel shift to the left of the patent monopoly's linear demand curve?

7.3 A patent grants the holder the exclusive right to use or sell an invention for the duration of the term of the patent. The Agreement on Trade-Related Aspects of Intellectual Property Rights was signed on April 15, 1994, by 123 nations to protect intellectual property rights. Article 33 of that World Trade Organization agreement provides for the length of patent protection to be no less than 20 years from the date a patent is filed. This term is now standard in the laws of most countries. Would a change in the term of patent protection change the price a monopoly charged during the patent period? For example, would prices and monopoly profit rise if the term were to be reduced to 15 years?

12 Pricing and Advertising

Everything is worth what its purchaser will pay for it.
—Publilius Syrus (first century B.C.)

Challenge

Sale Prices

Because many firms use *sales*—temporarily setting the price below the usual price—customers who always buy *on sale* pay less than do others who do not regularly buy on sale. Grocery stores are particularly likely to put products on sale frequently. In large U.S. supermarkets, some soft drink brand is on sale 94% of the time. Either Coke or Pepsi is on sale half the weeks in a year.

Heinz Ketchup controlled 62% of the U.S. ketchup market in 2015. Its market share was 84% in Canada, and nearly 78% in the United Kingdom. Kraft Heinz Co. sells over 650 million bottles of ketchup in more than 140 countries and had annual sales of more than $1.5 billion. When Heinz goes on sale, *switchers*—ketchup customers who normally buy whichever brand is least expensive—purchase Heinz rather than the low-price generic ketchup. How can Heinz's managers design a pattern of sales that maximizes Heinz's profit by obtaining extra sales from switchers without losing substantial sums by selling to its loyal customers at a discount price? Under what conditions does it pay for Heinz to have a policy of periodic sales?

Sales are not the only means that firms use to charge customers different prices. Why are airline fares often substantially less if you book in advance? Why do the spiritualists who live at the Wonewoc Spiritualist Camp give readings for $45 for half an hour, but charge seniors only $40 on Wednesdays?[1] Why are some goods, including computers and software, combined and sold as a bundle? To answer these questions, we need to examine how monopolies and other noncompetitive firms set prices.

uniform pricing
charging the same price for every unit sold of a particular good

nonuniform pricing
charging consumers different prices for the same product or charging a single customer a price that depends on the number of units the customer buys

In Chapter 11, we examined how a monopoly chooses a single price when it uses **uniform pricing**, charging the same price for every unit sold of a particular good. However, a monopoly can increase its profit if it can use **nonuniform pricing**, where a firm charges consumers different prices for the same product or charges a single customer a price that depends on the number of units the customer buys. In this chapter, we analyze nonuniform pricing for monopolies, but similar principles apply to any firm with market power.

As we saw in Chapter 11, a monopoly that sets a uniform price sells only to customers who value the good enough to buy it at the monopoly price, and those customers receive some consumer surplus. The monopoly does not sell the good to other customers who value the good at less than the single price, even if those consumers

[1] **http://www.campwonewoc.com** (viewed September 12, 2013).

would be willing to pay more than the marginal cost of production. These lost sales cause *deadweight loss*, which is the forgone value of these potential sales in excess of the cost of producing the good.

A firm with market power can earn a higher profit using nonuniform pricing than by setting a uniform price for two reasons. First, the firm captures some, or all, of the single-price consumer surplus. Second, the firm converts at least some of the single-price deadweight loss into profit by charging a price below the uniform price to some customers who would not purchase at the single-price level. A monopoly that uses nonuniform pricing can lower the price to these otherwise excluded consumers without lowering the price to consumers who are willing to pay higher prices.

We examine several types of nonuniform pricing including price discrimination, two-part pricing, and tie-in sales. The most common form of nonuniform pricing is *price discrimination*, where a firm charges various consumers different prices for a good. For example, for a full-year combination print and online subscription, the *Wall Street Journal* charges \$49 to students, who are price sensitive, and \$155 to other subscribers, who are less price sensitive.

Some firms with market power use other forms of nonuniform pricing to increase profits. A firm may use *two-part pricing*, where it charges a customer one fee for the right to buy the good and an additional fee for each unit purchased. For example, members of health or golf clubs typically pay an annual fee to belong to the club and then pay an additional amount each time they use the facilities. Similarly, cable television companies often charge a monthly fee for basic service and an additional fee for recent movies.

Another type of nonuniform pricing is a *tie-in sale* where customers can purchase one product only if they agree to buy another product as well. One example is *bundling*, where several products are sold together as a package. For example, many restaurants provide full-course dinners for a fixed price that is less than the sum of the prices charged if the items (appetizer, main dish, and dessert) are ordered separately (à la carte).

A monopoly may also increase its profit by advertising. A monopoly may advertise to shift its demand curve so as to raise its profit, taking into account the cost of advertising.

In this chapter, we examine seven main topics

1. **Conditions for Price Discrimination.** A firm can increase its profit using price discrimination if it has market power, if customers differ in their willingness to pay, if the firm can identify which customers are more price sensitive than others, and if it can prevent customers who pay low prices from reselling to those who pay high prices.
2. **Perfect Price Discrimination.** If a monopoly can charge the maximum each customer is willing to pay for each unit of output, the monopoly captures all potential consumer surplus, and the efficient (competitive) level of output is sold.
3. **Group Price Discrimination.** A firm that lacks the ability to charge each individual a different price may be able to charge different prices to various groups of customers that differ in their willingness to pay for the good.
4. **Nonlinear Price Discrimination.** A firm may set different prices for large purchases than for small ones, discriminating among consumers by inducing them to self-select the effective price they pay based on the quantity they buy.
5. **Two-Part Pricing.** By charging consumers a fee for the right to buy a good and then allowing them to purchase as much as they wish at an additional per-unit fee, a firm earns a higher profit than with uniform pricing.
6. **Tie-In Sales.** By selling a combination of different products in a package or bundle, a firm earns a higher profit than by selling the goods or services separately.
7. **Advertising.** A monopoly advertises in order to shift its demand curve and increase its profit.

12.1 Conditions for Price Discrimination

The prince travels through the forest for many hours and comes upon an inn, where he is recognized immediately. He orders a light meal of fried eggs. When he finishes, the prince asks the innkeeper, "How much do I owe you for the eggs?" The innkeeper replies, "Twenty-five rubles." "Why such an exorbitant price?" asks the prince. "Does this area have a shortage of eggs?" The innkeeper says, "We don't have a shortage of eggs, but we have a shortage of princes."[2]

price discrimination charging consumers different prices for the same good

We begin our discussion of firms' pricing practices by studying the most common form of nonuniform pricing, **price discrimination**: charging consumers different prices for the same good based on individual consumer characteristics, membership in an identifiable subgroup of consumers, or on the quantity consumers purchase.[3]

Why Price Discrimination Pays

For almost any good or service, some consumers are willing to pay more than others. A firm that sets a single price faces a trade-off between charging consumers who really want the good as much as they are willing to pay and charging a low enough price that the firm doesn't lose sales to less enthusiastic customers. As a result, the firm usually sets an intermediate price. A price-discriminating firm that varies its prices across customers avoids this trade-off.

As with any kind of nonuniform pricing, price discrimination increases profit above the uniform pricing level through two channels. First, a price-discriminating firm charges a higher price to customers who are willing to pay more than the uniform price, capturing some or all of their consumer surplus—the difference between what a good is worth to a consumer and what the consumer paid—under uniform pricing. Second, a price-discriminating firm sells to some people who were not willing to pay as much as the uniform price.

We use a pair of extreme examples to illustrate the two benefits of price discrimination to firms—capturing more of the consumer surplus and selling to more customers. These examples are extreme in the sense that the firm sets a uniform price at the price the most enthusiastic consumers are willing to pay or at the price the least enthusiastic consumers are willing to pay, rather than at an intermediate level.

Suppose that the only movie theater in town has two types of patrons: college students and senior citizens. The college students will see the Saturday night movie if the price is $10 or less, and the senior citizens will attend if the price is $5 or less. For simplicity, we assume that the theater incurs no cost to show the movie, so profit is the same as revenue. The theater is large enough to hold all potential customers, so the marginal cost of admitting one more customer is zero. Table 12.1 shows how pricing affects the theater's profit.

In panel a of the table, 10 college students and 20 senior citizens want to go to the movie. If the theater charges everyone $5, its profit is $\$150 = \$5 \times$ (10 college students + 20 senior citizens). If it charges $10, the senior citizens do not go to the movie, so the theater makes only $100. Thus, if the theater is going to charge everyone the same price, it maximizes its profit by setting the price at $5. Charging less than $5 lowers the theater's profit because the same number of people

[2]Thanks to Steve Salop.

[3]Price discrimination is legal in the United States unless it harms competition between firms, as specified in the Robinson-Patman Act.

Table 12.1 A Theater's Profit Based on the Pricing Method Used

(*a*) *No Extra Customers from Price Discrimination*

Pricing	Profit from 10 College Students	Profit from 20 Senior Citizens	Total Profit
Uniform, $5	$50	$100	$150
Uniform, $10	$100	$0	$100
Price discrimination*	$100	$100	$200

(*b*) *Extra Customers from Price Discrimination*

Pricing	Profit from 10 College Students	Profit from 5 Senior Citizens	Total Profit
Uniform, $5	$50	$25	$75
Uniform, $10	$100	$0	$100
Price discrimination*	$100	$25	$125

*The theater price discriminates by charging college students $10 and senior citizens $5.

Notes: College students go to the theater if they are charged no more than $10. Senior citizens are willing to pay at most $5. The theater's marginal cost for an extra customer is zero.

go to the movie when the price is less than $5 as when $5 is charged. Charging between $5 and $10 is less profitable than charging $10 because no extra seniors go and the college students are willing to pay $10. Charging more than $10 results in no customers.

At a price of $5, the seniors have no consumer surplus: They pay exactly what seeing the movie is worth to them. Seeing the movie is worth $10 to the college students, but they have to pay only $5, so each has a consumer surplus of $5, and their total consumer surplus is $50.

If the theater can price discriminate by charging senior citizens $5 and college students $10, its profit increases to $200. Its profit rises because the theater makes as much from the seniors as before but gets an extra $50 from the college students. By price discriminating, the theater sells the same number of seats but makes more money from the college students, capturing all the consumer surplus they received under uniform pricing. Neither group of customers has any consumer surplus if the theater price discriminates.

Panel b has 10 college students and 5 senior citizens who want to go to the movie. If the theater must charge a single price, it charges $10. Only college students buy tickets, so the theater's profit is $100. (If it charges $5, both students and seniors go to the theater, but its profit is only $75.) If the theater can price discriminate and charge seniors $5 and college students $10, its profit increases to $125. Here the gain from price discrimination comes from selling extra tickets to seniors (not from making more money on the same number of tickets, as in panel a). The theater earns as much from the students as before and makes more from the seniors, and neither group enjoys consumer surplus. Leslie (1997) found that Broadway theaters in New York increase their profits 5% by price discriminating rather than using uniform prices.

These examples illustrate the two channels through which price discrimination can increase profit: charging some existing customers more or selling extra units. The movie theater's ability to increase its profits by price discrimination arises from its ability to segment the market into two groups, students and senior citizens, with different levels of willingness to pay.

Application

Disneyland Pricing

In 2016, Disneyland charged local Southern Californians $149 for two visits to Disneyland or Disney's California Adventure park, but charged people from out of the area $185 for a two-day pass to only one of the parks. This policy of charging locals a discounted price makes sense if out-of-town visitors are willing to pay more than locals and if Disneyland can prevent locals from selling discounted tickets to nonlocals. Imagine a Midwesterner who's never been to Disneyland and wants to visit. Travel accounts for most of the trip's cost, so an extra few dollars for entrance to the park makes little percentage difference in the total cost of the visit and hence does not greatly affect that person's decision whether to go. In contrast, for a local who has gone to Disneyland often and for whom the entrance price is a larger share of the total cost, a slightly higher entrance fee might prevent a visit.[4]

Charging both groups the same price is not in Disney's best interest. If Disney were to charge the higher price to everyone, many locals wouldn't visit the park. If Disney were to use the lower price for everyone, it would be charging nonresidents much less than they are willing to pay.

Which Firms Can Price Discriminate

Not all firms can price discriminate. For a firm to price discriminate successfully, three conditions must be met.

First, a firm must have market power. Without market power, a firm cannot charge any consumer more than the competitive price. A monopoly, an oligopoly firm, or a monopolistically competitive firm might be able to price discriminate. However, a perfectly competitive firm cannot price discriminate because it must sell its product at the market price.

Second, for a firm to profitably discriminate, groups of consumers or individual consumers must have demand curves that differ, and the firm must be able to identify how its consumers' demand curves differ. The movie theater knows that college students and senior citizens differ in their willingness to pay for a ticket, and Disneyland knows that tourists and local residents differ in their willingness to pay for admission. In both cases, the firms can identify members of these two groups by using driver's licenses or other forms of identification. Similarly, if a firm knows that each individual's demand curve slopes downward, it may charge each customer a higher price for the first unit of a good than for subsequent units.

Third, a firm must be able to prevent or limit resale. The price-discriminating firm must be able to prevent consumers who buy the good at low prices from reselling the good to customers who would otherwise pay high prices. Price discrimination

[4]In 2012, a Southern Californian couple, Jeff Reitz and Tonya Mickesh, were out of work, so they decided to cheer themselves up by using their annual passes to visit Disneyland 366 days that year (a leap year).

doesn't work if resale is easy because the firm would be able to make only low-price sales. Disneyland and movie theaters can charge different prices for different groups of customers because those customers normally enter as soon as they buy their tickets and don't have time to resell them. For events that sell tickets in advance, firms can use other methods to prevent resale, such as having different colors for children's tickets and adults' tickets.

The first two conditions—market power and the ability to identify groups with different price sensitivities—are present in many markets. Usually, the biggest obstacle to price discrimination is a firm's inability to prevent resale.

Preventing Resale

In some industries, preventing resale is easier than in others. In industries where resale is initially easy, firms can act to make resale more difficult.

Resale is difficult or impossible for most *services* and when *transaction costs are high*. If a plumber charges you less than your neighbor for clearing a pipe, you cannot make a deal with your neighbor to resell this service. The higher the transaction costs that a consumer must incur to resell a good, the less likely that resale will occur. Suppose you are able to buy a jar of pickles for $1 less than the usual price. Could you practically find and sell this jar to someone else, or would the transaction costs be prohibitive? The more valuable or widely consumed a product is, the more likely it is that transaction costs are low enough that resale occurs.

Some firms act to raise transaction costs or otherwise make resale difficult. If your college requires that someone with a student ticket must show a student identification card with a picture on it before being admitted to a sporting event, you'll find it difficult to resell your low-price tickets to nonstudents who pay higher prices. When students at some universities buy computers at lower-than-usual prices, they must sign a contract that forbids them to resell the computer.

Similarly, a firm can prevent resale by *vertically integrating*: participating in more than one successive stage of the production and distribution chain for a good or service. Alcoa, the former aluminum monopoly, wanted to sell aluminum ingots to producers of aluminum wire at a lower price than was set for producers of aluminum aircraft parts. If Alcoa did so, however, the wire producers could easily resell their ingots. By starting its own wire production firm, Alcoa prevented such resale and was able to charge high prices to firms that manufactured aircraft parts (Perry, 1980).

Governments frequently aid price discrimination by preventing resale. State and federal governments require that milk producers, under penalty of law, price discriminate by selling milk for fresh use at a higher price than milk for processing (cheese, ice cream) and forbid resale. Government *tariffs* (taxes on imports) limit resale by making it expensive to buy goods in a low-price country and resell them in a high-price country. In some cases, laws prevent such reselling explicitly. Under U.S. copyright law, certain brand-name perfumes may not be sold in the United States except by their manufacturers.

Application

Preventing Resale of Designer Bags

During the holiday season, stores often limit how many of the hottest items—such as this year's best-selling toy—a customer can buy. But it may surprise you that websites of luxury-goods retailers such as Saks Fifth Avenue, Neiman Marcus, and Bergdorf Goodman limit how many designer handbags one can buy. For example, the Bergdorf Goodman site won't let you order more than one Prada Lizard Trimmed Nylon Shoulder Bag at $4,190.

Some websites explain that they impose limits due to "popular demand." The more plausible explanation is that the restriction facilitates international price discrimination. The handbag manufacturers force U.S. retailers to limit the number of bags that one person can buy to prevent people from buying large numbers of bags and reselling them in Europe or Asia where the same Prada and Gucci item often costs 20% to 40% more. When purchasing from Prada's U.S. online site, one must agree that the purchase is solely for private household use, that commercial resale or sale outside of the United States is not authorized, and that the company reserves the right to reject orders and to limit order quantities.

Not All Price Differences Are Price Discrimination

Not every seller who charges consumers different prices is price discriminating. Hotels charge newlyweds more for bridal suites. Is that price discrimination? Some hotel managers say no. They contend that honeymooners, unlike other customers, always steal mementos, so the price differential reflects an actual cost differential.

The 2016 subscription price for print and digital issues of the *Economist* magazine for a year is $190 for a standard subscription, and $115 for a college student subscription. The price difference between the standard subscription rate and the college student rate reflects pure price discrimination because the two subscriptions are identical in every respect except the price. The standard subscription is only 47% of the newsstand price, $407.49, which reflects in part the higher cost of selling the magazine at a newsstand, rather than mailing it directly to customers. Thus, this price difference does not reflect pure price discrimination.

perfect price discrimination (*first-degree price discrimination*) selling each unit at the maximum amount any customer is willing to pay for it, so prices differ across consumers, and a given consumer may pay a higher price for some units than for others

group price discrimination (*third-degree price discrimination*) charging each group of customers a different price, but charging the same price within the group

nonlinear price discrimination (*second-degree price discrimination*) charging a different price for large quantities than for small quantities, but charging all customers who buy a given quantity the same price

Types of Price Discrimination

Traditionally, economists focus on three types of price discrimination: perfect price discrimination, group price discrimination, and nonlinear price discrimination. With **perfect price discrimination** (also called *first-degree price discrimination*), a firm sells each unit at the maximum amount any customer is willing to pay for it. Under perfect price discrimination, price differs across consumers, and a given consumer may pay higher prices for some units than for others.

With **group price discrimination** (also called *third-degree price discrimination*), a firm charges each group of customers a different price, but it does not charge different prices within the group. The price that a firm charges a consumer depends on that consumer's membership in a particular group. Thus, not all customers pay different prices—the firm sets different prices only for a few groups of customers. Because group price discrimination is the most common type of price discrimination, the phrase *price discrimination* is often used to mean *group price discrimination*.

A firm engages in **nonlinear price discrimination** (also called *second-degree price discrimination*) when it charges a different price for large quantities than for small quantities, so that the price paid varies according to the quantity purchased. With pure nonlinear price discrimination, all customers who buy a given quantity pay the same price; however, firms can combine nonlinear price discrimination with group price discrimination, setting different nonlinear price schedules for different groups of consumers.

12.2 Perfect Price Discrimination

reservation price
the maximum amount a person would be willing to pay for a unit of output

If a firm with market power knows exactly how much each customer is willing to pay for each unit of its good and it can prevent resale, the firm charges each person his or her **reservation price**: the maximum amount a person would be willing to pay for a unit of output. Such an all-knowing firm *perfectly price discriminates*. By selling each unit of its output to the customer who values it the most at the maximum price that person is willing to pay, the perfectly price-discriminating monopoly captures all possible consumer surplus. For example, the managers of the Suez Canal set tolls on an individual basis, taking into account many factors such as weather and each ship's alternative routes.

Perfect price discrimination is rare because firms do not have perfect information about their customers. Nevertheless, it is useful to examine perfect price discrimination because it is the most efficient form of price discrimination and provides a benchmark against which we can compare other types of nonuniform pricing.

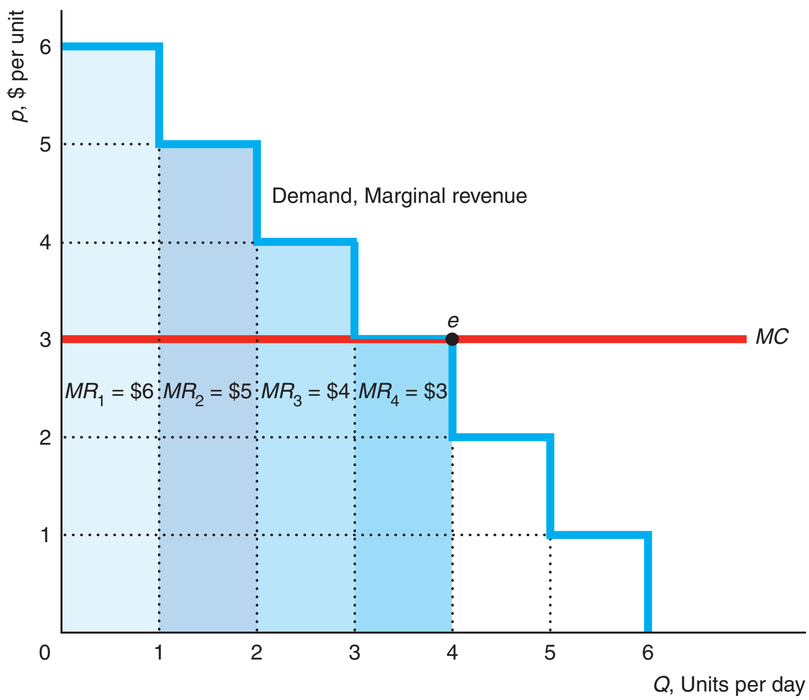
We now show how a firm with full information about consumer reservation prices can use that information to perfectly price discriminate. Then, we compare the market outcomes (price, quantity, surplus) of a perfectly price-discriminating monopoly to those of perfectly competitive and uniform-price monopoly firms. Finally, we discuss how firms obtain the information they need to perfectly price discriminate.

How a Firm Perfectly Price Discriminates

A firm with market power that can prevent resale and has full information about each customer's reservation price—the maximum amount that a customer is willing to pay—can price discriminate by selling each unit at its reservation price. We use the demand curve facing a monopoly in Figure 12.1 to illustrate how a perfectly price-discriminating firm maximizes its profit (see Appendix 12A for a mathematical treatment).

Figure 12.1 Perfect Price Discrimination

The monopoly can charge \$6 for the first unit, \$5 for the second, and \$4 for the third, as the demand curve shows. Its marginal revenue is $MR_1 = \$6$ for the first unit, $MR_2 = \$5$ for the second unit, and $MR_3 = \$4$ for the third unit. Thus, the demand curve is also the marginal revenue curve. Because the firm's marginal and average cost is \$3 per unit, it is unwilling to sell at a price below \$3, so it sells four units, point *e*, and breaks even on the last unit.



The maximum price for any unit of output equals the height of the demand curve at that output level. The figure shows that a perfectly price-discriminating firm sells its first unit of output for \$6. Having sold the first unit, the firm can get at most \$5 for its second unit. The firm must drop its price by \$1 for each successive unit it sells.

A perfectly price-discriminating monopoly's marginal revenue is the same as its price. As the figure shows, the firm's marginal revenue is $MR_1 = \$6$ on the first unit, $MR_2 = \$5$ on the second unit, and $MR_3 = \$4$ on the third unit. As a result, *the firm's marginal revenue curve is its demand curve*.

This firm has a constant marginal cost of \$3 per unit. The firm wants to produce the first unit because the firm sells that unit for \$6, so its marginal revenue exceeds its marginal cost by \$3. Similarly, the firm sells the second unit for \$5 and the third unit for \$4. The firm breaks even when it sells the fourth unit for \$3. The firm is unwilling to sell more than four units because its marginal cost would exceed its marginal revenue on all successive units. Thus, like any profit-maximizing firm, a perfectly price-discriminating firm produces at point e, where its marginal revenue curve intersects its marginal cost curve.

This perfectly price-discriminating firm earns revenues of $MR_1 + MR_2 + MR_3 + MR_4 = \$6 + \$5 + \$4 + \$3 = \$18$, which is the area under its marginal revenue curve up to the number of units, four, it sells. If the firm has no fixed cost, its cost of producing four units is $\$12 = \$3 \times 4$, so its profit is \$6.

Perfect Price Discrimination Is Efficient but Harms Some Consumers

Perfect price discrimination is efficient: It maximizes the sum of consumer surplus and producer surplus. Thus, both perfect competition and perfect price discrimination maximize total surplus. However, *with perfect price discrimination, the entire surplus goes to the firm, whereas buyers and firms share the surplus under competition*.

If the market in Figure 12.2 is competitive, the intersection of the demand curve and the marginal cost curve, MC, determines the competitive equilibrium at e_c, where price is p_c and quantity is Q_c. Consumer surplus is $A + B + C$, producer surplus is $D + E$, and society has no deadweight loss. The market is efficient because the price, p_c, equals the marginal cost, MC_c.

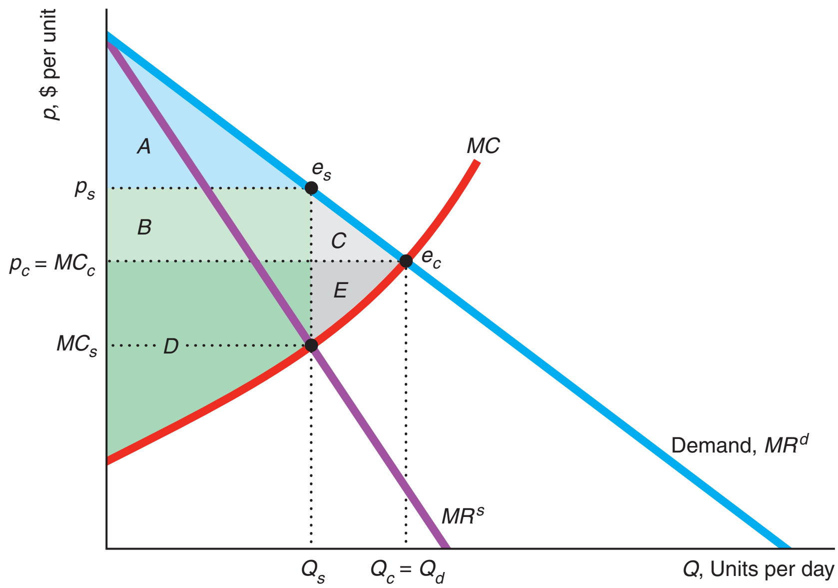
With a single-price monopoly (which charges all its customers the same price because it cannot distinguish among them), the intersection of the MC curve and the single-price monopoly's marginal revenue curve, MR^s, determines the output, Q_s.[5] The monopoly operates at e_s, where it charges p_s. The deadweight loss from single-price monopoly is $-C - E$. This efficiency loss is due to the monopoly's charging a price, p_s, that's above its marginal cost, MC_s, so less is sold than in a competitive market.

The quantity, Q_d, that the perfectly price-discriminating firm produces is determined by the intersection of the marginal cost curve, MC, and the demand curve or marginal revenue curve, MR^d. A perfectly price-discriminating firm's producer surplus from selling Q_d units is the area below its demand curve and above its marginal cost curve, $A + B + C + D + E$. Its profit is the producer surplus minus its fixed cost, if any. Consumers receive no consumer surplus because each consumer pays his or her reservation price. The perfectly price-discriminating firm's profit-maximizing solution has *no deadweight loss* because the last unit is sold at a price, p_c, that equals the marginal cost, MC_c, as in a competitive market. Thus, both a perfect price discrimination outcome and a competitive equilibrium are efficient.

[5]We assume that if we convert a monopoly into a competitive industry, the industry's marginal cost curve—the lowest cost at which any firm can produce an additional unit—is the same as the monopoly MC curve. The industry MC curve is the industry supply curve (Chapter 8).

Figure 12.2 Competitive, Single-Price, and Perfect Price Discrimination Equilibria MyLab Economics Video

In the competitive market equilibrium, e_c, price is p_c, quantity is Q_c, consumer surplus is $A + B + C$, producer surplus is $D + E$, and society has no deadweight loss. In the single-price monopoly equilibrium, e_s, price is p_s, quantity is Q_s, consumer surplus falls to A, producer surplus is $B + D$, and deadweight loss is $-C - E$. In the perfect price discrimination equilibrium, the monopoly sells each unit at the customer's reservation price on the demand curve, which is also its marginal revenue curve, MR^d. It sells Q_d (= Q_c) units, where the demand curve intersects the marginal cost curve, so that the last unit is sold at its marginal cost. Customers have no consumer surplus, but society suffers no deadweight loss.



	Competition	Monopoly: Single Price	Monopoly: Perfect Price Discrimination
Consumer Surplus, *CS*	$A + B + C$	A	0
Producer Surplus, *PS*	$D + E$	$B + D$	$A + B + C + D + E$
Welfare, $W = CS + PS$	$A + B + C + D + E$	$A + B + D$	$A + B + C + D + E$
Deadweight Loss, *DWL*	0	$-C - E$	0

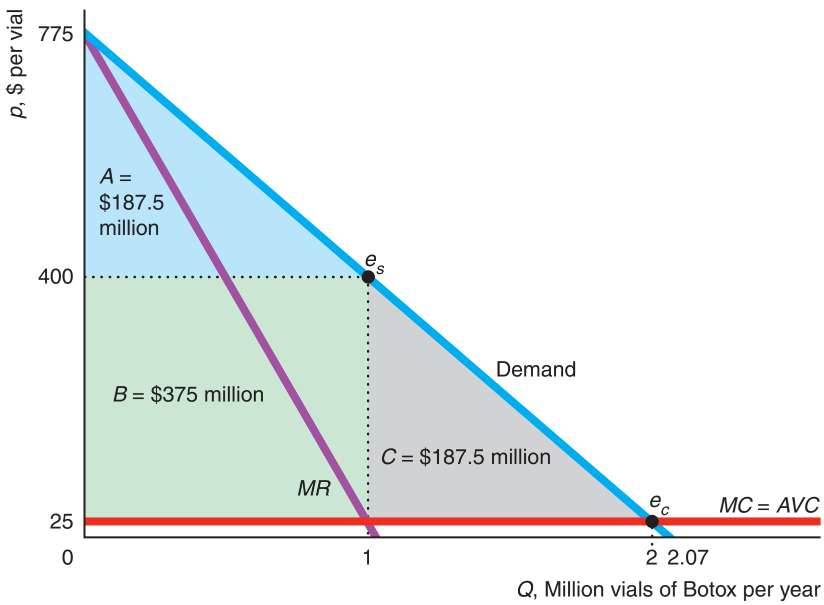
The perfect price discrimination solution differs from the competitive equilibrium in two important ways. First, in the competitive equilibrium, everyone is charged a price equal to the equilibrium marginal cost, $p_c = MC_c$; however, in the perfect price discrimination optimum, only the last unit is sold at that price. Customers buy the other units at their reservation prices, which are greater than p_c. Second, consumers receive some net benefit (consumer surplus, $A + B + C$) in a competitive market, whereas a perfectly price-discriminating monopoly captures all the surplus or potential gains from trade. Thus, perfect price discrimination does not reduce efficiency—both output and total surplus are the same as under competition—but it does redistribute income away from consumers. Consumers are much better off under competition.

Is a single-price or perfectly price-discriminating monopoly better for consumers? The perfect price discrimination optimum is more efficient than the single-price monopoly optimum because the discriminating monopoly produces more output than does the single-price monopoly. However, a single-price monopoly takes less consumer surplus from consumers than a perfectly price-discriminating monopoly. Consumers who put a very high value on the good are better off under single-price monopoly, where they have consumer surplus, than with perfect price discrimination, where they have none. Consumers with lower reservation prices who purchase from the perfectly price-discriminating monopoly but not from the single-price monopoly have no consumer surplus in either case. All the social gain from the extra output goes to the perfectly price-discriminating firm. Consumer surplus is greatest with competition, lower with single-price monopoly, and eliminated by perfect price discrimination.

Application

Botox and Price Discrimination

To show how perfect price discrimination differs from competition and single-price monopoly, we revisit the Application "Botox Patent Monopoly" in Chapter 11. The graph shows our estimated linear demand curve for Botox and a constant marginal cost (and average variable cost) of \$25 per vial. If the market had been competitive (so that price equals marginal cost at e_c), consumer surplus would have been triangle $A + B + C = \$750$ million per year, and producer surplus and deadweight loss would be zero. In the single-price monopoly optimum, e_s, the firm sells one million Botox vials at \$400 each. The corresponding consumer surplus is triangle $A = \$187.5$ million per year, producer surplus is rectangle $B = \$375$ million, and the deadweight loss is triangle $C = \$187.5$ million.



		Monopoly	
	Competition	Single Price	Perfect Price Discrimination
Consumer Surplus, *CS*	$A + B + C$	A	0
Producer Surplus, *PS*	0	B	$A + B + C$
Welfare, $W = CS + PS$	$A + B + C$	$A + B$	$A + B + C$
Deadweight Loss, *DWL*	0	$-C$	0

If Allergan, the manufacturer of Botox, could perfectly price discriminate, its producer surplus would double to $A + B + C = \$750$ million per year, and consumers would obtain no consumer surplus. The marginal consumer would pay the marginal cost of \$25, the same as in a competitive market.

Allergan's inability to perfectly price discriminate costs the company and society dearly. The profit of the single-price monopoly, $B = \$375$ million per year, is lower than what it could earn if it could use perfect price discrimination, $A + B + C = \$750$ million per year. Society's welfare under single-price monopoly is lower than from perfect price discrimination by the deadweight loss, C, of \$187.5 million per year. However, consumers have no surplus with perfect price discrimination.

Solved Problem 12.1

MyLab Economics Solved Problem

How does welfare change if the movie theater described in Table 12.1 goes from charging a single price to perfectly price discriminating?

Answer

1. *Calculate welfare for panel a (a) if the theater sets a single price and (b) if it perfectly price discriminates, and then (c) compare them.* (a) If the theater sets the profit-maximizing single price of \$5, it sells 30 tickets and makes a profit of \$150. The 20 senior citizen customers are paying their reservation price, so they have no consumer surplus. The 10 college students have reservation prices of \$10, so their consumer surplus is \$50. Thus, welfare is \$200: the sum of the profit, \$150, and the consumer surplus, \$50. (b) If the theater perfectly price discriminates, it charges seniors \$5 and college students \$10. Because the theater is charging all customers their reservation prices, they have no consumer surplus. The firm's profit rises to \$200. (c) Thus, *welfare is the same under both pricing systems where output stays the same.*

2. *Calculate welfare for panel b (a) if the theater sets a single price and (b) if it perfectly price discriminates, and then (c) compare them.* (a) If the theater sets the profit-maximizing single price of \$10, only college students attend and they have no consumer surplus. The theater's profit is \$100, so total welfare is \$100. (b) With perfect price discrimination, consumers receive no consumer surplus, but profit increases to \$125, so welfare rises to \$125. (c) Thus, *welfare is greater with perfect price discrimination where output increases*. (The result that welfare increases if and only if output rises holds generally.)

Application

Google Uses Bidding for Ads to Price Discriminate

Which ads appear next to your Google search results, depends on the terms in your search. That is, Google allows advertisers to *contextually target* people who search for particular phrases (Goldfarb, 2014). By making searches for unusual topics easy and fast, Google helps advertisers reach difficult-to-find potential customers with targeted ads. For example, a lawyer specializing in toxic mold lawsuits can place an ad that appears only when someone searches for "toxic mold lawyer."

Google uses auctions to price these ads. Advertisers are willing to bid higher for the first listing on Google's search page. Goldfarb and Tucker (2011) found that how much lawyers are willing to pay for context-based ads depends on the difficulty of making a match. Lawyers are willing to pay more to advertise, the fewer self-identified potential customers (the fewer people searching for a particular phrase).

They also found that lawyers bid more when other methods of reaching potential clients are limited. Some states have anti-ambulance-chaser regulations, which prohibit personal injury lawyers from directly contacting potential clients by snail mail, phone, or email for a few months after an accident. Search engine advertising prices per click are 5–7% higher in those states than in others.

By taking advantage of advertisers' desire to reach small, targeted segments of the population and varying the price according to advertisers' willingness to pay, Google is essentially perfectly price discriminating.

Transaction Costs and Perfect Price Discrimination

Although some firms come close to perfect price discrimination, many more firms set a single price or use another nonlinear pricing method. Transaction costs are a major reason why these firms do not perfectly price discriminate: It is too difficult or costly to gather information about each customer's price sensitivity. Recent advances in computer technologies, however, have lowered these costs, causing hotels, car and truck rental companies, cruise lines, and airlines to price discriminate more often.

Private colleges request and receive financial information from students, which allows the schools to nearly perfectly price discriminate. The schools give partial scholarships as a means of reducing tuition for relatively poor students.

Many other firms believe that, given the transaction costs, it pays to group price discriminate, or use nonlinear pricing rather than try to perfectly price discriminate. We now turn to these alternative approaches.

12.3 Group Price Discrimination

Most firms have no practical way to estimate the reservation price for each of their customers and to charge each customer a different price. However, many of these firms know which groups of customers are likely to have higher reservation prices on average than others. A firm engages in *group price discrimination* by dividing potential customers into two or more groups and setting different prices for each group. As with individual price discrimination, to engage in group price discrimination, a firm must have market power, be able to identify groups with different reservation prices, and prevent resale.

Consumer groups may differ by age (such as adults and children), by location (such as by country), or in other ways. All units of the good sold to customers within a group are sold at a single price.

For example, movie theaters with market power charge senior citizens a lower price than they charge younger adults because senior citizens are not willing to pay as much as others to see a movie. By admitting people as soon as they demonstrate their age and buy tickets, the theater prevents resale.

How does a monopoly set the prices for two groups of consumers?[6] To answer this question, we use our understanding of a single-price monopoly's behavior.

Suppose a monopoly sells its product in the United States and the United Kingdom and can prevent resale. Its marginal and average cost, m, is constant and identical in both countries; p_A and p_B are the American and British prices; and Q_A and Q_B are the American and British quantities. Its total profit, π, is the sum of its American and British profits, π_A and π_B:

$$\pi = \pi_A + \pi_B = (p_A Q_A - mQ_A) + (p_B Q_B - mQ_B),$$

where $p_A Q_A$ is the U.S. revenue, mQ_A is the U.S. cost, $p_B Q_B$ is the U.K. revenue, and mQ_B is the U.K. cost.

To maximize its total profit, the firm maximizes its profit in each country, setting the marginal revenue in each country equal to the marginal cost, m. This price-setting rule is profit maximizing if the firm does not want to change its price for either group. Would the monopoly want to lower its price and sell more output in the United States? If it did, its marginal revenue would be below its marginal cost, so this change would reduce its profit. Similarly, if the monopoly sold less output in the United States, its marginal revenue would be above its marginal cost, which would reduce its profit. The same argument holds for pricing in Britain.

Thus, the price-discriminating monopoly maximizes its profit by operating where its marginal revenue for each country equals its common marginal cost. Because the monopoly equates the marginal revenue for each group to its common marginal cost, m, the marginal revenues for the two countries are equal:

$$MR^A = m = MR^B. \tag{12.1}$$

The following Application illustrates this behavior.

Application

Warner Brothers Sets Prices for a *Harry Potter* DVD

A copyright gives Warner Brothers the legal monopoly to produce and sell the *Harry Potter and the Deathly Hallows, Part 2* DVD. Warner sells the movie in the United States, the United Kingdom, and other countries. Because the U.S. and U.K. DVD formats differ, Warner can charge different prices in the two countries without worrying about resale.

The DVD was released during the holiday season of 2011–2012 and sold $Q_A = 5.8$ million copies to American consumers at $p_A = \$29$ and $Q_B = 2.0$ million copies to British consumers at $p_B = \$39$ (£25). Thus, Warner used group price discrimination, charging different prices in various countries.

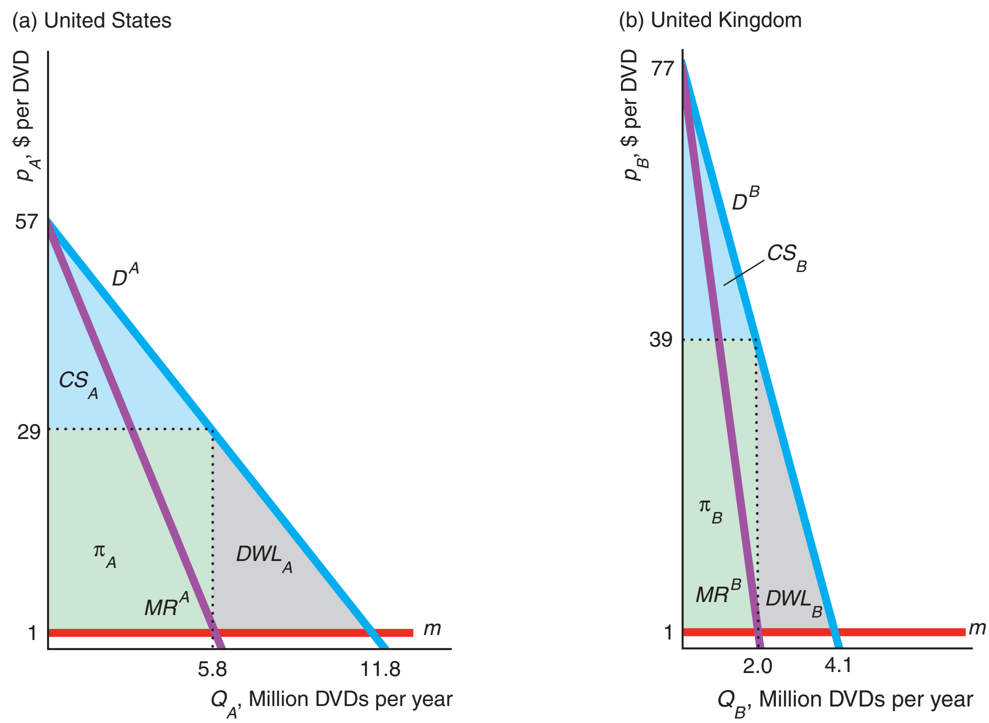
Warner Brothers had the same constant marginal cost, m, of about \$1 per unit in both countries. It set p_A and p_B to maximize its combined profit in the two countries.

Figure 12.3 shows my estimates of the linear demand curves in the two countries. In panel a, Warner Brothers maximized its U.S. profit by selling $Q_A \approx 5.8$ million DVDs, where its marginal revenue equaled its marginal cost $MR^A = m = 1$, as in Equation 12.1. It charged $p_A = \$29$. Similarly in panel b, the firm maximized its U.K. profit by selling $Q_B = 2$ million DVDs where $MR^B = m = \$1$. The corresponding price was $p_B = \$39$.

[6] See Appendix 12B for a calculus analysis.

Figure 12.3 Group Pricing of the *Harry Potter* DVD

Warner Brothers, the monopoly producer of the *Harry Potter and the Deathly Hallows, Part 2* DVD, charges more in the United Kingdom, p_B = \$39 (£25), than in the United States, p_A = \$29, because demand is more elastic in the United States. Warner Brothers sets the quantity independently in each country, where its relevant marginal revenue equals its common, constant marginal cost, m = \$1. As a result, it maximizes its profit by equating the two marginal revenues: $MR^A = 1 = MR^B$.



Solved Problem 12.2

I estimated that Warner faced inverse demand functions for its *Harry Potter and the Deathly Hallows, Part 2* DVD of $p_A = 57 - 4.8Q_A$ in the United States and $p_B = 77 - 19Q_B$ in the United Kingdom. Given that the marginal cost is 1 in both countries, solve for Warner's optimal prices and quantities in each country.

Answer

1. *Determine the marginal revenue functions.* The marginal revenue curve corresponding to a linear inverse demand curve has twice as steep a slope and the same price intercept (Chapter 11, Solved Problem 11.1). Thus, the marginal revenue function is $MR^A = 57 - 9.6Q_A$ in the United States and $MR^B = 77 - 38Q_B$ in the United Kingdom.

2. *Solve for Warner's optimal monopoly prices and quantities in each country separately.* Warner's optimal monopoly quantity is determined by equating the marginal revenue and marginal cost. The U.S. monopoly optimum condition is $57 - 9.6Q_A = 1$, so $Q_A = 56/9.6 \approx 5.833$. Using the inverse demand function, we learn that the corresponding price is $p \approx 29$. In the United Kingdom, the condition for an optimum is $77 - 38Q_B = 1$, so $Q_B = 76/38 = 2$, and $p_B = 39$.

Prices and Elasticities

In our DVD example, we can use Equation 12.1, $MR^A = m = MR^B$, to determine how the prices for the two groups vary with the price elasticities of demand at the profit-maximizing outputs. The marginal revenue for each group is a function of the corresponding price and the price elasticity of demand (as Chapter 11 showed in Equation 11.4). The U.S. marginal revenue is $MR^A = p_A(1 + 1/\ _A)$, where ${}_A$ is the price elasticity of demand for U.S. consumers, and the U.K. marginal revenue is $MR^B = p_B(1 + 1/\ _B)$, where ${}_B$ is the price elasticity of demand for British consumers.

Rewriting Equation 12.1 using these expressions for marginal revenue, we find that

$$MR^A = P_A\left(1 + \frac{1}{\ _A}\right) = m = p_B\left(1 + \frac{1}{\ _B}\right) = MR^B. \quad (12.2)$$

By rearranging Equation 12.2, we learn that the ratio of prices in the two countries depends only on demand elasticities in those countries:

$$\frac{p_B}{p_A} = \frac{1 + 1/\ _A}{1 + 1/\ _B}. \quad (12.3)$$

Solved Problem 12.3

MyLab Economics Solved Problem

What U.S. and U.K. elasticities did Warner Brothers believe it faced for its *Harry Potter and the Deathly Hallows, Part 2* DVD? Check that these elasticities are consistent with the observed price ratio, p_A/p_B.

Answer

1. *Use Equation 12.2 to express the elasticities in terms of the price and marginal cost.* According to Equation 12.2, $p_A(1 + 1/\ _A) = m = p_B(1 + 1/\ _B)$. Rearranging these expressions, ${}_A = p_A/(m - p_A)$, and ${}_B = p_B/(m - p_B)$.

2. *Use the equations we just derived to calculate the elasticities.* Given that $m = \$1$, $p_A = \$29$, and $p_B = \$39$ (see the Application "Warner Brothers Sets Prices for a *Harry Potter* DVD"), Warner must have believed that ${}_A = p_A/[m - p_A] = 29/[-28] \approx -1.0357$ and ${}_B = p_B/[m - p_B] = 39/[-38] \approx -1.0263$.

3. *Use Equation 12.3 to check that these equations are consistent with the observed prices.*

Substituting the prices and elasticities into Equation 12.3, we determine that

$$\frac{p_B}{p_A} = \frac{\$39}{\$29} \approx 1.345 \approx \frac{1 + 1/(-1.0357)}{1 + 1/(-1.0263)} = \frac{1 + 1/\ _A}{1 + 1/\ _B}.$$

Comment: Warner Brothers apparently believed that the British demand curve was less elastic at its profit-maximizing prices than the U.S. demand curve, as ${}_B \approx -1.0263$ is closer to zero than is ${}_A \approx -1.0357$. Consequently, Warner charged U.K. consumers 34% more than U.S. customers.[7]

---

[7] By mid-2012 as demand for the DVD fell, Amazon dropped the price for this DVD at its sites around the world, but maintained its price differentials. Amazon's U.S. price fell to \$7, while its U.K. price dropped to \$9.50, so that U.K. consumers paid about the same amount more than U.S. consumers, 36%.

Preventing Resale

As with all types of price discrimination, a monopoly can use group price discrimination only if it can prevent resale. In the *Harry Potter* example, Warner did not have to worry about resale between the United States and the United Kingdom because the U.S. and U.K. DVD formats differ. In many cases of international group price discrimination, the ability to prevent resale depends on trade, copyright, and patent laws. When these laws change, firms must develop other means to prevent resale, or they lose their ability to group price discriminate.

Application

Reselling Textbooks

When Supap Kirtsaeng, a Thai math student, was an undergraduate at Cornell University and then a Ph.D. student at the University of Southern California, he found a way to pay for his education. He had his friends and relatives ship him textbooks that they bought in Thailand, which he resold to U.S. college students on eBay and elsewhere, netting hundreds of thousands of dollars.

Why was reselling these books profitable? U.S. textbooks sell at much lower prices in foreign markets. Many of these books differ from their U.S. versions only by having a soft cover with an "international edition" label.

John Wiley & Sons, a publisher, sued Mr. Kirtsaeng for copyright infringement. The company claimed that by importing and selling its books, Mr. Kirtsaeng infringed the company's copyright. It asserted that the *first-sale* doctrine—which allows people who buy something to use or resell it however they want—did not apply to goods produced specifically for sale overseas.

The U.S. Court of Appeals for the 2nd Circuit in New York agreed with Wiley and upheld a $600,000 judgment against Mr. Kirtsaeng. However, in 2013, the U.S. Supreme Court reversed that ruling by a six-to-three vote, concluding that the first-sale rule holds generally. This decision also applies to records, movies, art, software, and other goods covered under copyright law.[8]

This decision could make it more difficult to maintain price differentials across countries. A possible consequence of this ruling is that poor foreign students will no longer be able to afford textbooks because the foreign price will rise. The U.S. and foreign price will differ by only the transaction cost of reselling the books. If those transaction costs are negligible, publishers will charge a single price throughout the world.

However, as of 2016, the prices of textbooks in the United States are still generally higher than in other countries (even other Western countries such as the United Kingdom). One reason why publishers are able to maintain price differentials is that they modify foreign editions to prevent reselling; however, doing so is expensive and time-consuming. Once electronic textbooks become common, students will rent the books for the term and be unable to resell them, which will facilitate price discrimination (though the prices of textbooks will fall substantially).

Solved Problem 12.4

A monopoly book publisher with a constant marginal cost (and average cost) of $MC = 1$ sells a novel in only two countries and faces a linear inverse demand curve of $p_1 = 6 - \frac{1}{2}Q_1$ in Country 1 and $p_2 = 9 - Q_2$ in Country 2. What price would a profit-maximizing monopoly charge in each country with and without a ban against shipments between the countries?

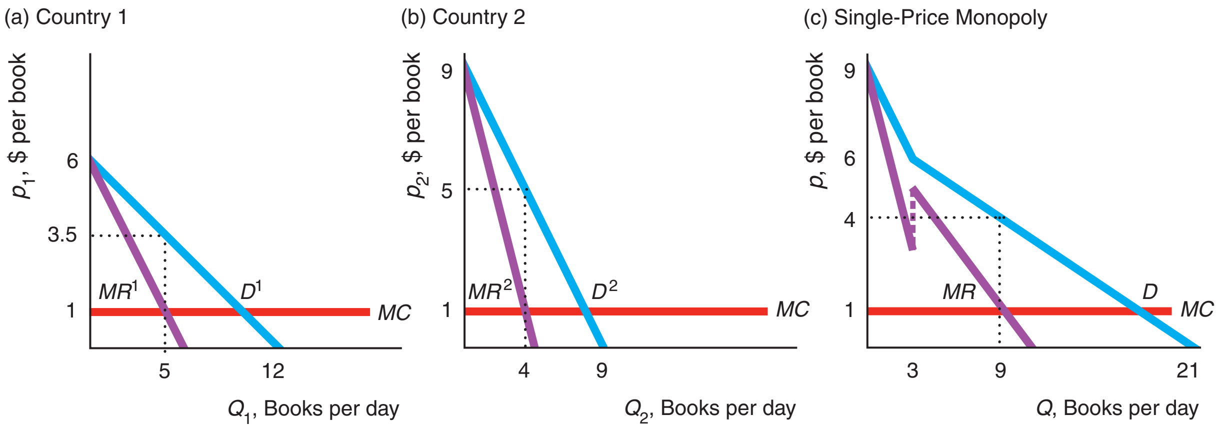
[8]However, the Supreme Court held in 2010 that Omega could prevent Costco from selling its watches produced outside the United States, citing a tiny trademark on each watch, so that they came under the jurisdiction of trademark laws, which provide owners more protection than do copyright laws.

Answer

***If resale across borders is banned so that price discrimination is possible*:**

1. *Determine the profit-maximizing price that the monopoly sets in each country by setting the relevant marginal revenue equal to the marginal cost.* If the monopoly can price discriminate, it sets a monopoly price (Section 11.1)independently in each country. The inverse demand function is $p_1 = 6 - \frac{1}{2}Q_1$, as panel a in the figure shows. The marginal revenue curve is twice as steeply sloped as is the linear inverse demand curve (Chapter 11): $MR^1 = 6 - Q_1$. The monopoly maximizes its profit where its marginal revenue equals its marginal cost,

$$MR^1 = 6 - Q_1 = 1 = MC.$$



Solving, we find that its profit-maximizing output is $Q_1 = 5$. Substituting this expression back into the monopoly's inverse demand curve, we learn that its profit-maximizing price is $p_1 = 3.5$, as panel a illustrates. In Country 2, the inverse demand curve is $p_2 = 9 - Q_2$, so the monopoly chooses Q_2 such that $MR^2 = 9 - 2Q_2 = 1 = MC$. Thus, it maximizes its profit in Country 2 where $Q_2 = 4$ and $p_2 = 5$, as panel b shows.

***If imports are permitted so that price discrimination is impossible*:**

2. *Derive the total demand curve.* If the monopoly cannot price discriminate, it charges the same price, p, in both countries. The monopoly faces the total demand curve in panel c, which is the horizontal sum of the demand curves for each of the two countries in panels a and b (Chapter 2). If the price is between 6 and 9, the quantity demanded is positive in only Country 2, so the total demand curve (panel c) is the same as Country 2's demand curve (panel b). If the price is less than 6 where both countries demand a positive quantity, the total demand curve (panel c) is the horizontal sum of the two individual countries' demand curves (panels a and b).[9] As panel c shows, the total demand curve has a kink at $p = 6$, because the quantity demanded in Country 1 is positive only below this price.

[9] Rearranging the inverse demand functions, we find that the Country 1 demand function is $Q_1 = 12 - 2p_1$ and the Country 2 demand function is $Q_2 = 9 - p_2$. As a result for prices below 6, the total demand function is $Q = (12 - 2p) + (9 - p) = 21 - 3p$, where $Q = Q_1 + Q_2$ is the total quantity that the monopoly sells in both countries.

3. *Determine the marginal revenue curve corresponding to the total demand curve.* Because the total demand curve has a kink at $p = 6$, the corresponding marginal revenue curve has two sections. At prices above 6, the marginal revenue curve is the same as that of Country 2 in panel b. At prices below 6, where the total demand curve is the horizontal sum of the two countries' demand curves, the marginal revenue curve has twice the slope of the linear total inverse demand curve. The inverse total demand function is $p = 7 - \frac{1}{3}Q$, and the marginal revenue function is $MR = 7 - \frac{2}{3}Q$.[10] Panel c shows that the marginal revenue curve *jumps*—is discontinuous—at the quantity where the total demand curve has a kink.

4. *Solve for the single-price monopoly solution.* The monopoly maximizes its profit where its marginal revenue equals its marginal cost. From inspecting panel c, we learn that the intersection occurs in the section where both countries are buying the good: $MR = 7 - \frac{2}{3}Q = 1 = MC$. Thus, the profit-maximizing output is $Q = 9$. Substituting that quantity into the inverse total demand function, we find that the monopoly charges $p = 4$. Thus, the price of the nondiscriminating monopoly, 4, lies between the two prices it would charge if it could price discriminate: $3.50 < 4 < 5$.

Identifying Groups

People often complain about airline pricing. Some argue that:

> **Common Confusion:** It doesn't make sense for airlines to charge lower prices to people who book in advance, as the airlines don't directly gain from this rule.

That statement is correct as far as it goes—the airlines don't get a large, direct benefit from having customers book a few weeks early. However, airlines do benefit indirectly. This rule allows airlines to identify groups of consumers with different elasticities and to price discriminate.

Firms use two main approaches to divide customers into groups. One method is to identify and divide consumers based on their *actions*. The firm allows consumers to self-select the group to which they belong. Airlines have discovered that people who can book in advance, such as vacationers, are likely to be more price sensitive than people who cannot book in advance, such as business travelers.

Similarly, a firm may identify customers by their willingness to spend time to buy a good at a lower price. Firms price discriminate by taking advantage of the differing values that customers place on their time. For example, in the case of customers who are willing to spend time to obtain a bargain price, firms may make people wait in line or use other time-intensive methods of selling goods. Because many high-wage people are unwilling to "waste their time shopping," store managers may run *in store* sales that allow customers who visit the store to get a low price, while customers who order by phone or over the Internet pay a higher price. This type of price discrimination increases profit if people who put a high value on their time also have a less elastic demand for the good.

---

[10] From the previous footnote, we know that the total demand function for prices less than 6 is $Q = 21 - 3p$. Rearranging this expression, we find that the inverse demand function is $p = 7 - \frac{1}{3}Q$. Because the marginal revenue function has twice as steep a slope, it is $MR = 7 - \frac{2}{3}Q$.

Yet another example concerns early adopters of a new product, who are often very enthusiastic and willing to pay premium prices. Firms can take advantage of early adopters by charging a high initial price for a new product and then lowering the price after a few weeks or months.

The second method firms use to divide buyers into groups is based on *observable characteristics* of consumers that the firm believes are associated with unusually high or low reservation prices or demand elasticities. For example, movie theaters price discriminate using the age of customers, charging higher prices for adults than for children.

Similarly, some firms charge customers in one country higher prices than those in another country. In 2015, the price of a round-trip flight between New York and London was $2,507 on average if a traveler started in New York but $1,672 if the traveler departed from London. This difference reflects group price discrimination.

Application

Buying Discounts

Firms use various approaches to induce consumers to indicate whether they have relatively high or low elasticities of demand. For each of these methods, consumers must incur some cost, such as their time, to receive a discount. Otherwise, all consumers would get the discount. By spending extra time to obtain a discount, price-sensitive consumers are able to differentiate themselves.

You've got to prove you really want a discount!

Coupons. Many firms use discount coupons to group price discriminate. Through this device, firms divide customers into two groups, charging coupon clippers less than nonclippers. Offering coupons makes sense if the people who do not clip coupons are less price sensitive on average than those who do. People who are willing to spend their time clipping coupons are able to buy cereals and other goods at lower prices than those who value their time more. According to one estimate, as much as 93% of U.S. households use coupons at least occasionally. In 2015, firms issued 286 billion "free standing insert" coupons with a face value of $515 billion. Consumers redeemed less than 1% of these coupons.

The introduction of digital coupons (for example, **EverSave.com**) has made it easier for firms to target appropriate groups, but has lowered consumers' costs of using coupons, which means that a larger share of people use them. The firm eMarketer estimated that 59% of U.S. adult Internet users redeemed online coupons in 2016.

Airline Tickets. By choosing between two different types of tickets, airline customers indicate whether they are likely to be business travelers or vacationers. Airlines give customers a choice between high-price tickets with no strings attached and low-price fares that must be purchased long in advance.

Airlines know that many business travelers have little advance warning before they book a flight and have relatively inelastic demand curves. In contrast, vacation travelers can usually plan in advance and have relatively high elasticities of demand for air travel. The airlines' rules ensure that vacationers, with relatively elastic demand, obtain low fares while most business travelers, with relatively inelastic demand, buy high-price tickets (often more than four times higher than the plan-ahead rate).

Reverse Auctions. Priceline.com and other online merchants use a name-your-own-price or "reverse" auction to identify price-sensitive customers. A customer enters a relatively low-price bid for a good or service, such as an airline ticket. Merchants decide whether to accept that bid. To prevent their less price-sensitive customers from using these methods, airlines force successful Priceline bidders to be flexible: to fly at off hours, to make one or more connections, and to accept any type of aircraft. Similarly, when bidding on groceries, a customer must list "one or two brands you like." As Jay Walker, Priceline's founder, explained, "The manufacturers would rather not give you a discount, of course, but if you prove that you're willing to switch brands, they're willing to pay to keep you."

Rebates. Why do many firms offer a rebate of, say \$5 instead of reducing the price on their product by \$5? The reason is that a consumer must incur an extra, time-consuming step to receive the rebate. Thus, only those consumers who are price sensitive or place a low value on their time will actually apply for the rebate. According to a *Consumer Reports* survey, 47% of customers always or often apply for a rebate, 23% sometimes apply, 25% never apply, and 5% responded that the question was not applicable to them.

Welfare Effects of Group Price Discrimination

Group price discrimination results in inefficient production and consumption. As a result, welfare under group price discrimination is lower than that under competition or perfect price discrimination. However, welfare may be lower or higher with group price discrimination than with a single-price monopoly.

Group Price Discrimination Versus Competition Consumer surplus is greater and more output is produced with perfect competition than with group price discrimination. In Figure 12.3, consumer surplus with group price discrimination is CS_A for American consumers in panel a and CS_B for British consumers in panel b. Under competition, consumer surplus is the area below the demand curve and above the marginal cost curve: $CS_A + \pi_A + |DWL_A|$ in panel a and $CS_B + \pi_B + |DWL_B|$ in panel b.

Thus, group price discrimination transfers some of the competitive consumer surplus, π_A and π_B, to the monopoly as additional profit and causes the deadweight loss, DWL_A and DWL_B, which is reduced consumer surplus that is simply lost or wasted. The deadweight loss is due to the group-price-discriminating monopoly charging prices above marginal cost, which results in reduced production from the optimal competitive level.

Group Price Discrimination Versus Single-Price Monopoly From theory alone, we can't tell whether welfare is higher if the monopoly uses group price discrimination or if it sets a single price. Both types of monopolies set price above marginal cost, so they produce too little relative to competition. Output may rise as the firm starts discriminating if groups that did not buy when the firm charged a single price start buying. In the movie theater example in panel b of Table 12.1, welfare is higher with discrimination than with single-price monopoly because the discriminating monopoly sells more tickets (see Solved Problem 12.1).

The closer the group-price-discriminating monopoly comes to perfectly price discriminating (by, for example, dividing its customers into many groups rather than just two), the more output it produces, which reduces the production inefficiency. However, total surplus falls if the firm switches to group price discrimination and total output falls.[11]

[11]An additional source of inefficiency is time spent by consumers trying to resell the product to high-willingness-to-pay customers or searching for low prices. These activities do not occur if everyone knows the firm sets a uniform price.

12.4 Nonlinear Price Discrimination

Many firms are unable to determine which customers have the highest reservation prices. However, such firms may know that most customers are willing to pay more for the first unit than for successive units. That is, a typical customer's demand curve is downward sloping. Such a firm can price discriminate by letting the price each customer pays vary with the number of units the customer buys. Here, the firm uses a type of *nonlinear pricing* called *second-degree price discrimination* or *quantity discrimination*.

Although the price varies with quantity, each customer faces the same nonlinear pricing schedule.[12] To use nonlinear pricing, a firm must have market power and be able to prevent customers who buy at a low price from reselling to those who would otherwise pay a high price.

A 64-ounce bottle of V8 vegetable juice sells for \$4.39 or 6.8¢ an ounce, while a 12-ounce bottle sells for \$2.79 or 23¢ an ounce. This difference in the price per ounce reflects nonlinear price discrimination unless the price difference is due to cost differences. The result of this quantity discount is that customers who make large purchases pay less per ounce than those who make small purchases.[13]

Another nonlinear pricing strategy is *block pricing*. Many utilities use block pricing schedules, by which they charge one price per unit for the first few units (a *block*) purchased and a different price per unit for subsequent blocks. Gas, electric, water, and other utility companies commonly use declining-block or increasing-block pricing.

The block-pricing utility monopoly in Figure 12.4 faces a linear demand curve for each (identical) customer. The demand curve hits the vertical axis at \$90 and the horizontal axis at 90 units. The monopoly has a constant marginal and average cost of $m = \$30$. Panel a shows how this monopoly maximizes its profit if it can quantity discriminate by setting two prices (and both prices lie on the demand curve).

The firm uses declining-block prices to maximize its profit. The monopoly charges a price of \$70 on any quantity between 1 and 20—the first block—and \$50 on any units beyond the first 20—the second block. The points that determine the blocks, \$70 and 20 units and \$50 and 40 units, lie on the demand curve. (See Appendix 12C for a mathematical analysis.)

Given each consumer's demand curve, a consumer decides to buy 40 units and pays \$1,400 (= \$70 × 20) for the first block and \$1,000 (= \$50 × 20) for the second block. The consumer gains consumer surplus equal to A on the first block and C on the second block, for a total of $A + C$. The quantity-discriminating monopoly's profit or producer surplus is area B. Society suffers a deadweight loss of $-D$ because price, \$50, is above marginal cost, \$30, on the last unit, 40, purchased.

In panel b, the firm can set only a single price. It produces where its marginal revenue equals its marginal cost, and sells 30 units at \$60 per unit. By using nonlinear price discrimination instead of setting a single price, the utility sells more units, 40 instead

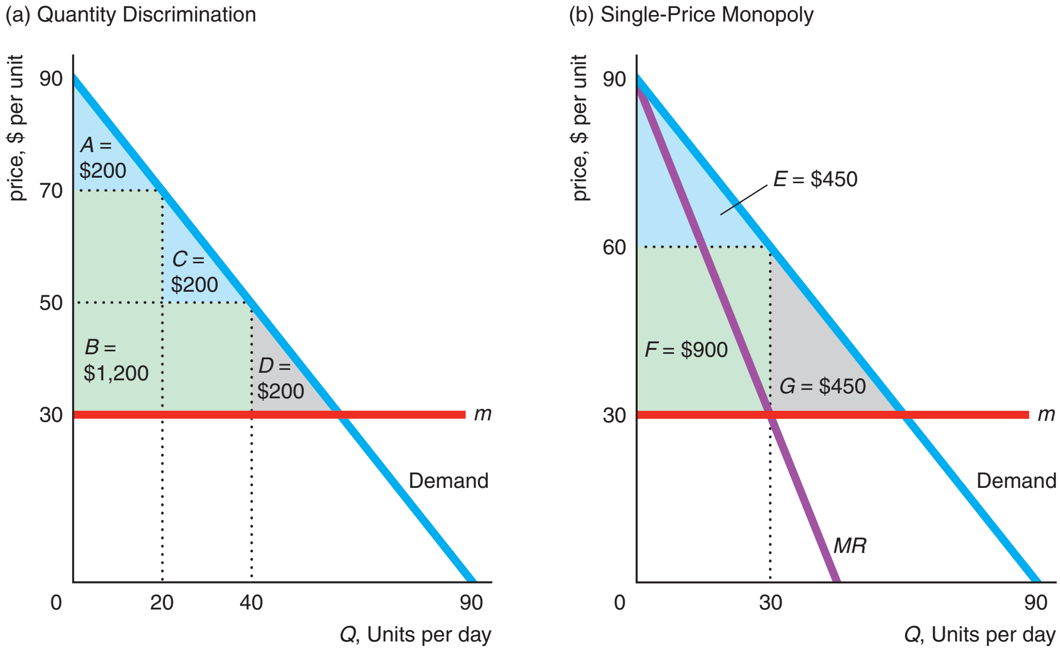
---

[12] A consumer's expenditure, pq, is a linear function of quantity, q, only if the price, p, is constant. Expenditure is *nonlinear* in quantity if the price varies with quantity: $qp(q)$.

[13] Not all quantity discounts are a form of price discrimination. Some reflect the reduction in a firm's cost with large-quantity sales. For example, the cost per ounce of selling a soft drink in a large cup is less than that of selling it in a smaller cup; the cost of cups varies little with size, and the cost of pouring and serving is the same. A restaurant offering quantity discounts on drinks may be passing on actual cost savings to larger purchasers rather than price discriminating.

Figure 12.4 Block Pricing MyLab Economics Video

If this monopoly engages in block pricing with quantity discounting, it makes a larger profit (producer surplus) than it does if it sets a single price, and welfare is greater. (a) With block pricing, its profit is $B = \$1{,}200$, welfare is $A + B + C = \$1{,}600$, and the deadweight loss is $D = \$200$. (b) If the monopoly sets a single price (so that its marginal revenue equals its marginal cost), the monopoly's profit is $F = \$900$, welfare is $E + F = \$1{,}350$, and the deadweight loss is $G = \$450$.



	Block Pricing	Single Price
Consumer Surplus, *CS*	$A + C = \$400$	$E = \$450$
Producer Surplus or Profit, $PS = \pi$	$B = \$1{,}200$	$F = \$900$
Welfare, $W = CS + PS$	$A + B + C = \$1{,}600$	$E + F = \$1{,}350$
Deadweight Loss, *DWL*	$-D = -\$200$	$-G = -\$450$

of 30, and makes a larger profit, $B = \$1{,}200$ instead of $F = \$900$. With quantity discounting, consumer surplus is lower, $A + C = \$400$ instead of $E = \$450$; total surplus (consumer surplus plus producer surplus) is higher, $A + B + C = \$1{,}600$ instead of $E + F = \$1{,}350$; and deadweight loss is lower, $-D = -\$200$ instead of $-G = -\$450$. Thus, in this example, the firm and society are better off with nonlinear price discrimination, but consumers as a group suffer.

The more block prices that a firm can set, the closer the firm gets to perfect price discrimination, where it captures all the potential consumer surplus, and its profit or producer surplus equals total surplus. Moreover, because the last unit sells at a price equal to marginal cost, total surplus is maximized and society suffers no deadweight loss.

12.5 Two-Part Pricing

We now turn to another form of nonuniform pricing, *two-part pricing*. It is similar to nonlinear price discrimination in that the average price per unit paid by a consumer varies with the number of units purchased by that consumer.

two-part pricing
charging each consumer a lump-sum *access fee* for the right to buy as many units of the good as the consumer wants at a per-unit *price*

With **two-part pricing**, the firm charges each consumer a lump-sum *access fee* for the right to buy as many units of the good as the consumer wants at a per-unit *price*.[14] Consumers pay an access fee, A, for the right to buy any number of goods and a per-unit price of p. So the total expenditure for a consumer who buys q units is $A + pq$. Because of the access fee, the average amount per unit that consumers pay is greater if they buy a small number of units than if they buy a larger number.

Two-part pricing is commonly used.[15] Many fitness clubs charge a yearly access fee and a price per session. Many warehouse stores require that customers buy an annual membership for the right to buy goods at relatively low prices. Some car rental firms charge a rental or access fee for the day and an additional price per mile driven. To buy 2015–2016 season tickets to the Dallas Cowboys football games (at a price from \$690 to \$1,250), a fan first must buy a *personal seat license* (PSL), giving the fan the right to buy season tickets for the next 30 years. Most PSLs sell for between \$10,000 and \$125,000.

To profit from two-part pricing, a firm must have market power, know how demand differs across customers or with the quantity that a single customer buys, and successfully prevent resale. We start by examining a firm's two-part pricing problem in the extreme case in which all customers have the same demand curve. We then consider what happens when the demand curves of individual customers differ.

Two-Part Pricing with Identical Customers

If all its customers are identical, a monopoly that knows its customers' demand curve can set a two-part price that possesses the same two important properties as perfect price discrimination. First, the quantity is efficient because the price of the last unit equals marginal cost. Second, all potential consumer surplus is transferred from consumers to the firm.

To illustrate these points we consider a monopoly that has a constant marginal cost of $MC = 10$ and no fixed cost, so its average cost is also constant at 10. All of the monopoly's customers have the same demand curve, $Q = 80 - p$. Panel a of Figure 12.5 shows the demand curve, D^1, of one such customer, Valerie.

If the monopoly sets its price, p, equal to its constant marginal cost of 10, total surplus is maximized. The firm breaks even on each unit sold and has no producer surplus and no profit. Valerie buys $q = 70$ units. Her consumer surplus is area $A = \frac{1}{2}(80 - p)q = \frac{1}{2}([80 - 10] \times 70) = 2{,}450$.

However, if the firm also charges a lump-sum access fee of 2,450, it captures this 2,450 as its producer surplus or its profit per customer, and leaves Valerie with no consumer surplus. The firm's total profit is 2,450 times the number of identical customers.

The firm maximizes its profit by setting its price equal to its marginal cost and charging an access fee that captures the entire potential consumer surplus. If the firm were to charge a price above its marginal cost of 10, it would sell fewer units and make a smaller profit.

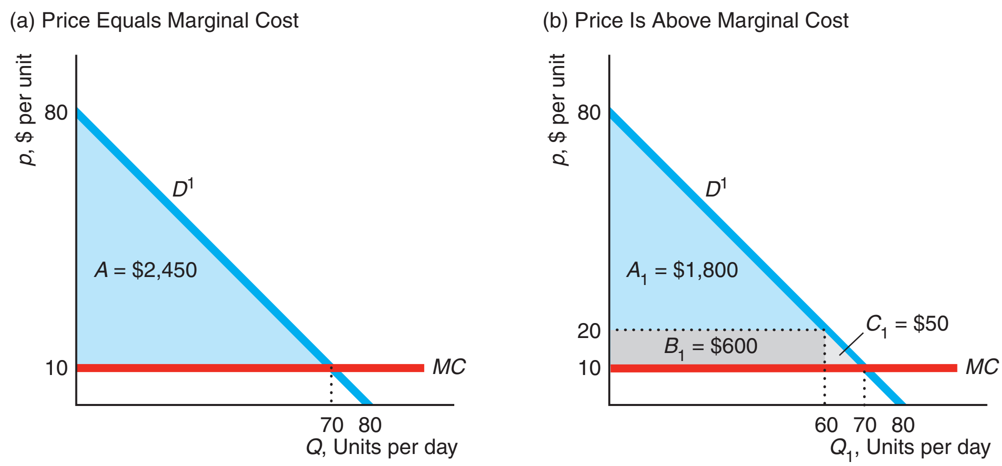
---

[14]*Jargon alert*: Many people refer to two-part pricing as *two-part tariffs*.

[15]For example, *venting* stores are springing up in shopping malls in China. A customer pays to enter, and then pays for each second-hand mobile phone, television set, or other product that the customer smashes.

Figure 12.5 Two-Part Pricing with Identical Consumers MyLab Economics Video

(a) Because all customers have the same individual demand curve as Valerie, D^1, the monopoly captures the entire potential consumer surplus using two-part pricing. The monopoly charges a per-unit fee price, p, equal to the marginal cost of 10, and a lump-sum access fee, $A = 2{,}450$, which is the blue triangle under the demand curve and above the per-unit price of $p = 10$. (b) Were the monopoly to set a price at 20, which is above its marginal cost, it would earn less. It makes a profit of $B_1 = 600$ from the 10 it earns on each of the 60 units that Valerie buys at this higher price. However, the largest access fee the firm can charge now is $A_1 = 1{,}800$, so its total profit is 2,400, which is less than the 2,450 it makes if it sets its price equal to marginal cost. The difference is a deadweight loss of $C_1 = 50$, which is due to fewer units being sold at the higher price.



For example, the firm charges $p = 20$ in panel b of Figure 12.5. At that higher price, Valerie buys only 60 units, which is less than the 70 units that she buys at a price of 10 in panel a. The firm's profit from selling these 60 units is $B_1 = (20 - 10) \times 60 = 600$.

For Valerie to agree to buy any units, the monopoly must lower its access fee to 1,800 $(= \frac{1}{2} \times 60 \times 60)$, the new potential consumer surplus, area A_1. The firm's total profit from Valerie is $A_1 + B_1 = 1{,}800 + 600 = 2{,}400$. This amount is less than the 2,450 ($= A$ in panel a) profit the firm earns if it sets price equal to marginal cost, 10, and charges the higher access fee.

Area A in panel a equals $A_1 + B_1 + C_1$ in panel b. By charging a price above marginal cost, the firm loses C_1, which is the deadweight loss due to selling fewer units.

Similarly, if the firm were to charge a price below its marginal cost, it would also earn less profit. It would sell too many units and make a loss on each unit that a higher access fee could not fully recapture.

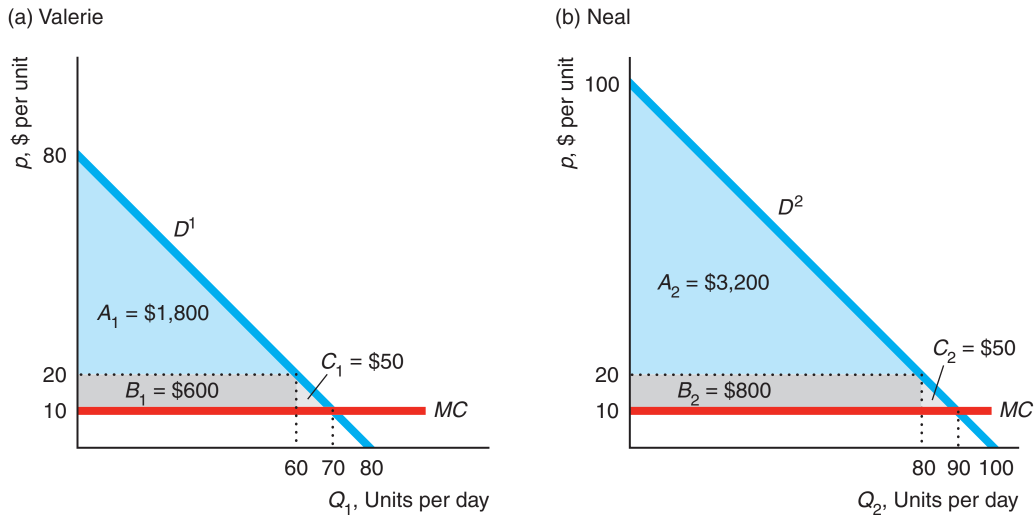
Two-Part Pricing with Nonidentical Consumers

Two-part pricing is more complex if consumers have different demand curves.[16] Suppose that the monopoly has two customers, Valerie, Consumer 1, and Neal, Consumer 2. Valerie's demand curve, $Q_1 = 80 - p$, is D^1 in panel a of Figure 12.6 (which is the same as panel b of Figure 12.5), and Neal's demand curve, $Q_2 = 100 - p$, is D^2 in panel b. The monopoly's marginal cost, MC, and average cost are constant at 10 per unit.

[16]See Appendix 12D for a calculus analysis.

Figure 12.6 Two-Part Pricing with Different Consumers

The monopoly faces two consumers. Valerie's demand curve is D^1 in panel a, and Neal's demand curve is D^2 in panel b. If the monopoly can set different prices and access fees for its two customers, it charges both a per-unit price of $p = 10$, which equals its marginal cost, and it charges an access fee of 2,450 ($= A_1 + B_1 + C_1$) to Valerie and 4,050 ($= A_2 + B_2 + C_2$) to Neal. If the monopoly cannot charge its customers different prices, it sets its per-unit price at $p = 20$, where Valerie purchases 60 and Neal buys 80 units. The firm charges both the same access fee of $1{,}800 = A_1$, which is Valerie's potential consumer surplus. The highest access fee that the firm could charge and have Neal buy is 3,200, but at that level, Valerie would not buy. By charging a price above its marginal cost, the firm captures $B_1 = 600$ from Valerie and $B_2 = 800$ from Neal. Thus, its total profit is 5,000 ($= [2 \times 1{,}800] + 600 + 800$), which is less than the 6,500 ($= 2{,}450 + 4{,}050$) it makes if it can charge separate access fees to each customer.



If the firm knows each customer's demand curve, can prevent resale, and can charge its customers different prices and access fees, it can capture the entire potential consumer surplus. The monopoly sets its price for both customers at $p = m = 10$ and sets its access fee equal to each customer's potential consumer surplus. At $p = 10$, Valerie buys 70 units (panel a), and Neal buys 90 units (panel b). If no access fee were charged, Valerie's consumer surplus, $CS_1 = \frac{1}{2}(80 - p)q_1 = \frac{1}{2}(80 - p)^2$, would equal the triangle below her demand curve and above the price line at 10, $A_1 + B_1 + C_1$, which is 2,450 ($= \frac{1}{2} \times 70 \times 70$). Similarly, Neal's consumer surplus, $CS_2 = \frac{1}{2}(100 - p)^2$, would be 4,050 ($= \frac{1}{2} \times 90 \times 90$), which is the triangle $A_2 + B_2 + C_2$.

Thus, the monopoly charges Valerie a lump-sum access fee of 2,450 and Neal an access fee of 4,050, so that the customers receive no consumer surplus. The firm's total profit, the sum of these access fees, $2{,}450 + 4{,}050 = 6{,}500$, is the maximum possible profit, because the monopoly has captured the maximum potential consumer surplus from both customers.

Now suppose that the firm cannot charge its customers different prices or access fees. The firm maximizes its profit by setting a price of 20, which exceeds its marginal cost, and collecting an access fee equal to Valerie's potential consumer surplus, $A_1 = 1{,}800$. Although this access fee captures all of Valerie's potential consumer surplus, it is less than Neal's potential consumer surplus, $3{,}200 = A_2$. Were the firm to charge an access fee of 3,200, it would sell only to Neal and make less money ($3{,}200 < 1{,}800 + 1{,}800 = 3{,}600$).

At $p = 20$, Valerie buys 60 units, and Neal buys 80 units. Because the firm's average cost is 10, the firm makes $20 - 10 = 10$ per unit, so it earns $B_1 = 600$ ($= 10 \times 60$) from Valerie and $B_2 = 800$ ($= 10 \times 80$) from Neal for a total $= 1,400$. Adding that to what it makes from the access fees, 3,600, the monopoly's total profit is 5,000 ($= [2 \times 1,800] + 600 + 800$). Valerie receives no consumer surplus, but Neal enjoys a consumer surplus of 1,400 ($= 3,200 - 1,800$).

This 5,000 profit obtained from pure two-part pricing is less than the 6,500 that it could obtain if it could set different access fees for each customer. On the other hand, its profit from pure two-part pricing exceeds the 3,200 profit that the firm could earn from uniform monopoly pricing.[17]

Why does the monopoly charge a price above marginal cost when using two-part pricing? By raising its price, the monopoly earns more per unit from both types of customers but lowers its customers' potential consumer surplus. Thus, if the monopoly can capture each customer's potential surplus by charging different lump-sum fees, it sets its price equal to marginal cost. However, if the monopoly cannot capture all the potential consumer surplus because it must charge everyone the same lump-sum fee, the increase in profit from Neal due to the higher price more than offsets the reduction in the lump-sum fee from Valerie.[18] Thus, a price above marginal cost maximizes profit in this case.

Application

iTunes for a Song

Prior to 2009, Apple's iTunes music store, the giant of music downloading, used *uniform pricing*, where it sold songs at 99¢ each. However, some of its competitors, such as Amazon MP3, did not use uniform pricing. Some record labels told Apple that they would not renew their contracts if Apple continued to use uniform pricing. Apparently responding to this pressure and the success of some of its competitors, Apple switched in 2009 to selling each song at one of three prices.

Did Apple's one-price-for-all-songs policy cost it substantial potential profit? How do consumer surplus and deadweight loss vary with pricing methods such as a single price, song-specific prices, price discrimination, and two-part pricing? To answer such questions, Shiller and Waldfogel (2011) surveyed nearly 1,000 students and determined each person's willingness to pay for each of 50 popular songs. Then they used this information to calculate optimal pricing under various pricing schemes.

First, under uniform pricing, Apple charges the same price for every song. Second, under variable pricing, Apple sells each song at its individual profit-maximizing price. Third, Apple uses two-part pricing, charging a monthly or annual fee for access and then a fixed price for each download.

---

[17] A single-price monopoly faces an aggregate demand function of the sum of the two individual demand functions: $Q = q_1 + q_2 = (80 - p) + (100 - p)$ or $Q = 180 - 2p$, for p less than 80, where both consumers demand a positive quantity. Its inverse demand function is $p(Q) = 90 - \frac{1}{2}Q$. Its revenue function is $R(Q) = p(Q) \times Q = 90Q - \frac{1}{2}Q^2$, so its marginal revenue function is $MR = \mathrm{d}R(Q)/\mathrm{d}Q = 90 - Q$. To maximize its profit given that it sets a uniform price, the monopoly equates its MR and its MC, so that $90 - Q = 10$, or $Q = 80$. At that quantity, the price is $p = 90 - (80/2) = 50$. The firm's profit is $\pi = (p - AC)Q = (50 - 10) \times 80 = 3,200$.

[18] If the monopoly lowers its price from 20 to the marginal cost of 10, it loses B_1 from Valerie, but it can raise its access fee from A_1 to $A_1 + B_1 + C_1$, so its total profit from Valerie increases by $C_1 = 50$. The access fee it collects from Neal also rises by $B_1 + C_1 = 650$, but its profit from unit sales falls by $B_2 = 800$, so its total profit decreases by 150. The loss from Neal, -150, more than offsets the gain from Valerie, 50. Thus, the monopoly makes 100 more by charging a price of 20 rather than 10.

If we know the demand curve and the marginal cost, we can determine the producer surplus (*PS*) or profit, the consumer surplus (*CS*), and the deadweight loss (*DWL*) from each pricing regime. By dividing each of these surplus measures by the total available surplus—the area under the demand curve and above the marginal cost curve—we can determine the shares of *PS*, *CS*, and *DWL* (which add to 100%). The following table shows Shiller and Waldfogel's estimates of the percentage shares of *CS*, *PS*, and the absolute value of the deadweight loss, $|DWL|$, under each of the three pricing methods:

Pricing	*PS*	*CS*	$\|DWL\|$
Uniform	28	42	29
Variable	29	45	26
Two-part pricing	37	43	20

If these students have tastes similar to those of the general market, then Apple raised its profit by switching from uniform pricing to variable pricing (see the *PS* column in the table). However, these results suggest that it could do even better using two-part pricing. Perhaps in response to this opportunity, Apple added iTunes Match (2011) and Apple Music (2015), which effectively use two-part pricing. Deadweight loss decreases under either of the alternatives to uniform pricing. Consumers do best with variable pricing, but two-part pricing is also better for consumers than uniform pricing.

12.6 Tie-In Sales

tie-in sale
allowing customers to buy one product only if they agree to purchase another product as well

Another type of nonlinear pricing is a **tie-in sale**, by which customers can purchase one product only if they agree to buy another product as well. The two forms of tie-in sales are a *requirement tie-in sale* and *bundling*.

Requirement Tie-In Sale

requirement tie-in sale
requiring customers who buy one product to make all their purchases of another product from that firm

With a **requirement tie-in sale**, customers who buy one product from a firm are required to make all their purchases of another product from that firm. Some firms sell durable machines such as copiers under the condition that customers buy copier services and supplies from them in the future. This requirement allows the firm to identify heavier users and charge them more per unit. For example, if a printer manufacturer can require that consumers buy their ink cartridges only from the manufacturer, then that firm can capture most of the consumers' surplus. Heavy users of the printer, who presumably have a less elastic demand for it, pay the firm more than light users because of the high cost of the ink cartridges.

Application

Ties That Bind

Unfortunately for printer manufacturers, the Magnuson-Moss Warranty Improvement Act forbids a manufacturer from using such tie-in provisions as a condition of warranty. To get around this Act, printer firms such as Brother, Canon, Epson, and Hewlett-Packard (HP) write their warranties to induce consumers to use only their cartridges and not to refill them. The warranty for an HP inkjet printer says that it does not apply if printer failure or damage is attributable to a non-HP or refilled cartridge.

Is this warning sufficient to induce most consumers to buy cartridges only from HP? Apparently so. HP sells its Deskjet 1112 printer for only $29.99 (with free shipping). That is, HP is virtually giving away an impressive machine that will print up to 4800 × 1200 optimized dots per inch in color. However, HP charges $21.99 for its color cartridge (rated for 165 pages) and $15.99 for its black cartridge (rated for 180 pages). If most customers bought inexpensive cartridges or refills from other firms, HP would not sell its printer at a rock-bottom price. Thus, HP has achieved the benefits of requirement tie-in sales through a carefully worded warranty.

Bundling

bundling (*package tie-in sale*)
selling multiple goods or services for a single price

A firm engages in **bundling** (or a *package tie-in sale*) by selling multiple goods or services for a single price. Indeed, most goods are bundles of many separate parts. Cars come assembled. Shoe companies such as Nike, Clarks, and Merrell sell left and right shoes together as a pair and include laces.

Usually goods are bundled for efficiency because combining goods in a bundle reduces the transaction costs incurred by consumers or the production costs associated with the product. For example, we buy shirts with buttons already attached. Rather than buying shirts without buttons, and then buying buttons, consumers prefer to buy assembled shirts, eliminating the need to make two separate purchases and then sew on buttons.

However, firms sometimes bundle even when they gain no production advantages and transaction costs are small. Bundling allows firms to increase their profit by charging different prices to different consumers based on the consumers' willingness to pay. For example, a computer firm may sell a package including a computer and a printer for a single price even if it has no cost savings from selling these products together.

A firm that sells two or more goods may sell the goods together in a *bundle* to raise its profit. In *pure bundling*, a firm only sells the goods together. For example, a restaurant may offer a soup and sandwich special but not allow customers to purchase the soup or the sandwich separately. In *mixed bundling*, the firm offers consumers the choice of buying the goods separately or as a bundle. A restaurant may offer the soup and sandwich special as well as sell each item separately.

Bundling allows firms that can't directly price discriminate to charge customers different prices. Whether either type of bundling is profitable depends on customers' tastes and the ability to prevent resale.

Pure bundling is very common. An example of a pure bundle is Microsoft Works. The primary components of this software bundle are a word processing program and a spreadsheet program. These programs have fewer features than Microsoft's flagship Word and Excel programs and are not sold individually but only as a bundle.

Whether it pays for Microsoft to sell a bundle or sell the programs separately depends on how reservation prices for the components vary across customers. We use an example of a firm selling word processing and spreadsheet programs to illustrate two cases, one in which pure bundling produces a higher profit than selling the components separately, and one in which pure bundling is not profitable.

The firm has two customers, Alisha and Bob. The first two columns of Table 12.2 show the reservation prices for each consumer for the two products. Alisha's

Table 12.2 Negatively Correlated Reservation Prices

	Word Processor	*Spreadsheet*	*Bundle*
Alisha	$120	$50	$170
Bob	$90	$70	$160
Profit-maximizing price	$90	$50	$160
Units sold	2	2	2

reservation price for the word processing program, $120, is greater than Bob's, $90. However, Alisha's reservation price for the spreadsheet program, $50, is less than Bob's, $70. The reservation prices are *negatively correlated*: The customer who has the higher reservation price for one product has the lower reservation price for the other product. The third column of the table shows each consumer's reservation price for the bundle, which is the sum of the reservation prices for the two underlying products.

If the firm sells the two products separately, it maximizes its profit by charging $90 for the word processor and selling to both consumers, so that its profit is $180, rather than charging $120 and selling only to Alisha. If it charges between $90 and $120, it still only sells to Alisha and earns less than if it charges $120. Similarly, the firm maximizes its profit by selling the spreadsheet program for $50 to both consumers, earning $100, rather than charging $70 and selling to only Bob. The firm's total profit from selling the programs separately is $280 (= $180 + $100).

If the firm sells the two products in a bundle, it maximizes its profit by charging $160, selling to both customers, and earning $320. This is a better outcome than charging $170 and selling only to Alisha. Pure bundling is more profitable for the firm because it earns $320 from selling the bundle and only $280 from selling the programs separately.

Pure bundling is more profitable because the firm captures more of the consumers' potential consumer surplus—their reservation prices. With separate prices, Alisha has consumer surplus of $30 (= $120 − $90) from the word processing program and none from the spreadsheet program. Bob receives no consumer surplus from the word processing program and $20 from the spreadsheet program. Thus, the total consumer surplus is $50. With pure bundling, Alisha gets $10 of consumer surplus and Bob gets none, so the total is only $10. Thus, the pure bundling approach captures $40 more potential consumer surplus than does pricing separately.

Whether pure bundling increases the firm's profit depends on the reservation prices. Table 12.3 shows the reservation prices for two different consumers, Carol and Dmitri. Carol has higher reservation prices for both products than does Dmitri. These reservation prices are *positively correlated*: A higher reservation price for one product is associated with a higher reservation price for the other product.

If the programs are sold separately, the firm charges $90 for the word processor, sells to both consumers, and earns $180. However, it makes more charging $90 for

Table 12.3 Positively Correlated Reservation Prices

	Word Processor	*Spreadsheet*	*Bundle*
Carol	$100	$90	$190
Dmitri	$90	$40	$130
Profit-maximizing price	$90	$90	$130
Units sold	2	1	2

the spreadsheet program and selling only to Carol, than it does charging $40 for the spreadsheet, selling to both consumers, and earning $80. The firm's total profit if it prices separately is $270 (= $180 + $90).

If the firm uses pure bundling, it maximizes its profit by charging $130 for the bundle, selling to both customers, and making $260. Because the firm earns more selling the programs separately, $270, than when it bundles them, $260, pure bundling is not profitable in this example. Even if Dmitri placed a higher value on the spreadsheet, as long as reservation prices are positively correlated, pure bundling cannot increase the profit.

Solved Problem 12.5

A firm that sells word processing and spreadsheet programs has four potential customers with the following reservation prices:

	Word Processor	*Spreadsheet*	*Bundle*
Aaron	$120	$30	$150
Brigitte	$110	$90	$200
Charles	$90	$110	$200
Dorothy	$30	$120	$150

The firm's cost of production is zero, so maximizing its profit is equivalent to maximizing its revenue. To maximize its profit, should the firm charge separate prices for each product, engage in pure bundling, or use mixed bundling?

Answer

1. *Calculate the profit-maximizing separate service prices and the resulting profit.* If the firm prices each program separately, it maximizes its profit by charging $90 for each product and selling each to three out of the four potential customers. It sells the word processing program to Aaron, Brigitte, and Charles. It sells the spreadsheet program to Brigitte, Charles, and Dorothy. Thus, it makes $270 (= 3 × $90) from each program or $540 total, which exceeds what it could earn by setting any other price per program.[19]

2. *Calculate the profit-maximizing pure bundle price and the resulting profit.* The firm can charge $150 for the bundle, sell to all four consumers, and make a profit of $600, $60 more than the $540 it makes from selling the programs separately.

3. *Determine how the firm maximizes its profit by using mixed bundling.* With mixed bundling, the firm charges $200 for the bundle and $120 for each product separately. The firm earns $400 from Brigitte and Charles, who buy the bundle. Aaron buys only the word processing program for $120, and Dorothy buys only the spreadsheet for another $120, so that the firm makes $240 from its individual program sales. Thus, its profit is $640 (= $400 + $240) from mixed bundling, which exceeds the $600 from pure bundling, and the $540 from individual sales.

[19]If it sets a price of a program as low as $30, it sells both programs to all four customers, but makes only $240. If it charges $110 it sells each program to two customers and earns $440. If it charges $120, it makes a single sale of each program, so it earns $240.

12.7 Advertising

The man who stops advertising to save money is like the man who stops the clock to save time.

In addition to setting prices or quantities, choosing investments, and lobbying governments, firms engage in many other strategic actions to boost their profits. One of the most important is advertising. By advertising, a monopoly can shift its demand curve, which may allow it to sell more units at a higher price. In contrast, a competitive firm has no incentive to advertise as it can sell as many units as it wants at the going price without advertising.

Advertising is only one way to promote a product. Other promotional activities include providing free samples and using sales agents. Some promotional tactics are subtle. For example, grocery stores place sugary breakfast cereals on lower shelves so that they are at children's eye level. According to a survey of 27 supermarkets nationwide by the Center for Science in the Public Interest, the average position of 10 child-appealing brands (44% sugar) was on the next-to-bottom shelf, while the average position of 10 adult brands (10% sugar) was on the next-to-top shelf.

A monopoly advertises to raise its profit. A successful advertising campaign shifts the market demand curve by changing consumers' tastes or informing them about new products. The monopoly may be able to change the tastes of some consumers by telling them that a famous athlete or performer uses the product. Children and teenagers are frequently the targets of such advertising. If the advertising convinces some consumers that they can't live without the product, the monopoly's demand curve may shift outward and become less elastic at the new optimum, at which the firm charges a higher price for its product (see Chapter 11).

If a firm informs potential consumers about a new use for the product, the demand curve shifts to the right. For example, a 1927 Heinz advertisement suggested that putting its baked beans on toast was a good way to eat beans for breakfast as well as dinner. By so doing, it created a British national dish and shifted the demand curve for its product to the right.

The Decision Whether to Advertise

Even if advertising succeeds in shifting demand, it may not pay for the firm to advertise. If advertising shifts demand outward or makes it less elastic, the firm's *gross profit*, ignoring the cost of advertising, must rise. The firm undertakes this advertising campaign, however, only if it expects its *net profit* (gross profit minus the cost of advertising) to increase.

If the monopoly does not advertise, it faces the demand curve D^1 in Figure 12.7. If it advertises, its demand curve shifts from D^1 to D^2.

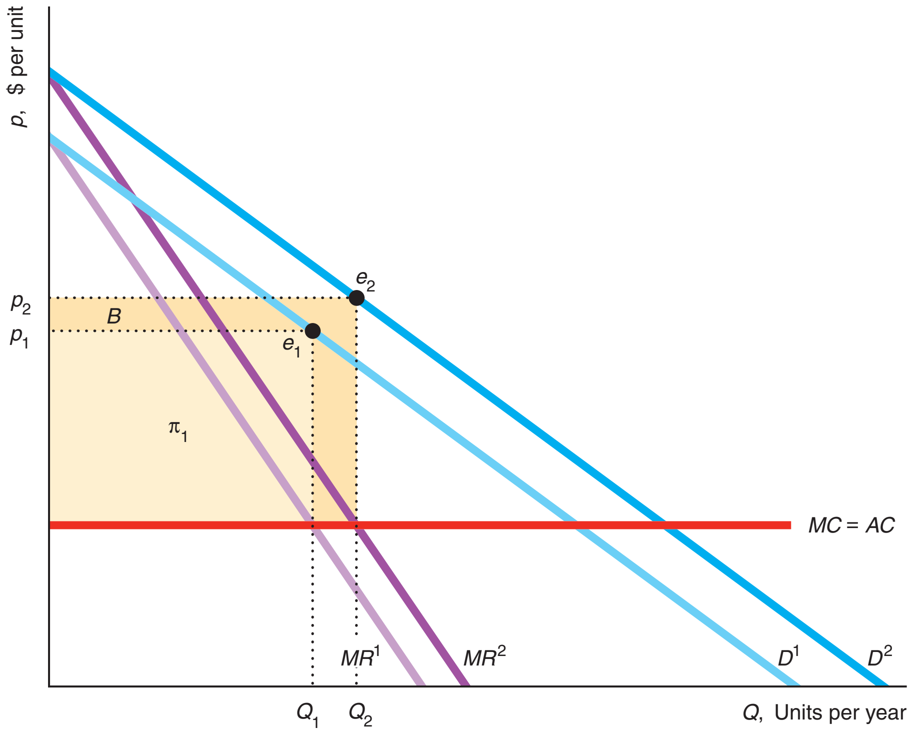
The monopoly's marginal cost, MC, is constant and equals its average cost, AC. Before advertising, the monopoly chooses its output, Q_1, where its marginal cost hits its marginal revenue curve, MR^1, which corresponds to the demand curve, D^1. The profit-maximizing optimum is e_1, and the monopoly charges a price of p_1. The monopoly's profit, π_1, is a box whose height is the difference between the price and the average cost and whose length is the quantity, Q_1.

After its advertising campaign shifts its demand curve to D^2, the monopoly chooses a higher quantity, Q_2 ($> Q_1$), where the MR^2 and MC curves intersect. In this new optimum, e_2, the monopoly charges p_2. Despite this higher price, the monopoly sells more units after advertising because of the outward shift of its demand curve.

Consequently, the monopoly's gross profit rises. Its new gross profit is the rectangle $\pi_1 + B$, where the height of the rectangle is the new price minus the average

Figure 12.7 Advertising MyLab Economics Video

If the monopoly does not advertise, its demand curve is D^1. At its actual level of advertising, its demand curve is D^2. Advertising increases the monopoly's gross profit (ignoring the cost of advertising) from π_1 to $\pi_2 = \pi_1 + B$. Thus, if the cost of advertising is less than the benefits from advertising, B, the monopoly's net profit (gross profit minus the cost of advertising) rises.



cost, and the length is the quantity, Q_2. Thus, the benefit, B, to the monopoly from advertising at this level is the increase in its gross profit. If its cost of advertising is less than B, its net profit rises, and it pays for the monopoly to advertise at this level rather than not to advertise at all.

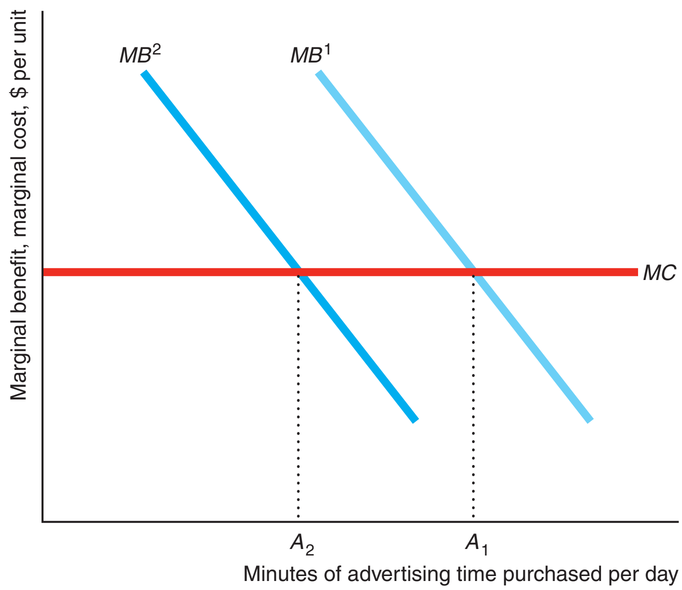
How Much to Advertise

How much should a monopoly advertise to maximize its net profit? To answer this question, we consider what happens if the monopoly raises or lowers its advertising expenditures by $1, which is its marginal cost of an additional unit of advertising. If a monopoly spends one more dollar on advertising and its gross profit rises by more than $1, its net profit rises, so the extra advertising pays. In contrast, the monopoly should reduce its advertising if the last dollar of advertising raises its gross profit by less than $1, so its net profit falls. Thus, the monopoly's level of advertising maximizes its net profit if the last dollar of advertising increases its gross profit by $1 (see Appendix 12E for an analysis using calculus). In short, the rule for setting the profit-maximizing amount of advertising is the same as that for setting the profit-maximizing amount of output: Set advertising or quantity where the marginal benefit (the extra gross profit from one more unit of advertising or the marginal revenue from one more unit of output) equals its marginal cost.

We can illustrate how firms use such marginal analysis to determine how much time to purchase from television stations for infomercials, those interminably long television advertisements sometimes featuring unique (and typically bizarre) plastic products: "Isn't that amazing?! It slices! It dices! . . . But wait! That's not all!" As Figure 12.8 shows, the marginal cost per minute of broadcast time, MC, on small television stations is constant. The firm buys A_1 minutes of advertising time, where its marginal benefit, MB^1, equals its marginal cost. If an event occurs that shifts down the marginal benefit curve to MB^2 (e.g., some regular viewers watch the Super Bowl or the World Cup instead of infomercials), the amount of advertising falls to A_2.

Figure 12.8 Shift in the Marginal Benefit of Advertising

If the marginal benefit of advertising curve is MB^1, a firm purchases A_1 minutes of infomercials, where MB^1 intersects the marginal cost (per minute of broadcast time) curve, MC. If a special event causes regular viewers to watch another show instead of infomercials so that the marginal benefit curve shifts to the left to MB^2, only A_2 minutes of advertising time is sold.



Application

Super Bowl Commercials

Super Bowl commercials are the most expensive commercials on U.S. television. A 30-second spot during the Super Bowl cost $5 million in 2016. The high price for these commercials is not surprising because the cost of commercials generally increases with the number of viewers (*eyeballs* in industry jargon), and the Super Bowl is the most widely watched show, with over 112 million viewers in 2016. What is surprising is that Super Bowl advertising costs 2.5 times as much per viewer as other TV commercials.

However, a Super Bowl commercial is much more likely to influence viewers than commercials on other shows. The Super Bowl is not only a premier sports event, it showcases the most memorable commercials of the year. People still discuss Apple's classic 1984 Macintosh ad, which is rebroadcast annually. Indeed, many

Super Bowl viewers are not even football fans—they watch to see these superior ads. Moreover, Super Bowl commercials receive extra exposure because these ads often *go viral* on the Internet.

Given that viewers are more likely to remember Super Bowl ads, are these commercials worth the extra price? Obviously many advertisers believe so, as their demand for these ads has bid up the price. Kim (2011) found that immediately after a Super Bowl commercial airs, the advertising firm's stock value rises. Thus, investors apparently believe that Super Bowl commercials raise a firm's profits despite the high cost of the commercial. Ho et al. (2009) found that, for the typical movie with a substantial advertising budget, a Super Bowl commercial advertising the movie raises theater revenues by more than the same expenditure on other television advertising. They concluded that movie firms' advertising during the Super Bowl was at (or close to) the profit-maximizing amount.

Challenge Solution

Sale Prices

By putting Heinz Ketchup on sale periodically, Heinz can price discriminate. How often should Heinz put its ketchup on sale? Under what conditions does Heinz make more money by having sales?

To answer these questions, we study a simplified market in which Heinz competes with one other ketchup brand, which we refer to as generic ketchup.[20] Every n days, the typical consumer buys either Heinz or generic ketchup.

Switchers are price sensitive and buy the least expensive ketchup. They pay attention to price information and always know when Heinz is on sale.

Heinz considers holding periodic sales to capture switchers' purchases. The generic is sold at a competitive price equal to its marginal cost of production of \$2.01 per unit. Suppose that Heinz's marginal cost is $MC = \$1$ per unit (due to its large scale) and that, if it only sold to its loyal customers, it would charge a monopoly price of $p = \$3$. Heinz faces a trade-off. If Heinz is infrequently on sale for less than the generic price, Heinz sells little to switchers. On the other hand, if Heinz is frequently on sale, it loses money on its sales to loyal customers.

We start by supposing that Heinz decides to charge a low sales price, \$2, once every n days. For the other $n - 1$ days, Heinz sells at the regular, nonsale (monopoly) price of \$3, which is the monopoly price given the demand curve of the loyal customers. During a sale, the switchers buy enough Heinz to last them for n days until it's on sale again. Consequently, the switchers never buy the generic product. (Some other customers are loyal to the generic, so they buy it even when Heinz is on sale.)

If the loyal customers find that Heinz is on sale, which happens $1/n$ of all days, they buy n days' worth at the sale price. Otherwise, they pay the regular (monopoly) price. If the other loyal customers were aware of this promotional pattern and shopped on a schedule such that they always bought on sale, this strategy would not maximize profit. However, their shopping schedules are determined independently: They buy many goods and are not willing to distort their shopping patterns solely to buy this one good on sale.[21]

---

[20]The rest of the U.S. market consists primarily of Hunt's Ketchup (15%) and generic or house brands (22%). In the following discussion, we assume that customers who are loyal to Hunt's or generic ketchup are unaffected by Heinz sales, and hence ignore those customers.

[21]We make this assumption for simplicity. In the real world, firms achieve a similar result by having random sales or by placing ads announcing sales where primarily switchers see them.

Could Heinz make more money by altering its promotion pattern? It does not want to place its good on sale more frequently because it would earn less from its loyal customers without making more sales to switchers. If it pays to hold sales at all, it does not want to have a sale less frequently because it would sell fewer units to switchers. During a promotion, Heinz wants to charge the highest price it can and yet still attract switchers, which is \$2. If it sets a lower price, the quantity sold is unchanged, so its profit falls. If Heinz sets a sale price higher than \$2, it loses all switchers.

Does Heinz increase its profit by having sales? Whether sales pay depends on the number of switchers, S, relative to the number of brand-loyal customers, B. If each customer buys one unit per day, then Heinz's profit per day if it sells only to loyals is $\pi = (p - MC)B = (3 - 1)B = 2B$, where $p = 3$ is Heinz's regular price and $MC = 1$ is its marginal and average cost. If Heinz uses the sale pricing scheme, its average profit per day is

$$\pi^* = 2B(n - 1)/n + (B + S)/n.$$

The first term is the profit it makes, \$2 per unit, selling B units to loyal customers for the fraction of days that Heinz ketchup is not on sale, $(n - 1)/n$. The second term is the profit it makes, \$1 per unit, selling $B + S$ units on the $1/n$ days that Heinz ketchup is on sale.

Thus, it pays to put Heinz on sale if $\pi < \pi^*$, or $2B < 2B(n - 1)/n + (B + S)/n$. Using algebra, we can simplify this expression to $B < S$. Thus, if the market has more switchers than loyal customers, the sales policy is more profitable than selling at a uniform price to only loyal customers.[22]

Summary

1. **Conditions for Price Discrimination.** A firm can price discriminate if it has market power, knows which customers will pay more for each unit of output, and can prevent customers who pay low prices from reselling to those who pay high prices. A firm earns a higher profit from price discrimination than from uniform pricing because (a) the firm captures some or all of the consumer surplus of customers who are willing to pay more than the uniform price and (b) the firm sells to some people who would not buy at the uniform price.

2. **Perfect Price Discrimination.** To perfectly price discriminate, a firm must know the maximum amount each customer is willing to pay for each unit of output. If a firm charges customers the maximum each is willing to pay for each unit of output, the monopoly captures all potential consumer surplus and sells the efficient (competitive) level of output. Compared to competition, total welfare is the same, consumers are worse off, and firms are better off under perfect price discrimination.

3. **Group Price Discrimination.** A firm that does not have enough information to perfectly price discriminate may know the relative elasticities of demand of various groups of its customers. Such a profit-maximizing firm charges groups of consumers prices in proportion to their elasticities of demand, the group of consumers with the least elastic demand paying the

[22]Hendel and Nevo (2013) examined the soft-drink market. They found that some price-sensitive consumers buy during sales and stockpile, while less price-sensitive consumers do not stockpile. As a result, sales capture 25–30% of the gap between the non-price-discriminating profit and the profit from (unattainable) price-discrimination in which the seller can identify consumer types and prevent arbitrage.

highest price. Welfare is less under group price discrimination than under competition or perfect price discrimination but may be greater or less than that under single-price monopoly.

4. **Nonlinear Price Discrimination.** Some firms charge customers different prices depending on how many units they purchase. A common example involves quantity discounts, so that the per-unit price for consumers who buy larger quantities is less than the per-unit price for consumers who buy smaller quantities.

5. **Two-Part Pricing.** By charging consumers an access fee for the right to buy a good and a separate price per unit, firms may earn higher profits than from using uniform pricing. If a firm knew all its customers' demand curves and could charge a different access fee to every customer, it could use two-part pricing to capture all potential consumer surplus. Even a firm that does not know each customer's demand curve or that cannot vary the access fee across customers can still use two-part pricing to earn a higher profit than it could earn using uniform pricing.

6. **Tie-In Sales.** A firm may increase its profit by using a tie-in sale that allows customers to buy one product only if they also purchase another one. In a requirement tie-in sale, customers who buy one good must make all of their purchases of another good or service from that firm. Some firms use pure bundling, in which they allow consumers to buy only the bundle. Others use mixed bundling, in which consumers may buy either the bundle or the individual goods separately. Bundling is likely to be a profitable pricing strategy when consumers have reservation prices that are negatively correlated—when consumers who have a relatively high willingness to pay for one good have a relatively low willingness to pay for the other good.

7. **Advertising.** A monopoly advertises or engages in other promotional activity to shift its demand curve to the right or make it less elastic so as to raise its profit net of its advertising expenses.

Questions

Select questions are available on MyLab Economics*; * = answer appears at the back of this book;* **A** *= algebra problem;* **C** *= calculus problem.*

1. Conditions for Price Discrimination

1.1 In the examples in Table 12.1, if the movie theater does not price discriminate, it charges either the highest price the college students are willing to pay or the one that the senior citizens are willing to pay. Why doesn't it charge an intermediate price? (*Hint*: Discuss how the demand curves of these two groups are unusual.)

*1.2 Many governments provide social housing to families with low or modest incomes. The rent charged for an accommodation is set according to the renter's income and rent increases are limited to keep it affordable. Explain why governments behave this way.

1.3 Australian citizens aged 60 or over who are not working more than 20 hours per week are eligible for a free Seniors Card. Issued by individual state governments, the card provides government concessions, business discounts, and special offers on a range of goods and services including travel, movies and entertainment, and personal and professional services. The names of businesses offering a discount of at least 10% are listed in the Discount Directory made available to card holders, and this is also available online. Why would Australian businesses want to participate in such an initiative?

1.4 Alexx's monopoly currently sells its product at a single price. What conditions must be met so that he can profitably price discriminate?

1.5 The European Youth Card offers discounts on culture, travel, accommodation, shopping, and services in most European countries. Anyone under the age of 30 or 26, depending on the country, can purchase one for under €20. Typically effective for one year, there are three versions: a classic card, a student card, and a co-brand card, the latter issued by a specific firm or local government. What is the main motivation for firms to offer discounts to holders of these cards?

1.6 Goods often sell for different prices in different countries. Suppose the inverse demand curves in two countries are $p_1 = 40 - 3Q_1$ and $p_2 = 90 - 2Q_2$, and a monopoly produces at constant marginal cost, $MC = 10$. What is the profit-maximizing price and quantity in each country? Will the monopolist necessarily be able to sell its output at those prices if shipping costs between the two countries are low? Explain.

1.7 Disneyland price discriminates by charging lower entry fees for children than adults and for local residents than for other visitors. Why does it not have a resale problem?

1.8 Renting an economy car in Munich and Berlin costs about the same, €19 per day. However, the rental price for a full size car is €10 euros per day more in Munich (€36) than in Berlin (€26). Is the difference in the rental prices of economy and full-size cars in those two locations evidence of price discrimination? Explain.

2. Perfect Price Discrimination

2.1 Using the information in the Application "Botox and Price Discrimination," determine how much Allergan loses by being a single-price monopoly rather than a perfectly price-discriminating monopoly. Explain your answer.

2.2 Suppose that the cost function curve for a natural monopoly is $C(Q) = 48 + 6Q$ and the inverse demand function it faces is $p = 38 - 2Q$. What are the firm's maximum profits if it cannot price discriminate? Compare that result with the firm's profits if it can price discriminate perfectly. Does the perfectly price-discriminating monopoly produce a level of output where price is less than average cost? Show your results for perfect price discrimination in a diagram.

2.3 Can Table 12.1 be modified so that the movie theater in Solved Problem 12.1 does not earn more by perfectly price discriminating than from charging a single price? What changes to the table would increase the extra profit from perfectly price discriminating?

2.4 How would the answers to Solved Problem 12.1 and Table 12.1 change if seniors' reservation price was $2.50?

2.5 Founded in 1998, eBid is one of the oldest and largest online auction websites. It operates in 23 countries, has over two million registered users, and offers a huge array of items for auction. Does eBay engage in price discrimination? If so, what type?

*2.6 If a monopoly faces an inverse demand curve of $p = 90 - Q$, has a constant marginal and average cost of 30, and can perfectly price discriminate, what is its profit? What are the consumer surplus, welfare, and deadweight loss? How would these results change if the firm were a single-price monopoly? **A**

2.7 To promote her platinum-selling CD, *Feels Like Home*, in 2005, singer Norah Jones toured the country for live performances. However, she sold an average of only two-thirds of the tickets available for each show, T^* (Robert Levine, "The Trick of Making a Hot Ticket Pay," *New York Times*, June 6, 2005, C1, C4).

a. Suppose that the local promoter is the monopoly provider of each concert. Each concert hall has a fixed number of seats. Assume that the promoter's cost is independent of the number of people who attend the concert (Ms. Jones received a guaranteed payment). Graph the promoter's marginal cost curve for the concert hall, where the number of tickets sold is on the horizontal axis (be sure to show T^*).

b. If the monopoly can charge a single market price, does the concert's failure to sell out prove that the monopoly set too high a price? Explain.

c. Would your answer in part b be the same if the monopoly can perfectly price discriminate? Use a graph to explain.

2.8 See the Application "Google Uses Bidding for Ads to Price Discriminate," which discusses how advertisers on Google's website bid for the right for their ads to be posted when people search for certain phrases. Should a firm that provides local services (such as plumbing or pest control) expect to pay more or less for an ad in a small town or a large city? Why?

3. Group Price Discrimination

3.1 Group price discrimination (also called third-degree price discrimination) is the most common type of price discrimination. Under it, a firm charges each group of customers a different price, but it does not charge different prices within the group. If the two countries in Question 1.6 were instead two different groups of customers, resale could be prevented, and average cost was also constant at 10, what would be the monopolist's total profits? Would the monopolist ever incur a loss in selling to one of its groups of customers? Explain.

3.2 There are two distinct groups of customers in the market for a firm's product. Their inverse demand curves are $p_1 = 5 - Q_1$ and $p_2 = 10 - Q_2$. The monopoly's constant marginal cost of production is zero.

a. If the monopoly can price discriminate, what is its total profit?

b. If the monopoly cannot prevent resale and the monopolist charges a uniform price, will it charge the higher price for Group 2's customers or the lower price for Group 1's customers to maximize total profit? What happens to consumer surplus if the monopolist charges the higher price for Group 2's customers?

3.3 Does a monopoly's ability to price discriminate between two groups of consumers depend on its marginal cost curve? To answer this, compare the situation for the monopolist when it engages in group price discrimination in Question 3.2 with the situation in which it has a higher constant marginal cost equal to 5.

3.4 Grocery stores often set consumer-specific prices by issuing frequent-buyer cards to willing customers and collecting information on their purchases. Grocery chains can use that data to offer customized discount coupons to individuals.

a. Are grocery stores engaging in perfect or group price discrimination, or some other type of pricing?

b. How should a grocery store use past-purchase data to set individualized prices to maximize its profit? (*Hint*: Refer to a customer's price elasticity of demand.)

*3.5 A patent gave Sony a legal monopoly to produce a robot dog called Aibo ("eye-BO"). The Chihuahua-size pooch robot can sit, beg, chase balls, dance, and play an electronic tune. When Sony started selling the toy, it announced that it would sell 3,000 Aibo robots in Japan for about \$2,000 each and a limited litter of 2,000 in the United States for \$2,500 each. Suppose that Sony's marginal cost of producing Aibos is \$500. Its inverse demand curve is $p_J = 3{,}500 - \frac{1}{2}Q_J$ in Japan and $p_A = 4{,}500 - 4{,}500Q_A$ in the United States. Solve for the optimal prices and quantities (assuming that U.S. customers cannot buy robots from Japan). Show how the profit-maximizing price ratio depends on the elasticities of demand in the two countries. What are the deadweight losses in each country, and in which is the loss from monopoly pricing greater? (*Hint*: See Solved Problem 12.2.) **A**

*3.6 A monopoly sells its good in the U.S. and Japanese markets. The American inverse demand function is $p_A = 100 - Q_A$, and the Japanese inverse demand function is $p_J = 80 - 2Q_J$, where both prices, p_A and p_J, are measured in dollars. The firm's marginal cost of production is $m = 20$ in both countries. If the firm can prevent resale, what price will it charge in both markets? (*Hint*: See Solved Problem 12.2.) **A**

*3.7 A copyright gave Universal Studios the legal monopoly to produce and sell the *Mama Mia!* DVD. The DVD sold for \$20 in the United States and \$36 (£22) in the United Kingdom. I estimate that the inverse demand functions for the United States and the United Kingdom were $p_A = 39 - 3Q_A$ and $p_B = 71 - 7Q_B$, respectively. The marginal cost in both countries was $m = 1$. Solve for Universal's optimal quantities and prices. (*Hint*: See Solved Problem 12.2.) **A**

3.8 Warner Home Entertainment sold the *Harry Potter and the Prisoner of Azkaban* two-DVD movie set around the world. Warner charged 33% more in Canada and 66% more in Japan than in the United States, where it charged \$15. Given that Warner's marginal cost was \$1, determine what the elasticities of demand must have been in the United States, Canada, and Japan if Warner was profit maximizing. (*Hint*: See Solved Problem 12.3.) **A**

*3.9 Warner Home Entertainment sold the *Harry Potter and the Prisoner of Azkaban* two-DVD movie set in China for about \$3, which was only one-fifth the U.S. price, and sold nearly 100,000 units. The price is extremely low in China because Chinese consumers are less wealthy than those in developed countries and because (lower-quality) pirated versions were available in China for \$0.72–\$1.20 (Jin Baicheng, "Powerful Ally Joins Government in War on Piracy," *China Daily*, March 11, 2005, 13). Assuming a marginal cost of \$1, what was the Chinese elasticity of demand? Derive the demand function for China and illustrate Warner's policy in China using a figure similar to panel a in Figure 12.3. **A**

3.10 Suppose a monopoly sells its product in two countries and resale is not possible. The inverse demand curves in the two countries are $p_1 = 200 - 10Q_1$ and $p_2 = 122 - 5Q_2$. The marginal cost of production in the first country is constant at 20 but is higher by 2 in the second country due to shipping costs. What is the monopoly's optimal price in each country, and what is the relationship between marginal revenue in the two countries? (*Hint*: See Solved Problem 12.2.) **A**

3.11 A monopoly sells its good in the United States, where the elasticity of demand is -2, and in Japan, where the elasticity of demand is -5. Its marginal cost is \$10. At what price does the monopoly sell its good in each country if resale is impossible? **A**

*3.12 In Solved Problem 12.4, calculate the firm's profit with and without a ban against shipments between the two countries.

3.13 How would the analysis in Solved Problem 12.3 change if $m = 7$ or if $m = 4$? (*Hint*: Where $m = 4$, the marginal cost curve crosses the *MR* curve three times—if we include the vertical section. The single-price monopoly will choose one of these three points where its profit is maximized.)

*3.14 Limited-time offers are a powerful marketing tool. Good ones are billed as the "best ever," create a sense of urgency (for example, "while supplies last"), provide incentives (for example, free shipping or gifts) for purchases exceeding a certain amount, and focus on how much the customer is saving as opposed to spending. Why do firms use limited-time offers? (*Hint*: See the Application "Buying Discounts.")

3.15 The price of a cappuccino at any given coffee shop can be more than double the price of a regular coffee with milk that is in a container of the same size. The basic ingredients of both products are identical, but the milk in the cappuccino will have been frothed up with pressurized steam. Is this one difference in the cost of production sufficient to explain the price differential, or is it more likely that consumers of cappuccinos subsidize the consumers of regular coffee by paying exorbitant prices?

3.16 In 2015, the European Commission charged six U.S. studios and a U.K. pay television company, Sky UK, with unfairly blocking access to films and other content. The charges challenge the studios' requirement under contracts that Sky UK block access for consumers outside Britain and Ireland. (James Kanter and Mark Scott, "E.U. Opens Antitrust Case Against Major U.S. Studios and Sky UK," *New York Times*, July 23, 2015.) The studios have separate contracts with broadcasters in other countries. Why do the studios want such restrictions?

3.17 In 2016, Tesla introduced a new version of its popular Model S electric sedan, the S 60, starting at $66,000. It comes with a battery pack that can store 75 kilowatt-hours of electricity, but its software limits it to using only 60. However, for $9,000, the owners can get a simple software upgrade to get the full 75, increasing the car's range by about 19 percent. Why is Tesla purposefully hobbling its cars? Explain its pricing strategy.

4. Nonlinear Pricing

4.1 Under a single-rate electricity tariff for households and small businesses in Australia, a flat usage rate is charged for "blocks" of electricity consumed. The rate for the first block is higher than that for the second, and the lowest rate applies to electricity usage in excess of the second block. Suppose the inverse demand curve for a typical consumer is $p = 30 - Q/3$, where p is cents per kilowatt hour (kWh) and Q is kWh per day, and the profit-maximizing monopoly has a constant marginal and average cost of 20 cents per kWh.

a. What is the quantity where price equals marginal cost?

b. What would be the optimal quantity and price and the firm's profit, consumer surplus, and deadweight loss if the monopoly could set only a single price?

c. What would be the firm's profit, consumer surplus, and deadweight loss if the consumer uses 20 kWh of electricity per day when the monopoly charges 26.67 cents/kWh on the first block of 10 kWh/day, 23.33 cents/kWh on the second block of 10 kWh/day, and 20 cents/kWh on the remaining usage?

d. Comparing your answers to parts b and c, would the typical consumer be worse off with declining block pricing?

4.2 The quantity-discriminating monopoly in panel a of Figure 12.4 sets three prices that depend on the quantity a consumer purchases. The firm's profit is $\pi = p_1Q_1 + p_2(Q_2 - Q_1) + p_3(Q_3 - Q_2) - mQ_3$, where p_1 is the high price charged on the first Q_1 units (first block), p_2 is a lower price charged on the next $Q_2 - Q_1$ units, p_3 is the lowest price charged on the $Q_3 - Q_2$ remaining units, Q_3 is the total number of units actually purchased, and $m = \$30$ is the firm's constant marginal and average cost. Use calculus to determine the profit-maximizing p_1, p_2, and p_3. **C**

*4.3 In our discussion of Figure 12.4, we assumed that the monopoly engaged in block-pricing by setting both block prices so that they were on the demand curve. However, suppose the monopoly sets the first block at 20 units but can choose a first-block price that is greater than $70. It then allows consumers to buy as many additional units as they want at $30 per unit. Can the monopoly choose a price for the first block such that consumers are willing to buy 60 units, the monopoly captures the entire potential surplus, and society does not suffer a deadweight loss? If so, what is the first-block price?

5. Two-Part Pricing

5.1 Using math, show why two-part pricing causes customers who purchase relatively few units to pay more per unit than customers who buy more units. **C**

5.2 Access to certain golf courses such as Golf de Morfontaine, north of Paris, France, is restricted to members and their guests. In addition to the membership fee, there is a charge (greens fee) to play each round of golf on private golf courses. Suppose the golf course sells memberships to two groups of 196 golfers each. The demand curve for the first group is $Q_1 = 1{,}860 - p$ and for the second group is $Q_2 = 1{,}600 - p$, where Q_1 and Q_2 are rounds per week and p is euros per round. The membership fee, δ, and the per-round price, p, is the same for each golfer, the marginal cost of a round of golf is $MC = €70$, and there are no fixed costs. What is the optimal price per round and the membership fee? How many rounds of golf does each group play? What is the consumer surplus for each? **A**

*5.3 A tennis club with a good reputation for providing high-quality programs is considering changing the way it charges its customers for lessons. Elena, one of those customers, is keen on improving her tennis skills. Her inverse demand curve for tennis lessons is $p = 100 - 4q$, where q is the number of one-hour lessons per month. The club has been charging her and all of its other customers the same price per lesson and has a constant marginal and average cost of €20. If the club were to instead charge Elena a membership fee and set her per-lesson price equal to its marginal cost, what membership fee would maximize profit for the club, and how much extra profit would the club earn? **A**

5.4 As described in the Application "iTunes for a Song," Shiller and Waldfogel (2011) estimated that if iTunes used two-part pricing, charging an annual access fee and a low price per song, it would raise its profit by about 30% relative to what it would earn using uniform or variable pricing. Assume that iTunes uses two-part pricing and assume that the marginal cost of an additional download is zero. How should iTunes set its profit-maximizing price per song if all consumers are identical? Illustrate profit-maximizing two-part pricing in a diagram for the identical consumer case. Explain why the actual profit-maximizing price per song is positive.

5.5 Explain why charging a higher or lower price than $p = 10$ reduces the monopoly's profit in Figure 12.5. Show the monopoly's profit if $p = 20$ and compare it to its profit if $p = 10$.

6. Tie-In Sales

6.1 Some people purchase expensive water conditioners to reduce the mineral content in their water, enhance the effectiveness of soaps and detergents, and reduce deposits in plumbing systems. A manufacturer of a certain brand will often designate a single plumbing company to be its sole supplier in a geographic area and provide special training for that dealer to service the product. The product warranty will also often state explicitly that the customer should request service from the dealer if there is a problem and use parts made only by the manufacturer. What are the benefits to the manufacturer and the dealer from entering into such an agreement?

6.2 A monopoly sells two products, of which consumers want only one. Assuming that it can prevent resale, can the monopoly increase its profit by bundling them, forcing consumers to buy both goods?

6.3 Explain why in Table 12.2 the firm does not use mixed bundling.

*6.4 An Indian restaurant has four potential customers with the following reservation prices:

	Main course	Dessert	Combination
Customer 1	£23	£9	£32
Customer 2	£30	£12	£42
Customer 3	£48	£5	£53
Customer 4	£60	£7	£67

a. Calculate the profit-maximizing prices and the revenue when the firm prices each course separately.

b. Is the profit-maximizing combination price and resulting revenue higher? (*Hint*: See Solved Problem 12.5.) **A**

7. Advertising

7.1 Using a graph similar to Figure 12.7, explain why a firm might not want to spend A on advertising, even though it shifts the firm's demand curve to the right. (*Hint*: Discuss what happens to the elasticity of demand or the price at the monopoly optimum.)

7.2 Google AdWords is a way for businesses to promote their products online; for example, when people are searching for resort packages, tours, or flights for a trip or vacation. Displaying products to customers at the time that they are looking for related content is called intent-based marketing. Being at the top of the first page of a Google search is a good way to be seen. AdWords help to do this by telling Google the searches for which a business wants its advertisement to be found. There is no sign-up cost and no cost for Google support to help clients find keywords that are right for them, but businesses do pay when someone either clicks to visit their website or calls them. How do AdWords affect the marginal benefit curve for an advertiser, and why?

*7.3 How might a natural disaster, violence, rising crime rates, or health alerts affect the optimal amount of advertising a tourism business chooses to purchase for the region affected, and why? Use a diagram to explain your answer. (*Hint*: See Figure 12.8.)

*7.4 A monopoly's inverse demand function is $p = 800 - 4Q + 0.2A^{0.5}$, where Q is its quantity, p is its price, and A is the level of advertising. Its marginal cost of production is 2, and its cost for a unit of advertising is 1. What are the firm's profit-maximizing price, quantity, and level of advertising? (*Hint*: See Appendix 12E.) **C**

7.5 Use a diagram to illustrate the effect of social media on the demand for Super Bowl commercials. (*Hint*: See the Application "Super Bowl Commercials.")

8. Challenge

8.1 When it is first released, a novel by a best-selling author is typically published as a hardcover book with a paper dust jacket covering it, and the paper used to print it is of a relatively high quality. However, after a few months, hardcover copies that have not been sold will be heavily discounted and paperback versions will be produced. Give an explanation based on price discrimination for this practice.

8.2 In the Challenge Solution, did the sales method achieve the same group-price-discrimination outcome that Heinz would achieve if it could set separate prices for loyal customers and for switchers? Why or why not?

13 Oligopoly and Monopolistic Competition

Three can keep a secret, if two of them are dead.
—Benjamin Franklin

Challenge

Government Aircraft Subsidies

Aircraft manufacturers lobby their governments for subsidies, which they use to compete better with rival firms. Airbus SAS, based in Europe, and the Boeing Co., based in the United States, are the only two major manufacturers of large commercial jet aircraft. France, Germany, Spain, and the United Kingdom subsidize Airbus, which competes in the wide-body aircraft market with Boeing. The U.S. government decries the European subsidies to Airbus despite giving lucrative military contracts to Boeing, which the Europeans view as implicit subsidies.

This government largesse does not magically appear. Managers at both Boeing and Airbus lobby strenuously for this support. For example, in 2015, Boeing spent $21,921,000 on lobbying and was represented by 95 lobbyists, including 3 former congressmen.

Washington and the European Union have repeatedly charged each other before the World Trade Organization (WTO) with illegally subsidizing their respective aircraft manufacturers. In 2010, the WTO ruled that Airbus received improper subsidies for its A380 superjumbo jet and several other aircraft, hurting Boeing, as the United States charged in 2005. In 2012, the WTO ruled that Boeing and Airbus both received improper subsidies. In 2015, the WTO agreed to investigate a complaint about Washington State subsidies to Boeing, and Boeing questioned government loans to Airbus. Thus, the cycle of subsidies, charges, agreements, and new subsidies continues.

If only one government subsidizes its firm, how does the firm use the subsidy to gain a competitive advantage? What happens if both governments subsidize their firms? Should Boeing and Airbus lobby for government subsidies that result in a subsidy war?

The major airlines within a country compete with relatively few other firms. Consequently, each firm's profit depends on the actions it and its rivals take. Similarly, three firms—Nintendo, Microsoft, and Sony—dominate the video game market, and each firm's profit depends on how its price stacks up against the prices of its rivals and whether its product has better features.

oligopoly
a small group of firms in a market with substantial barriers to entry

Airline and the video game firms are each an **oligopoly**: a small group of firms in a market with substantial barriers to entry. Because relatively few firms compete in such a market, each can influence the price, and hence each affects rival firms. The need to consider the behavior of rival firms makes an oligopolistic firm's profit-maximization decision more difficult than that of a monopoly or a competitive firm. A monopoly has no rivals, and a competitive firm ignores the behavior of individual rivals—it considers only the market price and its own costs in choosing its profit-maximizing output.

An oligopolistic firm that ignores or inaccurately predicts its rivals' behavior is likely to suffer a loss of profit. For example, as its rivals produce more cars, the price Ford can get for its cars falls. If Ford underestimates how many cars its rivals will produce, Ford may produce too many automobiles and lose money.

cartel
a group of firms that explicitly agree to coordinate their activities

Oligopolistic firms may act independently or may coordinate their actions. A group of firms that explicitly agree (collude) to coordinate their activities is called a **cartel**. These firms may agree on how much each firm will sell or on a common price. By cooperating and behaving like a monopoly, the members of a cartel collectively earn the monopoly profit—the maximum possible profit. In most developed countries, cartels are generally illegal.

If oligopolistic firms do not collude, they earn lower profits. However, because oligopolies consist of relatively few firms, oligopolistic firms that act independently may earn positive economic profits in the long run, unlike competitive firms.

monopolistic competition
a market structure in which firms have market power but no additional firm can enter and earn positive profits

In an oligopolistic market, one or more barriers to entry keep the number of firms small. In a market with no barriers to entry, firms enter the market until profits fall to zero. In perfectly competitive markets, enough entry occurs that firms face a horizontal demand curve and are price takers. However, in other markets, even after entry has driven profits to zero, each firm faces a downward-sloping demand curve. Because of this slope, the firm can charge a price above its marginal cost, creating a market failure due to inefficient (too little) consumption (Chapter 9). **Monopolistic competition** is a market structure in which firms have market power (the ability to raise price profitably above marginal cost) but no additional firm can enter and earn positive profits.

In this chapter, we examine cartelized, oligopolistic, and monopolistically competitive markets in which firms set quantities or prices. As noted in Chapter 11, the monopoly equilibrium is the same whether a monopoly sets price or quantity. Similarly, if colluding oligopolies sell identical products, the cartel equilibrium is the same whether they set quantity or price. The oligopolistic and monopolistically competitive equilibria differ, however, if firms set prices instead of quantities.

In this chapter, we examine six main topics

1. **Market Structures.** The number of firms, price, profits, and other properties of markets vary, depending on whether the market is monopolistic, oligopolistic, monopolistically competitive, or competitive.
2. **Cartels.** If firms successfully coordinate their actions, they can collectively behave like a monopoly.
3. **Cournot Oligopoly.** In a Cournot oligopoly, firms choose their output levels without colluding and the market output, price, and firms' profits lie between the competitive and monopoly levels.
4. **Stackelberg Oligopoly.** In a Stackelberg oligopoly, in which a leader firm chooses its output level before its identical-cost rivals, market output is greater than if all firms choose their output simultaneously, and the leader makes a higher profit than the other firms.
5. **Bertrand Oligopoly.** In a Bertrand oligopoly, in which firms choose prices, the equilibrium differs from the quantity-setting equilibrium and depends on the degree of product differentiation.
6. **Monopolistic Competition.** When firms can freely enter the market but, in equilibrium, face downward-sloping demand curves, firms charge prices above marginal cost but make no profit.

13.1 Market Structures

Markets differ according to the number of firms in the market, the ease with which firms may enter and leave the market, and the ability of firms in a market to differentiate their products from those of their rivals. Table 13.1 lists characteristics and properties of the four major market structures: monopoly, oligopoly, monopolistic competition, and perfect competition. In the table, we assume that the firms face many price-taking buyers.

The first row in the table describes the number of firms in each market structure. A monopoly is a single (*mono*) firm in a market. An *oligopoly* usually has a small number (*oligo*) of firms. A *monopolistically competitive* market may have a few or many firms, though typically few. A *perfectly competitive* market has many firms.

Monopolistic and oligopolistic markets have few firms due to insurmountable entry barriers, such as government licenses or patents (row 2). In contrast, in monopolistically competitive and perfectly competitive markets, entry occurs until no new firm can profitably enter, so that long-run economic profit is zero (row 3). Monopolistic and oligopolistic firms can earn positive long-run profits.

Perfectly competitive firms face horizontal demand curves so they are price takers. Monopolistically competitive markets have fewer firms than do perfectly competitive markets. Because monopolistically competitive firms have relatively few rivals, they are large relative to the market, so each monopolistically competitive firm faces a downward-sloping demand curve as do monopolistic and oligopolistic firms. Thus, noncompetitive firms are price setters (row 4). That is, all but perfectly competitive firms have some degree of market power—the ability to set price above marginal cost—so a market failure occurs in each of these noncompetitive market structures because the price is above marginal cost. Typically, the fewer the firms in a market, the higher is the price (row 5).

Oligopolistic and monopolistically competitive firms pay attention to rival firms' behavior, in contrast to monopolistic or perfectly competitive firms (row 6). A monopoly has no rivals. A perfectly competitive firm ignores the behavior of individual rivals

Table 13.1 Properties of Monopoly, Oligopoly, Monopolistic Competition, and Perfect Competition

	Monopoly	Oligopoly	Monopolistic Competition	Perfect Competition
1. Number of firms	1	Few	Few or many	Many
2. Entry conditions	No entry	Limited entry	Free entry	Free entry
3. Long-run profit	≥ 0	≥ 0	0	0
4. Ability to set price	Price setter	Price setter	Price setter	Price taker
5. Price level	Very high	High	High	Low
6. Strategy dependent on individual rival firms' behavior	No (has no rivals)	Yes	Yes	No (cares about market price only)
7. Products	Single product	May be differentiated	May be differentiated	Undifferentiated
8. Example	Local natural gas utility	Automobile manufacturers	Books, restaurants	Apple farmers

in choosing its output because the market price tells the firm everything it needs to know about its competitors.

Oligopolistic and monopolistically competitive firms may produce differentiated products (row 7). For example, oligopolistic car manufacturers produce automobiles that differ in size, weight, and various other dimensions. In contrast, perfectly competitive apple farmers sell undifferentiated (homogeneous) products.

13.2 Cartels

People of the same trade seldom meet together, even for merriment and diversion, but the conversation ends in a conspiracy against the public, or some contrivance to raise prices. —Adam Smith, 1776

Oligopolistic firms have an incentive to form cartels in which they collude in setting prices or quantities to increase their profits. The Organization of Petroleum Exporting Countries (OPEC) is a well-known example of an international cartel; however, many cartels operate within a single country.

Typically, each member of a cartel agrees to produce less output than it would if it acted independently. As a result, the market price rises and the firms earn higher profits. If the firms reduce market output to the monopoly level, they achieve the highest possible collective profit.

Luckily for consumers, cartels often fail because of government policies that forbid cartels or because members of the cartel "cheat" on the agreement. Each member has an incentive to cheat because it can raise its profit if it increases its output while other cartel members stick to the agreement.

Why Cartels Form

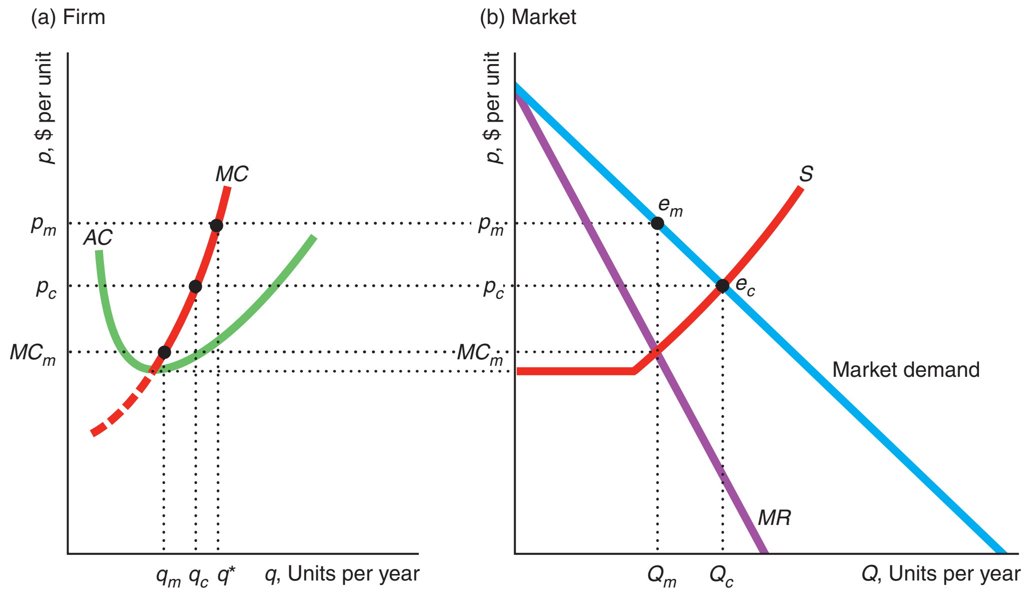
A cartel forms if members of the cartel believe that they can raise their profits by coordinating their actions. But if a firm maximizes its profit when acting independently, why should joining a cartel increase its profit? The answer involves a subtle argument. When a firm acts independently, it considers how increasing its output affects its own profit only. The firm does not care that when it increases its output, it lowers the profits of other firms. A cartel, in contrast, takes into account how changes in any one firm's output affect the profits of all members of the cartel. As a result, the aggregate profit of a cartel can exceed the combined profits of the same firms acting independently.

Although cartels are most common in oligopolistic markets, occasionally we see cartels formed in what would otherwise be highly competitive markets, such as in markets of professionals. If a competitive firm lowers its output, it raises the market price very slightly—so slightly that the firm ignores the effect not only on other firms' profits but also on its own. If all the identical competitive firms in an industry lower their output by this same amount, however, the market price will change noticeably. Recognizing this effect of collective action, a cartel chooses to produce a smaller market output than is produced by a competitive market.

Figure 13.1 illustrates this difference between a competitive market and a cartel. This oligopolistic market has n firms, and no further entry is possible. Panel a shows the marginal and average cost curves of a typical perfectly competitive firm. If all firms are price takers, the market supply curve, S in panel b, is the horizontal sum of the individual marginal cost curves above minimum average cost. At the competitive price, p_c, each price-taking firm produces q_c units of output (which is determined by

Figure 13.1 Competition Versus Cartel

(a) This panel shows the marginal cost and average cost of one of the n identical firms in the market. A competitive firm produces q_c units of output, whereas a cartel member produces $q_m < q_c$. At the cartel price, p_m, each cartel member has an incentive to increase its output from q_m to q^* (where the dotted line at p_m intersects the MC curve). (b) The competitive equilibrium, e_c, has more output and a lower price than the cartel equilibrium, e_m.



the intersection in panel a of MC and the dotted line at p_c).[1] The market output is $Q_c = nq_c$ (where S intersects the market demand curve in panel b).

Now suppose that the firms form a cartel. Should they reduce their output? At the competitive output, the cartel's marginal cost (which is the competitive industry supply curve, S in panel b) is greater than its marginal revenue, so the cartel's profit rises if it reduces output. The cartel's collective profit rises until output is reduced to where its marginal revenue equals its marginal cost at Q_m, the monopoly output. If the profit of the cartel increases, the profit of each of the n members of the cartel also increases. To achieve the cartel output level, each firm must reduce its output to $q_m = Q_m/n$, as panel a shows.

Why must the firms form a cartel to achieve these higher profits? A competitive firm produces q_c, where its marginal cost equals the market price. If only one firm reduces its output, it loses profit because it sells fewer units at essentially the same price. By getting all the firms to lower their output together, the cartel raises the market price and hence individual firms' profits. The less elastic the market demand the potential cartel faces, all else the same, the higher the price the cartel sets and the greater the benefit from cartelizing. If the penalty for forming an illegal cartel is relatively low, some unscrupulous people may succumb to the lure of extra profits and join.

[1]In the figure, the competitive price exceeds the minimum average cost. These competitive firms earn a profit because the number of firms is fixed.

Why Cartels Fail

In most developed countries, cartels are generally illegal, firms in the cartel may incur fines, and the owners or managers of these firms may be subject to individual fines and jail terms.[2] Further, many cartels fail even without legal intervention.

Some cartels fail because they do not control enough of the market to raise the price significantly. For example, copper producers tried four times to form an international cartel between 1918 and 1988. In the most recent attempt, the Intergovernmental Council of Copper Exporting Countries, controlled less than a third of the noncommunist world's copper production and faced additional competition from firms that recycle copper from scrap materials. Because of this competition from noncartel members, the cartel could not successfully increase copper prices and dissolved in 1988.

Members of a cartel have incentives to cheat on the cartel agreement. The owner of a participating firm may reason, "I joined the cartel to encourage others to reduce their output, which raises the market price and increases profits for everyone. However, I can make even more if I cheat on the cartel agreement by producing extra output. I can get away with cheating if the other firms can't tell who is producing the extra output because my firm is just one of many firms and my increase in output will hardly affect the market price." By this reasoning, it is in each firm's best interest for all *other* firms to honor the cartel agreement—thus driving up the market price—while it ignores the agreement and makes extra profitable sales at the high price.

Figure 13.1 illustrates why firms want to cheat. At the cartel output, q_m in panel a, each cartel member's marginal cost is MC_m. A firm that does not restrict its output to the cartel level can increase its profit. It can earn the market price, p_m, on each extra unit it sells because each individual firm's output has little effect on the market price. That is, the firm can act like a price taker, so its marginal revenue equals the market price. The firm maximizes its profit by selling q^* units, which is determined by the intersection of its marginal cost curve and the dotted line at p_m. Because its marginal revenue is above its marginal cost for all the extra units it sells (those between q_m and q^*), it makes extra money by violating the cartel agreement. As more and more firms leave the cartel, the cartel price falls. Eventually, if enough firms quit, the cartel collapses.

Laws Against Cartels

In the late nineteenth century, cartels (or, as they were called then, *trusts*) were legal and common in the United States. Oil, railroad, sugar, and tobacco trusts raised prices substantially above competitive levels.[3]

In response to the trusts' high prices, the U.S. Congress passed the Sherman Antitrust Act in 1890 and the Federal Trade Commission Act of 1914, which prohibit firms from *explicitly* agreeing to take actions that reduce competition.[4] In particular,

---

[2]With rare exceptions, it is illegal for firms to collude over prices, quantities, market areas, or the equivalent. However, in most jurisdictions, firms are allowed to coordinate R&D efforts or technical standards if they wish.

[3]Nineteenth century and early twentieth century robber barons who made fortunes due to these cartels include John Jacob Astor (real estate, fur), Andrew Carnegie (railroads, steel), Henry Clay Frick (steel), Jay Gould (finance, railroads), Mark Hopkins (railroads), J. P. Morgan (banking), John D. Rockefeller (oil), Leland Stanford (railroads), and Cornelius Vanderbilt (railroads, shipping).

[4]U.S. laws do not prohibit all cartels. A bizarre Supreme Court decision largely exempted Major League Baseball from antitrust laws. Unions are explicitly exempt from antitrust laws. Workers may act collectively to raise wages. A historical justification for exempting labor unions was that the workers faced employers that could exercise monopsony power (Chapter 15). As long as they do not discuss such issues as prices and quantities, firms may coordinate R&D efforts or technical standards.

these laws prohibit cartels from collectively setting prices. In legal jargon, such price-fixing is a *per se* violation: It is strictly against the law and firms have no possible mitigating justifications. By imposing penalties on firms caught colluding, government agencies seek to discourage cartels from forming.

The Antitrust Division of the Department of Justice (DOJ) and the Federal Trade Commission (FTC) divide the responsibility for U.S. antitrust policy. The U.S. Department of Justice, quoting the Supreme Court that collusion was the "supreme evil of antitrust," stated that prosecuting cartels was its "top enforcement priority." The FTC's objective is "to prevent unfair methods of competition in commerce" and "to administer . . . other consumer protection laws." Both U.S. agencies can use criminal and civil law to attack cartels, price-fixing, and other anticompetitive actions.

However, cartels persist despite these laws for three reasons. First, international cartels and cartels within certain countries operate legally. Some international cartels, such as OPEC, that are organized by countries rather than by firms operate legally.

Second, some illegal cartels operate believing that they can avoid detection or, if caught, that the punishment will be insignificant. At least until recently, they were generally correct. For example, in 1996, Archer Daniels Midland (ADM) paid to settle three civil price-fixing-related cases: $35 million in a case involving citric acid (used in many consumer products), $30 million to shareholders as compensation for lost stock value after the citric acid price-fixing scandal became public, and $25 million in a lysine (an animal feed additive) case. ADM paid a $100 million fine in a federal criminal case for fixing the price of lysine and citric acid in 1996, but only eight years later, ADM settled a fructose corn syrup price-fixing case for $400 million.

Third, some firms are able to coordinate their activities without explicitly colluding and thereby running afoul of competition laws. To determine guilt, U.S. antitrust laws use evidence of conspiracy—explicit agreements—rather than the economic effect of the suspected cartel. The law does not prohibit firms from charging monopoly-level prices—it prohibits explicitly agreeing to raise prices. As a result, some groups of firms charge monopoly-level prices without violating competition laws. These firms may *tacitly collude* without meeting by signaling to each other through their actions. If one firm raises its price and keeps it high only if other firms follow its lead, it is not necessarily violating the law because the firms did not explicitly communicate.

We can't legally discuss price. However, look at how many sugar cubes I can stack!

For example, shortly before Thanksgiving in 2012, United Airlines announced a fare increase. However, when rivals failed to match this increase, United rolled back its fares the next day. Shortly thereafter, the president of US Airways observed that if Southwest Airlines, the firm that carries the most passengers, fails to match an increase by other airlines, rivals cancel the increase.[5]

Canada enacted the world's first antitrust statute in 1889, one year before the U.S. Sherman Act. As under U.S. law, price-fixing cartels are *per se* illegal and are subject to civil and criminal punishment. Australia and New Zealand have laws on cartels that are similar to those in Canada and the United States. In recent

[5]Charisse Jones, "United Airlines Hikes Fares; Will Rivals Follow?" *USA Today*, October 11, 2012; "US Airways President Talks about Southwest Fares," *Businessweek*, October 24, 2012.

years, the European Union and most developed countries have followed Canada and the United States in strictly prohibiting cartels.

The DOJ, the FTC, the Canadian Competition Bureau, and the European Union authorities have become increasingly aggressive, prosecuting many more cases and increasing fines dramatically. Increasingly, antitrust authorities from around the world are cooperating. Cooperation agreements exist between authorities in Canada, Mexico, Europe, Australia, New Zealand, and the United States, among others. Such cooperation is critical given the increasingly global scope of the firms engaged in collusion and other anticompetitive activities.

Application

The Apple-Google-Intel-Adobe-Intuit-Lucasfilms-Pixar Wage Cartel

How would you get a higher wage than your current employer is paying? Probably you'd try to get a job offer from another firm in the same field. But if that other firm refuses to hire anyone employed by your current firm, you're out of luck. That's what happened to many skilled engineers.

In 2005, when demand for Silicon Valley engineers was skyrocketing, Apple's Steve Jobs agreed on a secret, illegal "no-poaching" deal with Google's Eric Schmidt (who was also on Apple's board) to keep their employees' wages low by agreeing not to recruit each other's workers, by sharing wage scale information, and by punishing a firm that violated the agreement. Internal emails show that, in 2007, Schmidt, to keep Jobs happy, fired a Google recruiter who had contacted an Apple engineer. Intuit, Pixar, and Lucasfilm joined the cartel. It is alleged that many other major tech firms joined the conspiracy, affecting over a million employees.

Apparently, this cartel stopped when the U.S. Justice Department opened an investigation of Apple, Google, and other companies in 2010. In 2014, Intuit, Pixar, and Lucasfilm agreed to a $20 million settlement of a class-action lawsuit alleging that they conspired to suppress wages. In 2015, Apple, Google, Intel, and Adobe agreed to pay $415 million to settle a similar lawsuit. These companies have denied any wrongdoing!

This conspiracy is an example of a buyers' cartel. Recent sellers' cartels are alleged or confirmed to have occurred in airline, book publishing, French dairy, Balkan trains, computer chips, air cargo, LCD panels, potatoes (!), and many other markets.

Maintaining Cartels

To keep firms from violating the cartel agreement, the cartel must be able to detect cheating and punish violators. Further, the members of the cartel must keep their illegal behavior hidden from customers and government agencies.

Detection and Enforcement Cartels use various techniques to detect cheating. Some cartels, for example, give members the right to inspect each other's accounts. Cartels may also divide the market by region or by customers, making it more likely that the cartel knows if a firm steals another firm's customers, as in the two-country mercury cartel (1928–1972) that allocated the Americas to Spain and Europe to Italy. Another option is for a cartel to turn to industry organizations that collect data on market share by firm. A cheating cartel's market share would rise, tipping off the other firms that it cheated.

You probably have seen "low price" ads in which local retail stores guarantee to meet or beat the prices of any competitors. These ads may in fact be a way for the firm to induce its customers to report cheating by other firms on an explicit or implicit cartel agreement (Salop, 1986).

Cartels use various methods to enforce their agreements. In the past when GE and Westinghouse were the two major sellers of large steam-turbine generators, they included "most-favored-customer" clauses in their contracts. These contracts stated that the seller would not offer a lower price to any other current or future buyer without offering the same price decrease to firms that had already signed these contracts. This type of rebate clause creates a penalty for cheating on the cartel: If either company cheats by cutting prices, it has to lower prices to all previous buyers as well. Threats of violence are another means of enforcing a cartel agreement.[6]

Governments often enable cartels indirectly:

> **Common Confusion:** Requiring government agencies to report which company had the lowest bid for a government contract and the level of the bid is good for the public.

Although society benefits in many ways from government transparency, disclosing this type of information can help a cartel enforce its agreement. If the government reports that the "wrong" cartel member low bid allowed it to win the contract, then the other cartel members know immediately that a firm cheated on the cartel agreement. Electric equipment and heavy construction cartels have made use of this government information.

Government Support Sometimes governments help create and enforce cartels, exempting them from antitrust and competition laws. By successfully lobbying the U.S. Congress for a special exemption, professional baseball teams have not been subject to most U.S. antitrust laws since 1922. As a result, they can use the courts to help enforce certain aspects of their cartel agreement.

The international airline market provides an example where governments first created a cartel and then later acted to end it. In 1944, 52 countries signed the Convention on International Civil Aviation, which established rules ("freedoms") that enabled airlines to fly between countries. Bilateral government agreements determined international airfares, and exempted airlines from cartel laws, which allowed them to discuss prices through the International Air Transport Association (IATA). In the late 1970s, the United States deregulated its airline industry. Soon thereafter, European countries started to deregulate, allowing nongovernment-owned airlines to enter the market. Countries negotiated bilateral *open skies* agreements that weakened IATA's price-fixing role.[7]

Application

Cheating on the Maple Syrup Cartel

Most maple syrup comes from Quebec (not Vermont as many Americans assume). Quebec has many, many trees and about 13,500 maple syrup producers. How could they band together and effectively run a cartel? The provincial government passed a law half a century ago that created a cartel: the Federation of Quebec Maple Syrup Producers. Simon Trépanier, the federation's executive director, has referred to the federation as the OPEC of maple syrup.

Technological change, such as the use of plastic pipes, caused a large increase in supply and a drop in price. In response, a majority of the federation's members

[6]See the Application "Bad Bakers" in MyLab Economics, Chapter 11.

[7]The European Court of Justice struck down the central provisions of aviation treaties among the United States and eight other countries in 2002.

voted to establish mandatory production quotas starting in 2004, which limit how much farmers can sell in a year. Moreover, farmers have to sell all their syrup through the federation. Thus, the federation restricts supply to raise the price of maple syrup. The price rose 36% from 2004 to 2015 (13% in real terms).

So are all the farmers happy? According to Mr. Trépanier, “Three-quarters of our members are happy or very happy with what we are doing.” What about the rest? Some of them are “cheating” on the cartel.

If the federation suspects a farmer is producing and selling outside the federation, it posts guards on their properties. Then it seeks fines, or, in extreme circumstances, it seizes production. In other words, it has powers that illegal cartels can only envy.

But does the federation stop all cheating? It is in a battle with farmers like Robert Hodge who break the law by not participating in the federation’s system. The federation did not catch Mr. Hodge from 2004 through 2008. However, in 2009, when they detected his activities, the federation demanded C$278,000 from Mr. Hodge for not joining and selling outside the system, which exceeded his annual sales of about C$50,000 by more than fivefold.

In 2015, the federation hired guards to keep watch over Mr. Hodge’s sugar farm. After watching the farm for several weeks, they seized his entire annual production of 20,400 pounds of maple syrup, worth about C$60,000 ($46,000). He remains intransigent, contending that he should be free to choose how much to produce and to whom to sell regardless of the law. He says, “They call us rebels, say we’re in a sugar war or something.” His 20-year-old daughter observed, “A war over maple syrup, like how pathetic can you get?”

Mr. Hodge is not the only one fighting the cartel. Despite harsh penalties and legal backing, the cartel is conducting about 400 investigations at any given time.

Barriers to Entry Barriers to entry that limit the number of firms help the cartel detect and punish cheating and keep prices high. The fewer the firms in a market, the more likely it is that other firms will know if a given firm cheats and the easier it is to impose costs on that firm. Cartels with a large number of firms are relatively rare, except those involving professional associations. Hay and Kelley (1974) examined Department of Justice price-fixing cases from 1963 to 1972 and found that only 6.5% involved 50 or more conspirators, the average number of firms was 7.25, and nearly half the cases (48%) involved 6 or fewer firms.

When new firms enter their market, cartels frequently fail. For example, when only Italy and Spain sold mercury, they were able to establish and maintain a stable cartel. When a larger group of countries joined them, their attempts to cartelize the world mercury market repeatedly failed (MacKie-Mason and Pindyck, 1986).

Mergers

If antitrust or competition laws prevent firms from colluding, firms could try to achieve the same end by merging to form a monopoly. To prevent this potential problem, most antitrust and competition laws restrict the ability of firms to merge if the net effect is to harm society.

U.S. laws restrict the ability of firms to merge if the effect would be anticompetitive. In recent years, the European Commission has actively reviewed and blocked some mergers. For example in 2011, the DOJ and the European Commission blocked a proposed merger between the world’s two largest stock exchanges, the New York Stock Exchange and NASDAQ.

Whether a merger helps or harms society depends on which of its two offsetting effects—reducing competition and increasing efficiency—is larger. Consider two extreme cases. In one, the merger of the only two firms in the market does not lower costs, but it increases monopoly power, so the newly merged firm substantially raises prices to consumers. Here, the merger hurts society. At the other extreme, two of a large number of firms in the market merge, with substantial cost savings, but with no noticeable increase in market power or market price. Here, the merger benefits society.

The contentious cases lie in the middle, where market power increases significantly *and* costs decrease. However, if the price falls after the merger because the cost reduction effect dominates the increased market power effect, the merger is desirable.

Application

Mergers to Monopolize

Why should governments worry about mergers? If a merger creates market power, increasing price and profit, won't new firms enter the market? This argument makes sense only if firms can easily enter the market without incurring substantial sunk costs or other entry barriers. Thus, in deciding whether to permit a merger, governments consider the difficulty of entry.

Collard-Wexler (2014) provided an example of why entry is important in the ready-mix concrete industry. Although the United States has about 5,000 ready-mix concrete plants, few of them directly compete. Given high shipping costs, the country has at least 449 small, local markets. New plants must incur substantial sunk entry costs. He found that it takes nine to ten years for a new firm to enter the market following a merger to monopoly. Thus, if the only two firms in a local market merge, the resulting firm can earn monopoly profits for the better part of a decade, generating damages nearly eight times the damage from one year of monopoly.

13.3 Cournot Oligopoly

How do oligopolistic firms behave if they do not collude? Although economists have only one model of competition and one model of monopoly, they have many models of noncooperative oligopolistic behavior with many possible equilibrium prices and quantities.

Which model is appropriate to use depends on the type of *actions* firms take—such as setting quantity or price—and whether firms act simultaneously or sequentially. We examine the three best-known oligopoly models in turn. In the *Cournot model*, firms simultaneously choose quantities without colluding. In the *Stackelberg model*, a leader firm chooses its quantity and then the other follower firms independently choose their quantities. In the *Bertrand model*, firms simultaneously and independently choose prices.

To illustrate these models as simply and clearly as possible, we start by making four restrictive assumptions, which we later relax.

duopoly
an oligopoly with two firms

First, we examine an oligopoly model for a **duopoly**: an oligopoly with two (*duo*) firms. Each of these models applies to markets with many firms. The Cournot and Stackelberg outcomes vary, whereas the Bertrand market outcome with undifferentiated goods does not vary as the number of firms increases.

Second, we assume that all firms are identical in the sense that they have the same cost functions and produce identical, *undifferentiated* products. Then, we show how the market outcomes change if costs differ or if consumers believe that the products differ across firms.

Third, we assume in this section that the firms act simultaneously. In our discussion of the Stackelberg model, we change this assumption so that one firm acts before the other.

Fourth, we assume that the market lasts for only one period. Consequently, each firm chooses its quantity or price only once. In Chapter 14, we examine markets that last for more than one period.

To compare market outcomes under the various models, we need to be able to characterize the oligopoly equilibrium. In Chapter 2, we defined an *equilibrium* as a situation in which no one wants to change his or her behavior. For a competitive market to be in equilibrium, no firm wants to change its output level given what the other firms are producing. As oligopolistic firms may take many possible actions—such as setting price or quantity, or choosing a level of advertising—the oligopoly equilibrium rule needs to refer to their behavior more generally than just setting output.

John Nash, a Nobel Prize-winning economist and mathematician, defined an equilibrium concept that has wide applicability including to oligopoly models (Nash, 1951). We will give a general definition of a Nash equilibrium in Chapter 14. In this chapter, we use a special case of that definition. It is appropriate for single-period oligopoly models where the only action that a firm can take is to set either its quantity or its price: A set of actions taken by the firms is a *Nash equilibrium* if, holding the actions of all other firms constant, no firm can obtain a higher profit by choosing a different action.

The Duopoly Nash-Cournot Equilibrium

The French economist and mathematician Antoine-Augustin Cournot introduced the first formal model of oligopoly in 1838. Cournot explained how oligopolistic firms behave if they simultaneously choose how much they produce. The firms act independently and have imperfect information about their rivals. Each firm must choose its output level before knowing what the other firms will choose. The quantity one firm produces directly affects the profit of the other firms because the market price depends on total output. Thus, in choosing its strategy to maximize its profit, each firm takes into account its beliefs about the output its rivals will sell. Cournot introduced an equilibrium concept that is the same as the Nash definition where the action that firms take is to choose quantities.

To illustrate the basic idea of the Cournot model, we turn to an actual market where American Airlines and United Airlines compete for customers on flights between Chicago and Los Angeles.[8] The total number of passengers flown by these two firms, Q, is the sum of the number of passengers flown on American, q_A, and those flown on United, q_U. We assume that no other companies can enter, perhaps because they cannot obtain landing rights at both airports.[9]

Cournot equilibrium (Nash-Cournot equilibrium)
a set of quantities chosen by firms such that, holding the quantities of all other firms constant, no firm can obtain a higher profit by choosing a different quantity

How many passengers does each airline choose to carry? To answer this question, we determine the Nash equilibrium for this model. This Nash equilibrium, in which firms choose quantities, is also called a **Cournot equilibrium or Nash-Cournot equilibrium** (or *Nash-in-quantities equilibrium*): a set of quantities chosen by firms such that, holding the quantities of all other firms constant, no firm can obtain a higher profit by choosing a different quantity.

---

[8]This example is based on Brander and Zhang (1990). In calculating the profits, we assume that Brander and Zhang's estimate of the firms' constant marginal cost is the same as the firms' relevant long-run average cost.

[9]With the end of deregulation, existing firms had the right to buy, sell, or rent landing slots. By controlling landing slots, existing firms can make entry difficult.

To determine the Nash-Cournot equilibrium, we need to establish how each firm chooses its quantity. We start by using the total demand curve for the Chicago–Los Angeles route and a firm's belief about how much its rival will sell to determine its *residual demand curve*: the market demand that is not met by other sellers at any given price (Chapter 8). Next, we examine how a firm uses its residual demand curve to determine its best response: the output level that maximizes its profit, given its belief about how much its rival will produce. Finally, we use the information contained in the firms' best-response functions to determine the Nash-Cournot equilibrium quantities.

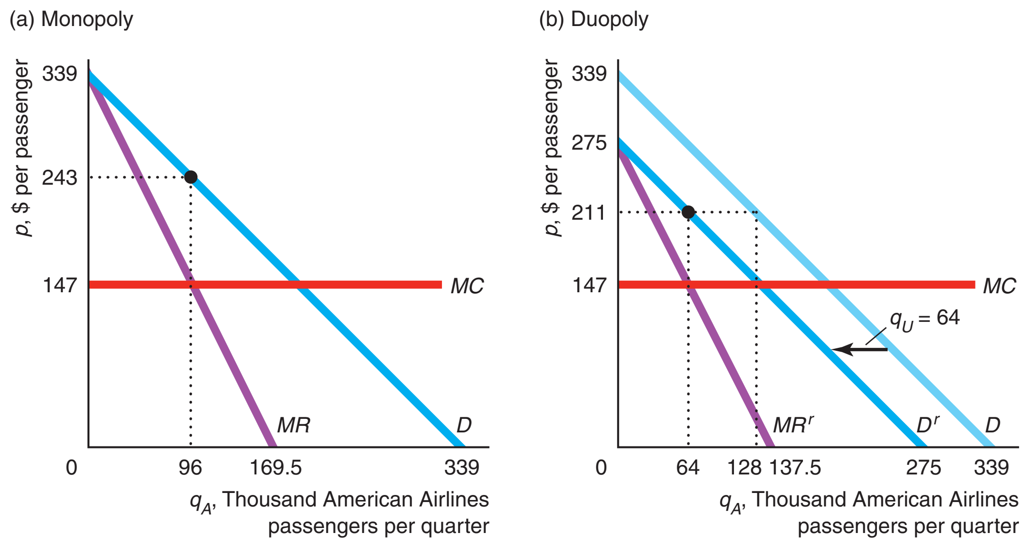
Graphical Approach The strategy that each firm uses depends on the demand curve it faces and its marginal cost. American Airlines' profit-maximizing output depends on how many passengers it believes United will fly. Figure 13.2 illustrates two possibilities.

If American were a monopoly, it wouldn't have to worry about United's strategy. American's demand would be the market demand curve, D, in panel a. To maximize its profit, American would set its output so that its marginal revenue curve, MR, intersected its marginal cost curve, MC, which is constant at \$147 per passenger. Panel a shows that the monopoly output is 96 units (thousands of passengers) per quarter and the monopoly price is \$243 per passenger (one way).

Because American competes with United, American must take account of United's behavior when choosing its profit-maximizing output. American's demand is not the entire market demand. Rather, American is concerned with its *residual demand curve* (Chapter 8): the market demand that is not met by other sellers at any given price.

Figure 13.2 American Airlines' Profit-Maximizing Output MyLab Economics Video

(a) If American is a monopoly, it picks its profit-maximizing output, $q_A = 96$ units (thousand passengers) per quarter, so that its marginal revenue, MR, equals its marginal cost, MC. (b) If American believes that United will fly $q_U = 64$ units per quarter, its residual demand curve, D^r, is the market demand curve, D, minus q_U. American maximizes its profit at $q_A = 64$, where its marginal revenue, MR^r, equals MC.



In general, if the market demand function is $D(p)$, and the supply of other firms is $S^o(p)$, then the residual demand function, $D^r(p)$, is

$$D^r(p) = D(p) - S^o(p).$$

Thus, if United flies q_U passengers regardless of the price, American transports only the residual demand, $Q = D(p)$, minus the q_U passengers, so $q_A = Q - q_U$.

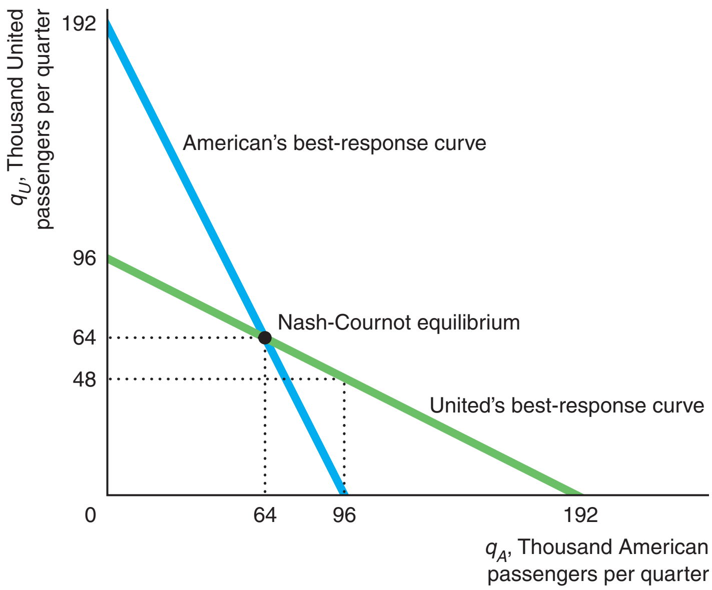
Suppose that American believes that United will fly $q_U = 64$. Panel b shows that American's residual demand curve, D^r, is the market demand curve, D, moved to the left by $q_U = 64$. For example, if the price is \$211, the total number of passengers who want to fly is $Q = 128$. If United transports $q_U = 64$, American flies $Q - q_U = 128 - 64 = 64 = q_A$.

What is American's best-response, profit-maximizing output if its managers believe that United will fly q_U passengers? American can think of itself as having a monopoly with respect to the people who don't fly on United, which its residual demand curve, D^r, shows. To maximize its profit, American sets its output so that its marginal revenue corresponding to this residual demand, MR^r, equals its marginal cost. Panel b shows that if $q_U = 64$, American's best response is $q_A = 64$.

By shifting its residual demand curve appropriately, American can calculate its best response to any given q_U using this type of analysis. Figure 13.3 plots American Airlines' best-response curve, which shows how many tickets American sells for each possible q_U.[10] As this curve shows, American will sell the monopoly number of tickets, 96, if American thinks United will fly no passengers, $q_U = 0$. The negative slope of the best-response curve shows that American sells fewer tickets the more people American thinks that United will fly. American sells $q_A = 64$ if it thinks q_U will be 64. American shuts down, $q_A = 0$, if it thinks q_U will be 192 or more, because operating wouldn't be profitable.

Figure 13.3 American's and United's Best-Response Curves MyLab Economics Video

The best-response curves show the output each firm picks to maximize its profit, given its belief about its rival's output. The Nash-Cournot equilibrium occurs at the intersection of the best-response curves.



[10]*Jargon alert*: Some economists refer to the best-response curve as the *reaction curve*.

Similarly, United's best-response curve shows how many tickets United sells if it thinks American will sell q_A. For example, United sells $q_U = 0$ if it thinks American will sell $q_A = 192$, $q_U = 48$ if $q_A = 96$, $q_U = 64$ if $q_A = 64$, and $q_U = 96$ if $q_A = 0$.

A firm wants to change its behavior if it is selling a quantity that is not on its best-response curve. In a Nash-Cournot equilibrium, neither firm wants to change its behavior. Thus, in a Nash-Cournot equilibrium, each firm is on its best-response curve: Each firm is maximizing its profit, given its correct belief about its rival's output.

These firms' best-response curves intersect at $q_A = q_U = 64$. If American expects United to sell $q_U = 64$, American wants to sell $q_A = 64$. Because this point is on its best-response curve, American doesn't want to change its output from 64. Similarly, if United expects American to sell $q_A = 64$, United doesn't want to change q_U from 64. Thus, this pair of outputs is a Cournot (Nash) equilibrium: Given its correct belief about its rival's output, each firm is maximizing its profit, and neither firm wants to change its output.

Any pair of outputs other than the pair at an intersection of the best-response functions is *not* a Nash-Cournot equilibrium. If either firm is not on its best-response curve, it changes its output to increase its profit. For example, the output pair $q_A = 96$ and $q_U = 0$ is not a Nash-Cournot equilibrium. American is perfectly happy producing the monopoly output if United doesn't operate at all: American is on its best-response curve. United, however, would not be happy with this outcome because it is not on United's best-response curve. As its best-response curve shows, if it knows that American will sell $q_A = 96$, United wants to sell $q_U = 48$. Only at $q_A = q_U = 64$ does neither firm want to change its behavior.

Algebraic Approach We can also use algebra to solve for the Nash-Cournot equilibrium for these two airlines. (See Appendix 13A for a more general analysis.) We use estimates of the market demand and firms' marginal costs to determine the equilibrium.

Our estimate of the market demand function is

$$Q = 339 - p, \tag{13.1}$$

where price, p, is the dollar cost of a one-way flight, and total quantity of the two airlines combined, Q, is measured in thousands of passengers flying one way per quarter. Panels a and b of Figure 13.2 show that this market demand curve, D, is a straight line that hits the price axis at \$339 and the quantity axis at 339 units (thousands of passengers) per quarter. Each airline has a constant marginal cost, MC, and average cost, AC, of \$147 per passenger per flight. Using only this information and our economic model, we can find the Nash-Cournot equilibrium for the two airlines.

If American believes that United will fly q_U passengers, American expects to fly only the total market demand minus q_U passengers. At a price of p, the total number of passengers, $Q(p)$, is given by the market demand function, Equation 13.1. Thus, the residual demand American faces is

$$q_A = Q(p) - q_U = (339 - p) - q_U.$$

Using algebra, we can rewrite this inverse residual demand function as

$$p = 339 - q_A - q_U. \tag{13.2}$$

In panel b, the linear residual demand, D^r, is parallel to the market demand, D, and lies to the left of D by $q_U = 64$.

If a demand curve is linear, the corresponding marginal revenue curve is twice as steep (Chapter 11). The slope of the residual demand curve, Equation 13.2, is $\Delta p/\Delta q_A = -1$, so the slope of the corresponding marginal revenue curve, MR^r in panel b in Figure 13.2, is -2. Therefore, the marginal revenue function is[11]

$$MR^r = 339 - 2q_A - q_U. \tag{13.3}$$

American Airlines' best response—its profit-maximizing output, given q_U—is the output that equates its marginal revenue, Equation 13.3, and its marginal cost:

$$MR^r = 339 - 2q_A - q_U = 147 = MC. \tag{13.4}$$

By rearranging Equation 13.4, we can write American's best-response output, q_A, as a function of q_U:

$$q_A = 96 - \tfrac{1}{2}q_U. \tag{13.5}$$

Figure 13.3 shows American's best-response function, Equation 13.5. According to this best-response function, $q_A = 96$ if $q_U = 0$ and $q_A = 64$ if $q_U = 64$. By the same reasoning, United's best-response function is

$$q_U = 96 - \tfrac{1}{2}q_A. \tag{13.6}$$

A Nash-Cournot equilibrium is a pair of quantities, q_A and q_U, such that Equations 13.5 and 13.6 both hold: Each firm is on its best-response curve. This statement is equivalent to saying that the Nash-Cournot equilibrium is a point at which the best-response curves cross.

One way to determine the Nash-Cournot equilibrium is to substitute Equation 13.6 into Equation 13.5,

$$q_A = 96 - \tfrac{1}{2}(96 - \tfrac{1}{2}q_A),$$

and solve for q_A. Doing so, we find that $q_A = 64$ is the Nash-Cournot equilibrium quantity for American. Substituting $q_A = 64$ into Equation 13.6, we find that $q_U = 64$ is the Nash-Cournot equilibrium quantity for United. As a result, the total output in the Nash-Cournot equilibrium is $Q = q_A + q_U = 128$. Setting $Q = 128$ in the market demand Equation 13.1, we learn that the Nash-Cournot equilibrium price is \$211.

Equilibrium, Elasticity, and the Number of Firms

Our airlines example illustrates that, if two Cournot firms set output independently, the price to consumers is lower than the monopoly (or cartel) price. The price to consumers is even lower if more than two Cournot firms produce independently, as we now show.

Each Cournot firm maximizes its profit by operating where its marginal revenue equals its marginal cost. A firm's marginal revenue depends on the price and the elasticity of demand it faces where it maximizes its profit (Chapter 11). The marginal revenue for a typical Cournot firm is $MR = p(1 + 1/{}_r)$, where ${}_r$ is the elasticity of the residual demand curve the firm faces. Appendix 13A shows that ${}_r = n$, where is the market elasticity of demand and n is the number of firms

[11] American's revenue is $R = pq_A = (339 - q_A - q_U)q_A$. If American treats q_U as a constant and differentiates R with respect to its output, it finds that its marginal revenue is $MR = \partial R/\partial q_A = 339 - 2q_A - q_U$.

with identical costs. Thus, we can write a typical Cournot firm's profit-maximizing condition as

$$MR = p\left(1 + \frac{1}{n}\right) = MC. \tag{13.7}$$

If $n = 1$, the Cournot firm is a monopoly, and Equation 13.7 is the same as the profit-maximizing monopoly condition, Equation 11.7. The more firms, the larger the residual demand elasticity, n , that a single firm faces. As n grows very large, the residual demand elasticity approaches negative infinity ($-\infty$), and Equation 13.7 becomes $p = MC$, which is the profit-maximizing condition of a price-taking competitive firm.

The Lerner Index, $(p - MC)/p$, is a measure of market power: the firm's ability to raise price above marginal cost. By rearranging the terms in Equation 13.7, we find that a Cournot firm's Lerner Index depends on the elasticity the firm faces:

$$\frac{p - MC}{p} = -\frac{1}{n}. \tag{13.8}$$

Thus, a Cournot firm's Lerner Index equals the monopoly level, $-1/$, with only one firm: Setting $n = 1$ in Equation 13.8, we obtain the monopoly expression (Equation 11.9). Again, as the number of firms grows large, the residual demand elasticity a firm faces approaches $-\infty$, so the Lerner Index approaches zero, which is the same as with price-taking, competitive firms.

We can illustrate these results using our airlines example. Suppose that other airlines with identical marginal cost, $MC = \$147$, were to fly between Chicago and Los Angeles. Table 13.2 shows how the Nash-Cournot equilibrium price and the Lerner Index vary with the number of firms.[12]

As we already know, one "Cournot" firm would produce the monopoly quantity, 96, at the monopoly price, \$243. We also know that each duopoly firm's output is 64, so market output is 128 and price is \$211. The duopoly market elasticity is

Table 13.2 Nash-Cournot Equilibrium Varies with the Number of Firms

Number of Firms, n	Firm Output, q	Market Output, Q	Price, p	Market Elasticity,	Residual Demand Elasticity, n ,	Lerner Index, $(p - m)/p = -1/(n\ )$
1	96	96	243	−2.53	−2.53	0.40
2	64	128	211	−1.65	−3.30	0.30
3	48	144	195	−1.35	−4.06	0.25
4	38.4	154	185.40	−1.21	−4.83	0.21
5	32	160	179	−1.12	−5.59	0.18
10	17.5	175	164.45	−0.94	−9.42	0.11
50	3.8	188	150.76	−0.80	−40.05	0.02
100	1.9	190	148.90	−0.78	−78.33	0.01

[12]In Appendix 13A, we derive the Nash-Cournot equilibrium quantity and price for a general linear demand. Given our particular demand curve, Equation 13.1, and marginal cost, \$147, each firm's Nash-Cournot equilibrium output is $q = (339 - 147)/(n + 1) = 192/(n + 1)$ and the Cournot market price is $p = (339 + 147n)/(n + 1)$.

$= -1.65$, so the residual demand elasticity each firm faces is twice as large as the market elasticity, $2 \;\; = -3.3$.

As the number of firms increases, each firm's output falls toward zero, but total output approaches 192, the quantity on the market demand curve where price equals the marginal cost of \$147. Although the market elasticity of demand falls as the number of firms grows, the residual demand curve for each firm becomes increasingly horizontal (perfectly elastic). As a result, the price approaches the marginal cost, \$147. Similarly, as the number of firms increases, the Lerner Index approaches the price-taking level of zero.[13]

The table shows that having extra firms in the market benefits consumers. When the number of firms rises from 1 to 4, the price falls by a quarter and the Lerner Index is cut nearly in half. At ten firms, the price is one-third less than the monopoly level, and the Lerner Index is a quarter of the monopoly level.

Application

Mobile Number Portability

An inability of consumers to switch from an old provider of a service to a new provider reduces the competitive effect of entry of new firms. In most European countries, a monopoly initially provided mobile phone service. After governments opened their mobile phone markets to new entrants, customers were slow to switch firms because of large switching costs such as having to obtain a new phone number and get new handsets. Preventing customers from transferring their phone number if they switch carriers makes the demand curve facing a given firm less elastic.

To reduce switching costs and increase competition by new firms, many governments require Mobile Number Portability (MNP), which allows consumers to move their phone number to another mobile phone carrier.[14] Cho et al. (2015) estimated that the introduction of MNP in European countries—and the increase in effective competitors—decreased phone service prices by 7.9% and increased consumer surplus by 2.86€ (\$3.07) per person per quarter.

Nonidentical Firms

We initially assumed that the firms were identical in the sense that they faced the same cost functions and produced identical products. However, costs often vary across firms, and firms often differentiate the products they produce from those of their rivals.

Unequal Costs In the Cournot model, the firm sets its output to equate its marginal revenue to its marginal cost, which determines its best-response function. If the firms' marginal costs vary, then the firms' best-response functions will as well. In the resulting Nash-Cournot equilibrium, the relatively low-cost firm produces more. As long as the products are undifferentiated, they both charge the same price.

We can illustrate the effect of unequal costs using our duopoly airlines example. Suppose that American Airlines' marginal cost remains at \$147, but United's marginal cost drops to \$99. The Cournot model still applies, but we have relaxed the assumption that the firms have identical costs. How does the Nash-Cournot equilibrium change? Your intuition probably tells you that United's output increases relative to that of American, as we now show.

[13]As the number of firms goes to infinity, the Nash-Cournot equilibrium goes to perfect competition only if average cost is nondecreasing (Ruffin, 1971).

[14]The United States has required wireless local number portability nationwide since 2003, and Canada since 2007. In 2002, the European Commission mandated that each European Community country enact MNP. At least 73 countries have MNP as of 2016.

Nothing changes for American, so its best-response function is unchanged. United's best response to any given American output is the output at which its marginal revenue corresponding to its residual demand, MR^r, equals its new, lower marginal cost. Because United's marginal cost curve fell, United wants to produce more than before for any given level of American's output.

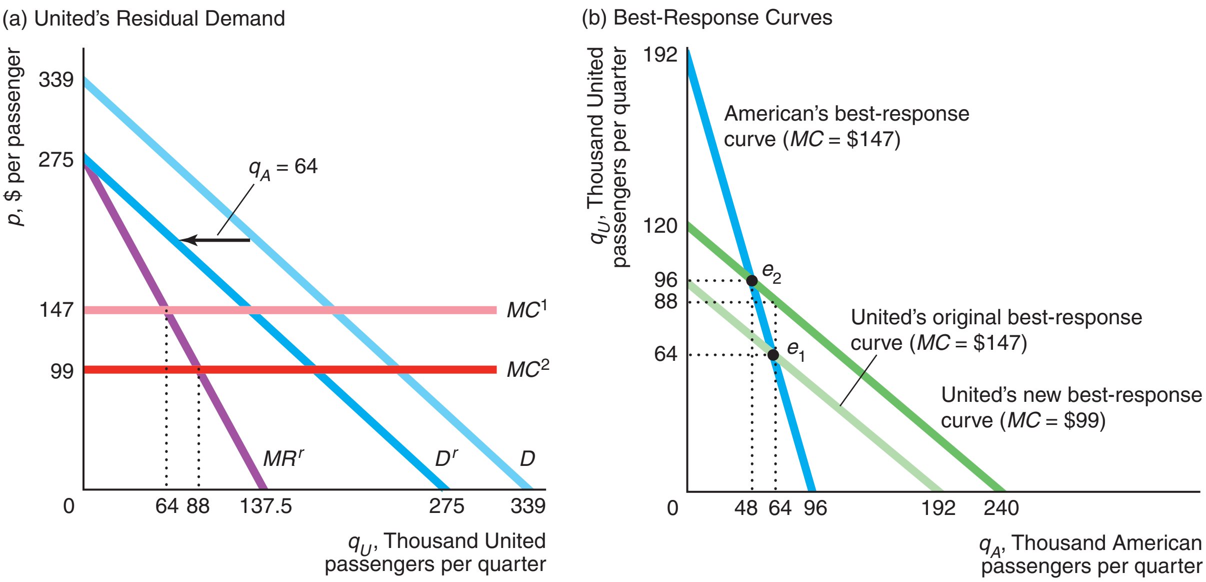
Panel a of Figure 13.4 illustrates this reasoning. United's MR^r curve is unaffected, but its marginal cost curve shifts down from MC^1 to MC^2. Suppose we fix American's output at 64 units. Consequently, United's residual demand, D^r, lies 64 units to the left of the market demand, D. United's corresponding MR^r curve intersects its original marginal cost curve, $MC^1 = \$147$, at 64 and its new marginal cost, $MC^2 = \$99$, at 88. Thus, if we hold American's output constant at 64, United produces more as its marginal cost falls.

Because this reasoning applies for any level of output American picks, United's best-response function in panel b shifts outward as its marginal cost falls. United's best response to any given quantity that American sells is to sell more than at its previous, higher cost. As a result, the Nash-Cournot equilibrium shifts from e_1, at which both firms sold 64, to e_2, at which United sells 96 and American sells 48.

Using the market demand curve, Equation 13.1, we find that the market price falls from \$211 to \$195, benefiting consumers. United's profit increases from \$4.1 million to \$9.2 million, while American's profit falls to \$2.3 million.[15] Thus, United and consumers gain and American loses from the fall in United's marginal cost.

Figure 13.4 Effect of a Drop in One Firm's Marginal Cost on a Duopoly Nash-Cournot Equilibrium
MyLab Economics Video

(a) United's marginal cost falls from $MC^1 = \$147$ to $MC^2 = \$99$. If American produces $q_a = 64$, United's best response is to increase its output from $q_U = 64$ to 88 given its lower marginal cost. (b) If both airlines' marginal cost is \$147, the Nash-Cournot equilibrium is e_1. After United's marginal cost falls to \$99, its best-response function shifts outward. It now sells more tickets in response to any given American output than previously. At the new Nash-Cournot equilibrium, e_2, United sells $q_U = 96$, while American sells only $q_A = 48$.



[15] Each firm's profit per passenger is price minus average cost, $p - AC$, so the firm's profit is $\pi = (p - AC)q$, where q is the number of passengers the firm flies. The Cournot price is \$211 and the average cost is \$147, so the Cournot profit per firm is $\pi = (211 - 147) \times 64$ units per quarter $= \$4.1$ million per quarter in the original symmetric case.

Solved Problem 13.1

MyLab Economics Solved Problem

Derive United Airlines' best-response function if its marginal cost falls to \$99 per unit due to a government subsidy. Given that American's marginal cost does not change, what is the new Nash-Cournot equilibrium?[16]

Answer

1. *Determine United's post-subsidy marginal revenue function corresponding to its residual demand curve*. Luckily, we already know that. The shift in its marginal cost curve does not affect United's residual demand curve, hence its marginal revenue function is the same as before: $MR^r = 339 - 2q_U - q_A$. (The same expression as American's marginal revenue function in Equation 13.3, where the A and U subscripts are reversed.)

2. *Equate United's post-subsidy marginal revenue function and its marginal cost to determine its best-response function*. For a given level of American's output, q_A, United chooses its output, q_U, to equate its marginal revenue and its marginal cost, m:

$$MR^r = 339 - 2q_U - q_A = 99 = m.$$

We can use algebra to rearrange this expression for its best-response function to express q_U as a function of q_a:

$$q_U = 120 - \tfrac{1}{2}q_A.$$

This equation corresponds to the dark green *new best-response curve* in panel b of Figure 13.4.

3. *Find the new Nash-Cournot equilibrium pair of quantities by solving the two best-response functions for q_A and q_U*. Because American Airlines' marginal cost is unchanged, its best-response function is the same as in Equation 13.5, $q_A = 96 - \tfrac{1}{2}q_U$. The intersection of that best-response function and United's new best-response function determines the new Nash-Cournot equilibrium. To find it, we substitute the expression for q_A from American's best-response function into United's best-response function:

$$q_U = 120 - \tfrac{1}{2}q_A = 120 - \tfrac{1}{2}(96 - \tfrac{1}{2}q_U).$$

Solving, we find that $q_U = 96$. Substituting $q_U = 96$ into either best-response function, we learn that $q_A = 48$.

Differentiated Products By differentiating its product from those of a rival, an oligopolistic firm can shift its demand curve to the right and make it less elastic. The less elastic the demand curve, the more the firm can charge. Loosely speaking, consumers are willing to pay more for a product that they perceive as being superior.

One way to differentiate a product is to give it unique, "desirable" attributes, such as the Lexus car that parks itself. In 2010, Kimberly-Clark introduced a new Huggies disposable diaper with a printed denim pattern, including seams and back pockets, which sent their sales up 15%. A firm can differentiate its product by advertising, using colorful labels, and engaging in other promotional activities to convince consumers that its product is superior in some (possibly unspecified) way even though it is virtually identical to its rivals physically or chemically. Economists call this practice

[16]See Appendix 13A for a mathematical approach to a more general case.

spurious differentiation. Bayer charges more for its chemically identical aspirin than other brands because Bayer has convinced consumers that its product is safer or superior in some other way. Clorox's bottle may be superior, but the bleach inside is chemically identical to that from rival brands costing much less.

Because differentiation makes demand curves less elastic, price markups over marginal cost are usually higher when products are differentiated than when they're identical. We know that consumer surplus falls as the gap between price and marginal cost rises for a given good. Does it follow that differentiating products lowers total surplus? Not necessarily. Although differentiation leads to higher prices, which harms consumers, differentiation is desirable in its own right. Consumers value having a choice, and some may greatly prefer a new brand to existing ones.

If consumers think products differ, the Cournot quantities and prices will differ across firms. Each firm faces a different inverse demand function and hence charges a different price. For example, suppose that Firm 1's inverse demand function is $p_1 = a - b_1q_1 - b_2q_2$, where $b_1 > b_2$ if consumers believe that Good 1 is different from Good 2 and $b_1 = b_2 = b$ if the goods are identical. Given that consumers view the products as differentiated and Firm 2 faces a similar inverse demand function, we replace the single market demand with these individual demand functions in the Cournot model. Solved Problem 13.2 shows how to solve for the Nash-Cournot equilibrium in an actual market with differentiated products.

Solved Problem 13.2

MyLab Economics Solved Problem

Intel and Advanced Micro Devices (AMD) are the only two firms that produce central processing units (CPUs)—the brains—for personal computers. Both because the products differ physically and because Intel's "Intel Inside" advertising campaign has convinced some consumers of its superiority, consumers view their CPUs as imperfect substitutes. Consequently, the two firms' estimated inverse demand functions differ:

$$p_A = 197 - 15.1q_A - 0.3q_I, \quad (13.9)$$

$$p_I = 490 - 10q_I - 6q_A, \quad (13.10)$$

where price is dollars per CPU, quantity is in millions of CPUs, the subscript I indicates Intel, and the subscript A represents AMD.[17] Each firm faces a constant marginal cost of $m = \$40$ per unit. (For simplicity, we assume that firms have no fixed costs.) Solve for the Nash-Cournot equilibrium quantities and prices.

Answer

1. *Using our rules for determining the marginal revenue for linear demand functions, calculate each firm's marginal revenue function*. For a linear demand curve,

[17] I thank Hugo Salgado for estimating these inverse demand functions for me and for providing evidence that this market is well described by a Nash-Cournot equilibrium.

we know that the marginal revenue curve is twice as steeply sloped as is the demand curve. Thus, the marginal revenue functions that correspond to the inverse demand Equations 13.9 and 13.10 are[18]

$$MR^A = 197 - 30.2q_A - 0.3q_I, \tag{13.11}$$

$$MR^I = 490 - 20q_I - 6q_A. \tag{13.12}$$

2. *Equate the marginal revenue functions to the marginal cost to determine the best-response functions*. We determine AMD's best-response function by equating MR^A from Equation 13.11 to its marginal cost of $m = \$40$,

$$MR^A = 197 - 30.2q_A - 0.3q_I = 40 = m,$$

and solving for q_A to obtain AMD's best-response function:

$$q_A = \frac{157 - 0.3\,q_I}{30.2}. \tag{13.13}$$

Similarly, Intel's best-response function is

$$q_I = \frac{450 - 6q_A}{20}. \tag{13.14}$$

3. *Use the best-response functions to solve for the Nash-Cournot equilibrium*. By simultaneously solving the system of best-response functions 13.13 and 13.14, we find that the Nash-Cournot equilibrium quantities are $q_A = 15{,}025/3{,}011 \approx 5$ million CPUs, and $q_I = 63{,}240/3{,}011 \approx 21$ million CPUs. Substituting these values into the inverse demand functions in Equations 13.9 and 13.10, we obtain the corresponding prices: $p_A = \$115.20$ and $p_I = \$250$ per CPU.

13.4 Stackelberg Oligopoly

In the Cournot model, both firms make their output decisions at the same time. Suppose, however, that one of the firms, called the *leader*, can set its output before its rival, the *follower*, sets its output. Having one firm act before another arises naturally if one firm enters the market first. Past examples of leaders include IBM in the mainframe computer market, General Electric in the turbine generator industry, and General Foods among coffee roasters.

Would the firm that acted first have an advantage? Heinrich von Stackelberg showed how to modify the Cournot model to answer this question.

How does the leader decide to set its output? The leader realizes that once it sets its output, the rival firm will use its Cournot best-response curve to pick a best-response output. Thus, the leader predicts what the follower will do before the follower acts. Using this knowledge, the leader manipulates the follower, benefiting at the follower's expense.

We illustrate this model graphically using our airlines market example (Appendix 13B analyzes the model mathematically). Although it is difficult to imagine that either American Airlines or United Airlines actually has an advantage that would allow it to act before its rival, we assume (arbitrarily) that American Airlines can act before United Airlines.

---

[18]We can use calculus to derive these marginal revenue functions. For example, by multiplying both sides of AMD's inverse demand function (Equation 13.9) by q_A, we learn that its revenue function is $R_A = p_A q_A = 197q_A - 15.2(q_A)^2 - 0.3q_I q_A$. Holding q_I constant and differentiating with respect to q_A, we obtain $MR^A = \mathrm{d}R_A/\mathrm{d}q_A = 197 - 30.2q_A - 0.3q_I$.

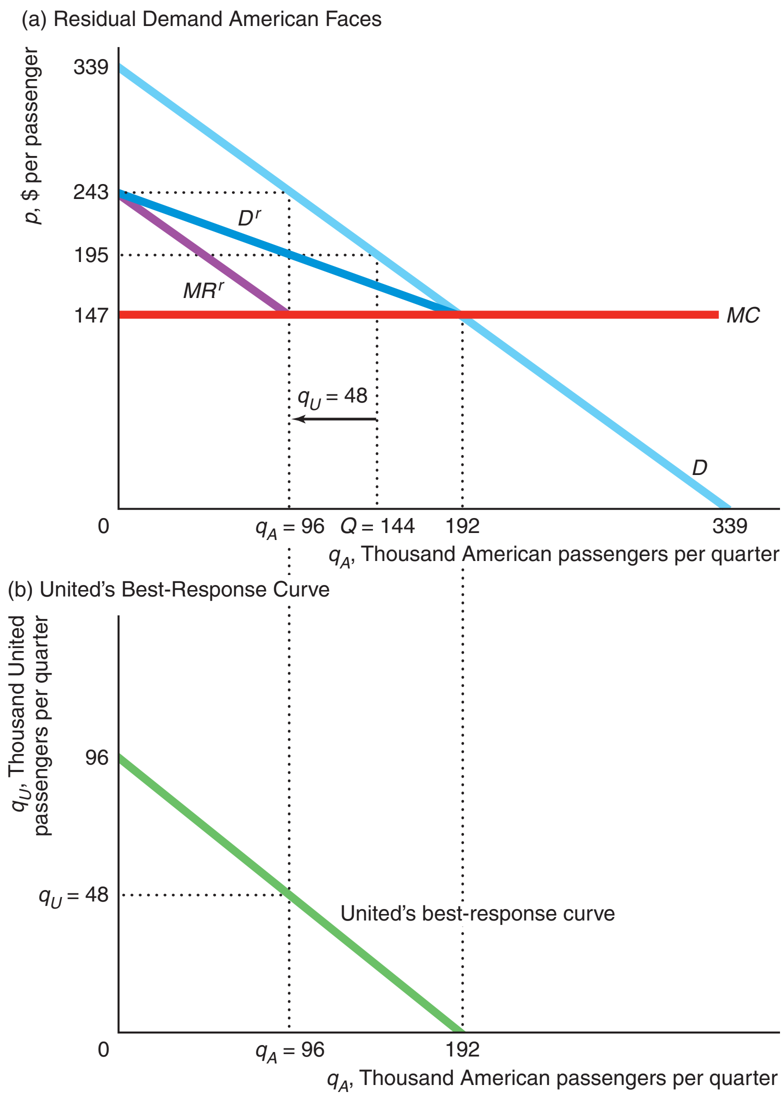
Graphical Model

Given that American Airlines chooses its output first, how does American decide on its optimal policy? American uses its residual demand curve to determine its profit-maximizing output. American knows that when it sets q_A, United will use its Cournot best-response function to pick its best-response q_U. Thus, American's residual demand curve, D^r (panel a of Figure 13.5), is the market demand curve, D (panel a), minus the output United will produce, as summarized by United's best-response curve (panel b). For example, if American sets $q_A = 192$, United's best response is $q_U = 0$ (as United's best-response curve in panel b shows). As a result, the residual demand curve and the market demand curve are identical at $q_A = 192$ (panel a).

Similarly, if American set $q_A = 0$, United would choose $q_U = 96$, so the residual demand at $q_A = 0$ is 96 less than demand. The residual demand curve hits the vertical axis, where $q_A = 0$, at $p = \$243$, which is 96 units to the left of demand at that

Figure 13.5 Nash-Stackelberg Equilibrium MyLab Economics Video

(a) The residual demand the Stackelberg leader faces is the market demand minus the quantity produced by the follower, q_U, given the leader's quantity, q_A. The leader chooses $q_A = 96$ so that its marginal revenue, MR^r, equals its marginal cost. The total output, $Q = 144$, is the sum of the output of the two firms. (b) The quantity the follower produces is its best response to the leader's output, as given by its Cournot best-response curve.



price. When $q_A = 96$, $q_U = 48$, so the residual demand at $q_A = 96$ is 48 units to the left of the demand curve where $p = \$195$.

The marginal revenue curve, MR^r, that corresponds to this residual demand curve hits the vertical axis at the same price, \$243, as the residual demand curve, and has twice as steep a slope.

American chooses its profit-maximizing output, $q_A = 96$, where its marginal revenue curve, MR^r, equals its marginal cost, \$147. At $q_A = 96$, the price, which is the height of the residual demand curve, is \$195. Total demand at \$195 is $Q = 144$. At that price, United produces $q_U = Q - q_A = 48$, its best response to American's output of $q_A = 96$.

Thus, in this Nash-Stackelberg equilibrium, the leader produces twice as much as the follower, as Figure 13.5 shows.[19] The total Stackelberg output, 144, is greater than the total Cournot, 128, output. As a result, the Stackelberg price, \$195, is less than the Cournot price, \$211. Thus, consumers prefer the Nash-Stackelberg equilibrium to the Nash-Cournot equilibrium.

The Stackelberg leader earns \$4.6 million, which is more than it could earn in a Nash-Cournot equilibrium, \$4.1 million. Total Stackelberg profit is less than total Cournot profit because the Stackelberg follower, earning \$2.3 million, is much worse off than in the Nash-Cournot equilibrium.

Solved Problem 13.3

Use algebra to solve for the Nash-Stackelberg equilibrium quantities and market price if American Airlines were a Stackelberg leader and United Airlines were a follower. (*Hint*: As the graphical analysis shows, American Airlines, the Stackelberg leader, maximizes its profit as though it were a monopoly facing a residual demand function.)

Answer

1. *Determine the inverse residual demand function facing American Airlines*. The residual demand function facing American Airlines is the market demand function (Equation 13.1), $Q = 339 - p$, minus the best-response function of United Airlines (Equation 13.6), $q_U = 96 - \frac{1}{2}q_A$:

$$q_A(p) = Q(p) - q_U(q_A) = 339 - p - [96 - \tfrac{1}{2}q_A] = 243 - p + \tfrac{1}{2}q_A. \quad (13.15)$$

Using algebra, we can rewrite Equation 13.15 as the inverse residual demand function (which is the D^r line in panel a of Figure 13.5):

$$p = 243 - \tfrac{1}{2}q_A. \quad (13.16)$$

2. *Solve for American Airlines' profit-maximizing output by equating its marginal revenue and marginal cost*. American Airlines, the Stackelberg leader, acts like a monopoly with respect to its residual demand. From Chapter 11, we know that its marginal revenue function is the same as its inverse residual demand function, Equation 13.16, except it has twice the slope: $MR_A = 243 - q_A$ (which is the MR^r line in panel a of Figure 13.5). To maximize its profit, American Airlines picks its output to equate its marginal revenue to its marginal cost:

$$MR_A = 243 - q_A = 147 = MC. \quad (13.17)$$

Solving Equation 13.17 for American Airlines' output, we find that $q_A = 96$.

---

[19]Here the leader produces the same quantity as a monopoly would, and the follower produces the same quantity as it would in the cartel equilibrium. These relationships are due to the linear demand curve and the constant marginal cost—they do not hold more generally.

3. *Use United Airlines' best-response function to solve for its quantity and the total output*. Substituting $q_A = 96$ into United Airlines' best-response function, Equation 13.6, we learn that United Airlines sells half as many seats as American Airlines: $q_U = 96 - \frac{1}{2}q_A = 48$. Thus, total output is

$$Q = q_A + q_U = 96 + 48 = 144.$$

4. *Use the market demand function to solve for the market price*. Substituting $Q = 144$, total output, into the market demand function, we determine that the market price is \$195.

Comment: See Appendix 13B for the solution to a more general case.

Why Moving Sequentially Is Essential

Why don't we get the Nash-Stackelberg equilibrium when both firms move simultaneously? Why doesn't one firm—say, American—announce that it will produce the Stackelberg leader output to induce United to produce the Stackelberg follower output level? The answer is that when the firms move simultaneously, United doesn't view American's warning that it will produce a large quantity as a *credible threat*.

If United believed that threat, it would indeed produce the Stackelberg follower output level. But United doesn't believe the threat because it is not in American's best interest to produce that large a quantity of output. If American were to produce the leader level of output and United produced the Cournot level, American's profit would be lower than if it too produced the Cournot level. Because American cannot be sure that United will believe its threat and reduce its output, American produces the Cournot output level.

Indeed, each firm may make the same threat and announce that it wants to be the leader. Because neither firm can be sure that the other will be intimidated and produce the smaller quantity, both produce the Cournot output level. In contrast, when one firm moves first, its threat to produce a large quantity is credible because it has already *committed* to producing the larger quantity, carrying out its threat.

Comparison of Competitive, Stackelberg, Cournot, and Collusive Equilibria

The Nash-Cournot and Nash-Stackelberg equilibrium quantities, prices, consumer surplus, and profits lie between those for the competitive (price-taking) and collusive equilibria, as Table 13.3 summarizes. In our airlines example, if the firms were to act as price takers, they would each produce where their residual demand curve intersects their marginal cost curve, so price would equal the marginal cost of \$147. The price-taking equilibrium is $q_A = q_U = 96$.

If American and United were to collude, they would maximize joint profits by producing the monopoly output, $q_A + q_U = 96$. Colluding airlines could split the monopoly quantity in many ways. American could act as a monopoly and serve all the passengers, $q_A = 96$ and $q_U = 0$, and give United some of the profits. Or they could reverse roles so that United served everyone, $q_A = 0$ and $q_U = 96$. Or the two airlines could share the passengers in any combination such that the sum of the airlines' passengers equals the monopoly quantity, or, equivalently,

$$q_U = 96 - q_A. \tag{13.18}$$

Figure 13.6 plots Equation 13.18, the set of possible collusive outcomes, which it labels the *Contract curve*. In the figure, we assume that the collusive firms split the market equally so that $q_A = q_U = 48$.

Table 13.3 Comparison of the Duopoly Airline Competitive, Stackelberg, Cournot, and Collusive Equilibria

	Competition	Stackelberg	Cournot	Collusion
Total output, Q (thousands)	192	144	128	96
Price, p ($)	147	195	211	243
Consumer Surplus ($ million)	18.4	10.4	8.2	4.6
Profit, π ($ million)	0	6.9	8.2	9.2

The quantities produced by both the firms are 192 (= 96 + 96) in the competitive equilibrium, 128 (= 64 + 64) in the Nash-Cournot equilibrium, 144 (= 96 + 48) in the Stackelberg equilibrium, and 96 in the collusive outcome.

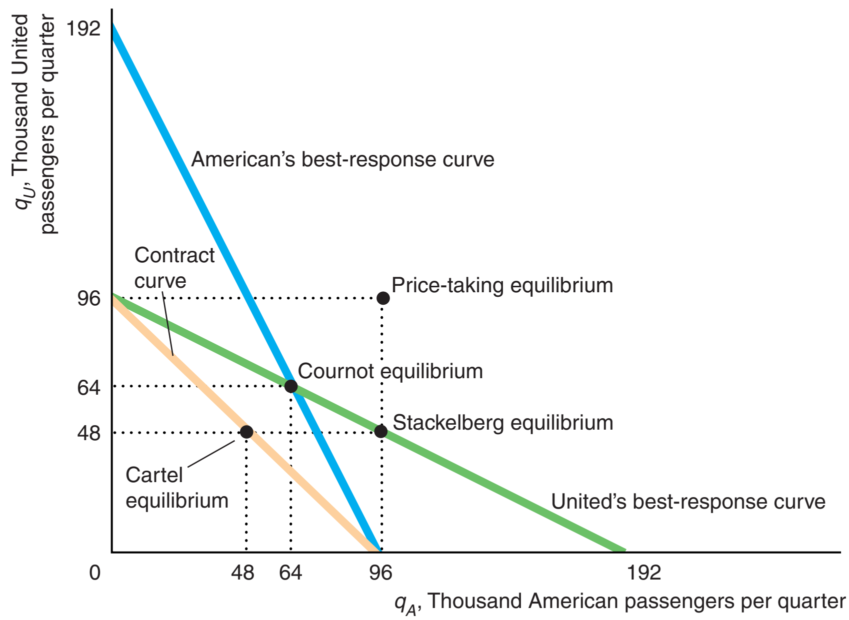
As a result, the competitive price, $147, is less than the Stackelberg price, $195, which is less than the Nash-Cournot price, $211, which is less than the collusive price, $243. Consumers, of course, prefer low prices. Consumer surplus is $18.4 million per quarter under competition, $10.4 million in the Stackelberg equilibrium, $8.2 million in the Nash-Cournot equilibrium, and only $4.6 million in the collusive outcome.

The cartel profits are the highest-possible level of profits that the firms can earn. The monopoly profit is $9.2 million per quarter, so each firm earns $4.6 million if they split the profit equally. In contrast, if the firms act independently, each earns the Cournot profit of approximately $4.1 million. The Stackelberg leader earns more than the Cournot profit, $4.6 million, while the follower earns less, $2.3 million.

We showed that the Nash-Cournot equilibrium approaches the competitive, price-taking equilibrium as the number of firms grows. Similarly, we can show that the Nash-Stackelberg equilibrium approaches the price-taking equilibrium as the number of Stackelberg followers grows. As a result, the differences between the Cournot, Stackelberg, and price-taking market structures shrink as the number of firms grows.

Figure 13.6 Duopoly Equilibria

The intersection of the best-response curves determines the Nash-Cournot equilibrium, where each firm produces 64. The possible cartel equilibria lie on the contract curve: $q_U = 96 - q_A$. The figure shows the symmetric cartel case where each firm produces 48. If the firms act as price takers, each firm produces where its residual demand equals its marginal cost, 96. In the Nash-Stackelberg equilibrium, the leader produces more, 96, than the follower, 48.



13.5 Bertrand Oligopoly

We have examined how oligopolies set quantities to try to maximize their profits. However, many oligopolistic firms set prices instead of quantities and allow consumers to decide how much to buy. The market equilibrium is different if firms set prices rather than quantities.

In monopolistic and competitive markets, the issue of whether firms set quantities or prices does not arise. Competitive firms have no choice: They cannot affect price and hence can choose only quantity (Chapter 8). The monopoly equilibrium is the same whether the monopoly sets price or quantity (Chapter 11).

Bertrand equilibrium (Nash-Bertrand equilibrium)
a *Nash equilibrium* in prices; a set of prices such that no firm can obtain a higher profit by choosing a different price if the other firms continue to charge these prices

In 1883, Joseph Bertrand argued that oligopolies set prices and then consumers decide how many units to buy. The resulting Nash equilibrium is called a **Bertrand equilibrium** or **Nash-Bertrand equilibrium** (or Nash-in-prices equilibrium): a set of prices such that no firm can obtain a higher profit by choosing a different price if the other firms continue to charge these prices.

We will show that the price and quantity in a Nash-Bertrand equilibrium are different from those in a Nash-Cournot equilibrium. We will also show that a Nash-Bertrand equilibrium depends on whether firms are producing identical or differentiated products.

Identical Products

We start by examining a price-setting oligopoly in which firms have identical costs and produce identical goods. The resulting Nash-Bertrand equilibrium price equals the marginal cost, as in the price-taking equilibrium. To show this result, we use best-response curves to determine the Nash-Bertrand equilibrium, as we did in the Cournot model.

Best-Response Curves Suppose that each of the two price-setting oligopolistic firms in a market produces an identical product and faces a constant marginal and average cost of \$5 per unit. What is Firm 1's best response—what price should it set—if Firm 2 sets a price of $p_2 = \$10$? If Firm 1 charges more than \$10, it makes no sales because consumers will buy from Firm 2. Firm 1 makes a profit of \$5 on each unit it sells if it also charges \$10 per unit. If the market demand is 200 units and both firms charge the same price, we'd expect Firm 1 to make half the sales, so its profit is \$500.

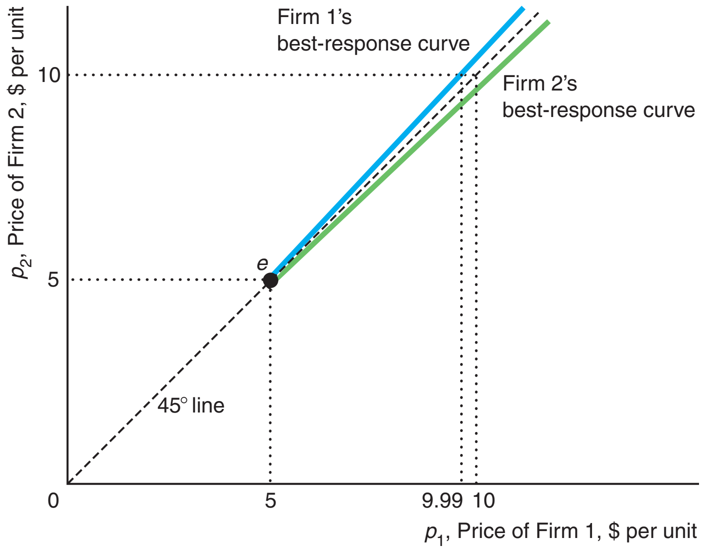
Suppose, however, that Firm 1 slightly undercuts its rival's price by charging \$9.99. Because the products are identical, Firm 1 captures the entire market. Firm 1 makes a profit of \$4.99 per unit and a total profit of \$998. Thus, Firm 1's profit is higher if it slightly undercuts its rival's price. By similar reasoning, if Firm 2 charges \$8, Firm 1 also charges slightly less than Firm 2.

Now imagine that Firm 2 charges $p_2 = \$5$. If Firm 1 charges more than \$5, it makes no sales. The firms split the market and make zero profit if Firm 1 charges \$5. If Firm 1 undercuts its rival, it captures the entire market, but it makes a loss on each unit. Therefore, Firm 1 will undercut only if its rival's price is higher than Firm 1's marginal and average cost of \$5. By similar reasoning, if Firm 2 charges less than \$5, Firm 1 chooses not to produce.

Figure 13.7 shows that Firm 1's best response is to produce nothing if Firm 2 charges less than \$5. Firm 1's best response is \$5 if Firm 2 charges \$5. If Firm 2 charges prices above \$5, Firm 1's best response is to undercut Firm 2's price slightly. Above \$5, Firm 1's best-response curve is above the 45° line by the smallest amount possible. (The figure exaggerates the distance of the best-response curve from the 45° line for clarity.) By the same reasoning, Firm 2's best-response curve starts at \$5 and lies slightly below the 45° line.

Figure 13.7 Nash-Bertrand Equilibrium with Identical Products MyLab Economics Video

With identical products and constant marginal and average costs of $5, Firm 1's best-response curve starts at $5 and then lies slightly above the 45° line. That is, Firm 1 undercuts its rival's price as long as its price remains above $5. The best-response curves intersect at *e*, the Nash-Bertrand equilibrium, where both firms charge $5.



The two best-response functions intersect only at *e*, where each firm charges $5. It does not pay either firm to change its price as long as the other firm charges $5, so *e* is the Nash-Bertrand equilibrium. In this equilibrium, each firm makes zero profit. Thus, *the Nash-Bertrand equilibrium when firms produce identical products is the same as the price-taking, competitive equilibrium.*

Bertrand Versus Cournot This Nash-Bertrand equilibrium differs substantially from the Nash-Cournot equilibrium. When firms produce identical products and have a constant marginal cost, firms receive positive profits and the price is above marginal cost in the Nash-Cournot equilibrium, whereas firms earn zero profits and price equals marginal cost in the Nash-Bertrand equilibrium.

When firms' products are identical, the Cournot model seems more realistic than the Bertrand model. The Bertrand model appears inconsistent with actual oligopolistic markets in at least two ways. First, the Bertrand model's "competitive" equilibrium price is implausible. In a market with few firms, why would the firms compete so vigorously that they would make no profit, as in the Nash-Bertrand equilibrium? In contrast, the Nash-Cournot equilibrium price with a small number of firms lies between the competitive price and the monopoly price. Because oligopolies typically charge a higher price than competitive firms, the Nash-Cournot equilibrium is more plausible.

Second, the Nash-Bertrand equilibrium price, which depends only on cost, is insensitive to demand conditions and the number of firms. In contrast, the Nash-Cournot equilibrium price depends on demand conditions and the number of firms as well as on cost. Consequently, economists are much more likely to use the Cournot model than the Bertrand model to study homogeneous goods markets.

Differentiated Products

If firms in most markets produced homogeneous goods, economists probably would have forgotten the Bertrand model. However, markets with differentiated goods—automobiles, stereos, computers, toothpastes, and spaghetti sauces—are extremely

common, as is price setting by firms in such markets. In differentiated-goods markets, the Nash-Bertrand equilibrium is plausible because the two "problems" of the homogeneous-goods model disappear: Firms set prices above marginal cost, and prices are sensitive to demand conditions and the number of firms.

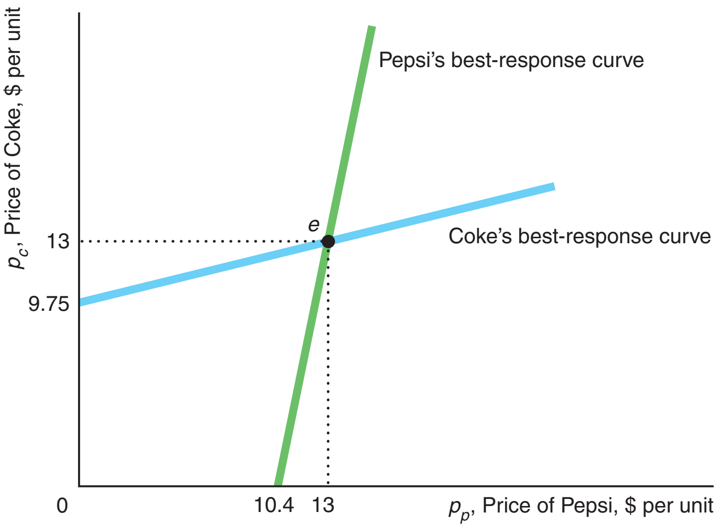
Indeed, many economists believe that price-setting models are more plausible than quantity-setting models in differentiated-goods markets. In such markets, firms set prices, then consumers decide how much to buy. In contrast, if firms set quantities, it is not clear how the prices of the differentiated goods are determined in the market.

Cola Market We illustrate a Nash-Bertrand equilibrium with the differentiated products in the cola market. We use best-response curves in a figure to solve for the equilibrium.

Figure 13.8 shows the firms' best-response curves. Quantities are in tens of millions of cases (a case consists of 24 twelve-ounce cans) per quarter, and prices (to retailers) and costs are in real 1982 dollars per 10 cases. The best-response curves in the figure were derived (see Appendix 13C) from demand functions estimated by Gasmi et al. (1992).[20] Coke and Pepsi produce similar but not identical products, so many consumers prefer one of these products to the other. If the price of Pepsi were to fall slightly relative to that of Coke, some consumers who prefer Coke to Pepsi would not switch. Thus, neither firm has to match a rival's price exactly. As a result, neither firm's best-response curve lies along a 45° line through the origin (as in Figure 13.7).

Figure 13.8 Nash-Bertrand Equilibrium with Differentiated Products MyLab Economics Video

Both Coke and Pepsi, which set prices, have upward-sloping best-response curves. These best-response curves of Coke and Pepsi intersect at *e*, the Nash-Bertrand equilibrium, where each sets a price of $13 per unit.



[20]Their estimated model allows the firms to set both prices and advertising. In the figure, we hold the firms' advertising constant. The Coke equations are the authors' estimates (with slight rounding). We rescaled the Pepsi equation so that the equilibrium prices of Coke and Pepsi are equal.

The Bertrand best-response curves have different slopes than the Cournot best-response curves in Figure 13.3. The Cournot curves—which plot relationships between quantities—slope downward, showing that a firm produces less the more its rival produces. In Figure 13.8, the Bertrand best-response curves—which plot relationships between prices—slope upward, indicating that a firm charges a higher price the higher the price its rival charges.

The Nash-Bertrand equilibrium, *e* in Figure 13.8, occurs where each firm's price is $13 per unit. In this Nash-Bertrand equilibrium, each firm sets its best-response price *given the price the other firm is charging*. Neither firm wants to change its price because neither firm can increase its profit by doing so.

Product Differentiation and Welfare Because differentiation makes demand curves less elastic, prices are likely to be higher when products are differentiated than when they're identical. We also know that welfare falls as the gap between price and marginal cost rises. Does it follow that differentiating products lowers welfare? Not necessarily. Although differentiation leads to higher prices, which harm consumers, differentiation is desirable in its own right. Consumers value having a choice, and some may greatly prefer a new brand to existing ones.

One way to illustrate the importance of this second effect is to consider what the value is of introducing a new, differentiated product. This value reflects how much extra income consumers would require to be as well off without the good as with it.

Application

Bottled Water

Bottled water is a dramatic example of *spurious product differentiation*, where the products do not significantly differ physically. You might think it is difficult to differentiate water. How do firms differentiate their uncarbonated, unflavored water? Primarily through marketing. According to a 2016 report, 53% of U.S. consumers preferred to drink "premium" bottled water.

Pepsico's top-selling bottled water, Aquafina, has a colorful blue label and a logo showing the sun rising over the mountains. From that logo, consumers may infer that the water comes from some bubbling spring high in an unspoiled wilderness. If so, they're wrong. Pepsi's best-selling bottled water comes from the same place as tap water: public-water sources. Pepsi also claims that it adds value by filtering the water using a state-of-the-art "HydRO-7 purification system," implying that such filtering (which removes natural minerals) is desirable. Similarly, Coke's marketing distinguishes its Dasani bottled water, even though it too is basically bottled public water.

In a "blind" taste test reported in *Slate*, no one could distinguish between Aquafina and Dasani, and both are equally clean and safe. However, many consumers, responding to perceived differences created by marketing, strongly prefer one or the other of these brands and pay a premium for these products.[21]

[21]Having succeeded in differentiating water, Coca-Cola turned to milk. In 2015, it started selling its Fairlife "super milk." Using a special filtration process, it has more "natural" protein and calcium, and less sugar. Sandy Douglas, President of Coca-Cola North America, said, "It's basically the premiumization of milk . . . We'll charge twice as much for it as the milk we're used to buying in a jug."

13.6 Monopolistic Competition

We now turn to monopolistic competition, which is a market structure that has the price setting characteristics of monopoly or oligopoly and the free entry characteristic of perfect competition. Monopolistically competitive firms have market power because they face downward-sloping demand curves, as do oligopolistic firms. However, in contrast to oligopolistic markets in which entry is very difficult, firms can freely enter a monopolistically competitive market, so firms earn zero economic profit, as do perfectly competitive firms.

If both competitive and monopolistically competitive firms make zero profits, what distinguishes these two market structures? In contrast to competitive firms (which face horizontal residual demand curves and charge prices equal to marginal cost), monopolistically competitive firms face downward-sloping residual demand curves, so they charge prices above marginal cost.

Monopolistically competitive firms face downward-sloping demand curves because the market is small or because the firms differentiate their products. Even if the firms produce identical products, if the market demand curve is close to the origin, the market may be able to support only a few firms, so the residual demand curve facing a single firm is downward sloping. For example, in a small town the market may be large enough to support only a few plumbing firms, each of which provides a similar service. If firms differentiate their products, each firm can retain those customers who particularly like that firm's product even if its price is higher than those of its rivals. Examples of monopolistically competitive firms are books, hotels, movies, plumbers in a small town, and restaurants.

Application

Monopolistically Competitive Food Truck Market

One of the hottest food phenomena in the United States is gourmet food trucks, which started in major West Coast cities such as Los Angeles, Portland, and Seattle. Gourmet food trucks serve differentiated food in monopolistically competitive markets. Now, flocks of food trucks ply their business in previously underserved areas of town in cities across the country.

Nouveau food trucks like Chairman Bao, Curry Up Now, and Liba Falafel sell high-quality lunches in San Francisco's blighted mid-Market area, which has few traditional, high-quality lunch restaurants. Because some customers prefer Chinese food to Indian food, Chairman Bao could raise its price without losing all its customers to Curry Up Now. Consequently, each of these trucks faces a downward-sloping demand curve.

The mobile restaurant business has been exploding. As William Bender, a food service consultant in Santa Clara, California, said, "The limited menu approach, high quality, and low operating costs have opened up an entirely new sector." Even top restaurant chefs have entered this business. Celebrity Los Angeles chef Ludovic Lefebvre created LudoTruck, a mobile fried chicken outlet. San Francisco's Chez Spencer has a "French takeaway," Spencer on the Go, which serves bistro food such as foie gras torchon and toast for $12.

The cost of entry is very low, ranging from $50,000 to lease the equipment and pay ancillary expenses, to $250,000 or more for a deluxe truck and top-of-the-line cooking and refrigeration facilities. Potential entrants can learn about the business at **mobilefoodnews.com**, which reports on local laws, where to buy equipment and obtain insurance, and a host of other topics.

Opening a new brick-and-mortar restaurant is very risky. If demand is less than anticipated, the firm loses its (large) fixed cost. However, if the manager of a food truck's first guess as to where to locate is wrong, it is easy to drive to another neighborhood.

How do firms identify profit opportunities? "Lunch is our consistent bread-and-butter market," said Matthew Cohen, proprietor of Off the Grid, a food truck promoter and location finder in the San Francisco Bay Area. When lines in front of his trucks grow longer at lunch time, he sets up additional trucks. Having started with a dozen trucks in 2010, Cohen had over 200 vendors by 2016.

Equilibrium

To examine the monopolistically competitive equilibrium, we initially assume that firms have identical cost functions and produce identical products. Two conditions hold in a long-run monopolistically competitive equilibrium: *marginal revenue equals marginal cost* because firms set output to maximize profit, and *price equals average cost*—that is, profit is zero—because firms enter until no further profitable entry is possible.

Figure 13.9 shows a monopolistically competitive market equilibrium. A typical monopolistically competitive firm faces a residual demand curve D^r. To maximize its profit, the firm sets its output, q, where its marginal revenue curve corresponding to the residual demand curve intersects its marginal cost curve: $MR^r = MC$. At that quantity, the firm's average cost curve, AC, is tangent to its residual demand curve. Because the height of the residual demand curve is the price, at the point of tangency, price equals average cost, $p = AC$, and the firm makes zero profit.

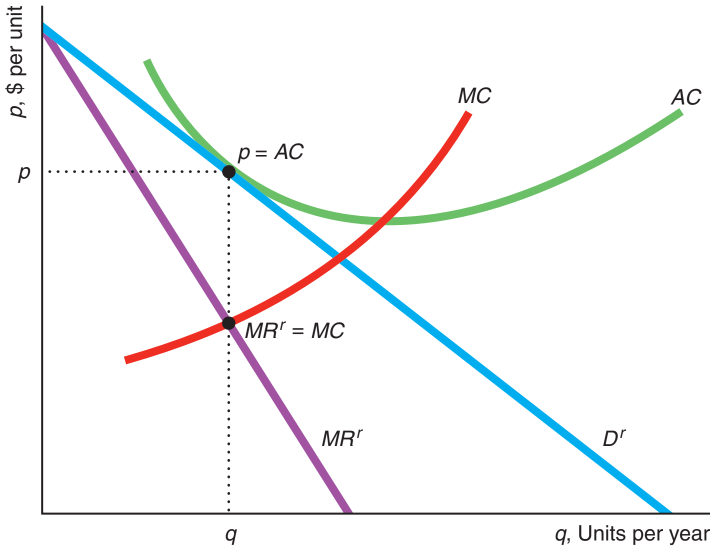
minimum efficient scale (full capacity)
the smallest quantity at which the average cost curve reaches its minimum

The smallest quantity at which the average cost curve reaches its minimum is referred to as *full capacity* or **minimum efficient scale**. The firm's full capacity or minimum efficient scale is the quantity at which the firm no longer benefits from economies of scale. Because a monopolistically competitive equilibrium occurs in the downward-sloping section of the average cost curve (where the average cost curve is tangent to the downward-sloping demand curve), a monopolistically competitive firm operates at less than full capacity in the long run.

The fewer monopolistically competitive firms, the less elastic the residual demand curve each firm faces. As we saw, the elasticity of demand for an individual Cournot firm is n , where n is the number of firms and is the market elasticity. Thus, the fewer the firms in a market, the less elastic the residual demand curve and the higher

Figure 13.9 Monopolistically Competitive Equilibrium MyLab Economics Video

A monopolistically competitive firm, facing residual demand curve D^r, sets its output where its marginal revenue equals its marginal cost: $MR^r = MC$. In response to a profit opportunity, firms enter the market, driving the profit of the marginal firm to zero, so price equals the firm's average cost: $p = AC$.



the price. Similarly, the more differentiated are firms' products, the less elastic the residual demand curve and the higher the price.

Fast-food restaurants are an example of such a monopolistically competitive industry. These restaurants differentiate their food, so each may face a downward-sloping demand curve. However, restaurants can easily enter and exit the market, so the marginal firm earns zero economic profit. Most restaurants have empty seats much of the time and hence are operating at less than full capacity. The following Solved Problem provides an explanation for this phenomenon.

Solved Problem 13.4

Show that a monopolistically competitive firm maximizes its profit where it is operating at less than *full capacity* or *minimum efficient scale*, which is the smallest quantity at which the average cost curve reaches its minimum (the bottom of a U-shaped average cost curve). The firm's minimum efficient scale is the quantity at which the firm no longer benefits from economies of scale.

Answer

Use the properties of the demand curve to show that a monopolistically competitive firm operates in the increasing-returns to scale section of its average cost curve (the downward-sloping section) in the long-run equilibrium. In the long-run equilibrium, a monopolistically competitive firm operates where its downward-sloping demand curve is tangent to its average cost curve, as Figure 13.9 illustrates. Because its demand curve is downward sloping, its average cost curve must also be downward sloping in the equilibrium. Thus, the firm chooses to operate at less than full capacity in equilibrium.

Fixed Costs and the Number of Firms

The number of firms in a monopolistically competitive equilibrium depends on firms' costs. The larger each firm's fixed cost, the smaller the number of monopolistically competitive firms in the market equilibrium.

Although entry is free, if the fixed costs are high, few firms may enter. In the automobile industry, just to develop a new fender costs \$8 to \$10 million. Developing a new pharmaceutical drug may cost \$350 million or more.

We can illustrate this relationship using the airlines example, in which we modify our assumptions about entry and fixed costs. American and United are the only airlines providing service on the Chicago–Los Angeles route. Until now, we have assumed that a barrier to entry—such as an inability to obtain landing rights at both airports—prevented entry and that the firms had no fixed costs. If fixed cost is zero and marginal cost is constant at \$147 per passenger, average cost is also constant at \$147 per passenger. As we showed earlier, each firm in this oligopolistic market flies $q = 64$ per quarter at a price of $p = \$211$ and makes a profit of \$4.1 million per quarter.

Now suppose that the market has no barriers to entry, but each airline incurs a fixed cost to enter of F due to airport fees, capital expenditure, or other factors. Each firm's marginal cost remains \$147 per passenger, but its average cost,

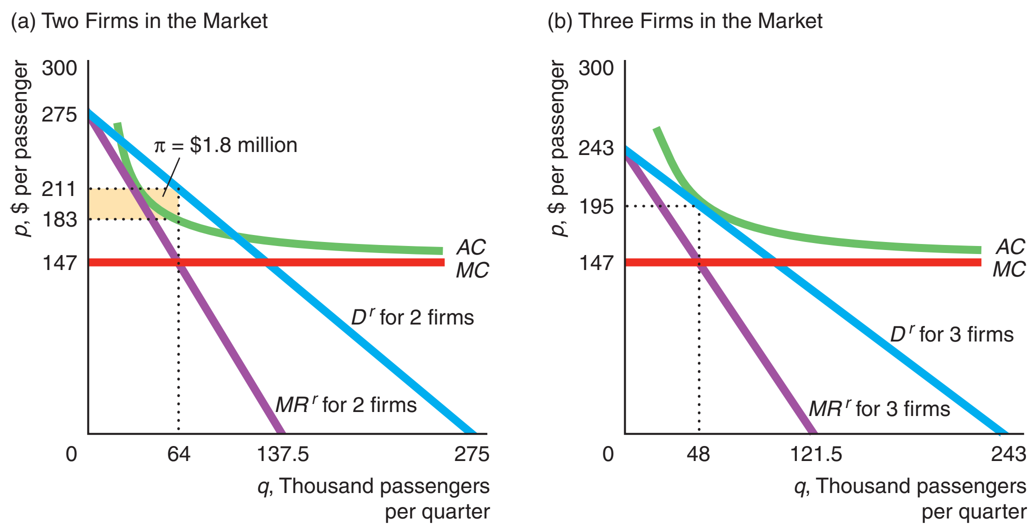
$$AC = 147 + \frac{F}{q},$$

falls as the number of passengers rises, as panels a and b of Figure 13.10 illustrate for $F = \$2.3$ million.

If the monopolistically competitive market has only two firms, what must the fixed costs be so that the two firms earn zero profit? We know that these firms receive a

Figure 13.10 Monopolistic Competition Among Airlines

(a) In a market with two identical airlines that each incurs a fixed cost of \$2.3 million, each firm flies $q = 64$ units (thousands of passengers) per quarter at a price of $p = \$211$ per passenger and makes a profit of \$1.8 million. This profit attracts entry. (b) After a third firm enters, the residual demand curve shifts, so each firm flies $q = 48$ units at $p = \$195$ and makes zero profit, which is the monopolistically competitive equilibrium.



profit of \$4.1 million per firm in the absence of fixed costs. As a result, the fixed cost must be \$4.1 million per firm for the firms to earn zero profit. With this fixed cost, the monopolistically competitive price and quantity are the same as in the oligopolistic equilibrium, $q = 64$ and $p = \$211$, and the number of firms is the same, but now each firm's profit is zero.

If the fixed cost is only \$2.3 million and the market has only two firms, each firm makes a profit, as panel a shows. Each duopoly firm faces a residual demand curve (labeled "D^r for 2 firms"), which is the market demand minus its rival's Nash-Cournot equilibrium quantity, $q = 64$. Given this residual demand, each firm produces $q = 64$, which equates its marginal revenue, MR^r, and its marginal cost, MC. At $q = 64$, the firm's average cost is $AC = \$147 + (\$2.3 \text{ million})/(64 \text{ units}) \approx \$183$, so each firm makes a profit of $\pi = (p - AC)q \approx (\$211 - \$183) \times 64$ units per quarter $\approx$ \$1.8 million per quarter.

This substantial economic profit attracts an entrant. The entry of a third firm causes the residual demand for any one firm to shift to the left in panel b. In the new equilibrium, each firm sets $q = 48$ and charges $p = \$195$. At this quantity, each firm's average cost is \$195, so the firms break even. No other firms enter because if one did, the residual demand curve would shift even farther to the left and all the firms would lose money. Thus, if fixed cost is \$2.3 million, the market has three firms in the monopolistically competitive equilibrium. This example illustrates a general result: *The lower the fixed costs, the more firms in the monopolistically competitive equilibrium.*

Solved Problem 13.5

What is the monopolistically competitive airline equilibrium if each firm has a fixed cost of \$3 million?

Answer

1. *Determine the number of firms.* We already know that the monopolistically competitive equilibrium has two firms if the fixed cost is \$4.1 million and three firms if the fixed cost is \$2.3 million. With a fixed cost of \$3 million and two firms in the market, each firm makes a profit of \$1.1 (= \$4.1 − 3) million. If another firm enters, though, each firm's loss equals −\$0.7 (= 2.3 − 3) million. Thus, the monopolistically competitive equilibrium has two firms, each of which earns a positive profit that is too small to attract another firm. This outcome is a monopolistically competitive equilibrium because no other firm wants to enter.

2. *Determine the equilibrium quantities and prices.* Because each duopoly firm produces $q = 64$, $Q = 128$, and $p = \$211$.

Application

Zoning Laws as a Barrier to Entry by Hotel Chains

U.S. local governments restrict land use through zoning. The difficulty of getting permission (generally from many agencies) to build a new commercial structure is a barrier to entry, which limits the number of firms in a monopolistically competitive market. Suzuki (2013) examined the effect of Texas municipalities' zoning laws on chain hotels, such as Best Western, Comfort Inn, Holiday Inn, La Quinta Inn, Quality Inn, and Ramada.

According to his estimates, construction costs are large even in the absence of zoning regulations. Construction costs are \$2.4 million for a new Best Western hotel and \$4.5 million for a new La Quinta hotel. Going from a lenient to a stringent zoning policy increases a hotel's variable cost by 21% and its sunk entry cost by

19%. The average number of hotels in a small market falls from 2.3 under a lenient policy to 1.9 with a stringent policy due to the higher entry cost. Consequently, a stringent policy reduces the number of rooms by 15%, which increases the revenue per room by 7%. The change from the most lenient policy to the most stringent policy decreases producer surplus by $1.5 million and consumer surplus by $1 million. Thus, tougher zoning laws raise entry costs and reduce the number of hotels and rooms, which causes the price to rise and lowers total surplus.

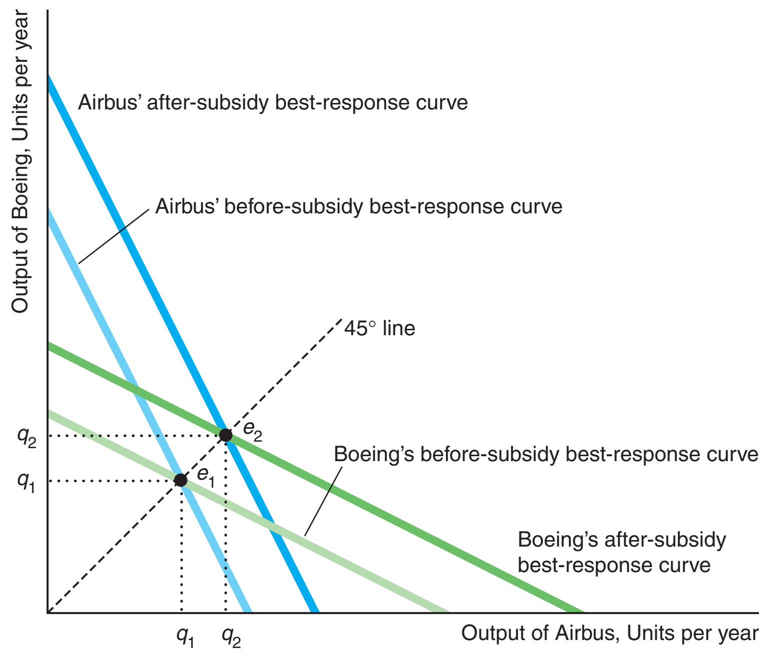
Therefore, owners of existing hotels often strongly support making zoning laws more stringent. These oligopolistic firms want to continue to earn positive economic profits, rather than see these profits driven to zero in a monopolistically competitive market with free entry.

Challenge Solution

Government Aircraft Subsidies

If only one government subsidizes its aircraft-manufacturing firm, how does the firm use the subsidy to gain a competitive advantage? What happens if both firms are government subsidized? Should Boeing and Airbus lobby for government subsidies that result in a subsidy war?

To keep our answers to these questions as simple as possible, we use a Cournot model in which Airbus and Boeing produce identical products with identical costs and face a linear demand curve.[22] A government per-unit subsidy to only one firm would cause its marginal cost to be lower than its rival's.



[22]We would reach the same conclusions if we used a Cournot or Bertrand model with differentiated products.

To maximize profit, a firm in a Cournot market should respond by increasing its output for any expected output level by its rival. That is, its best-response curve shifts out. In panel a of Figure 13.4, we saw how the equilibrium changes if United Airline's marginal cost falls while American's stays constant. As its marginal cost drops, United wants to produce more for any given output of its rival, so that its best-response function shifts out, away from the origin in panel b.

The market equilibrium shifts from e_1 to e_2 in panel b, so that United's Nash-Cournot equilibrium output increases and American's falls. Because total output rises, the equilibrium price falls. United benefits at the expense of American. Indeed, United's profit rises by \$5.1 million, which exceeds the actual cost saving of \$4.6 million. That is, United's managers used the cost savings to gain a competitive advantage.

The same analysis applies to the aircraft market. If Airbus is subsidized and Boeing is not, Airbus should produce more given any expected output from Boeing. Its equilibrium quantity and profit rise while Boeing's quantity and profit fall.

How much should the government subsidize Airbus? The answer depends on the government's objective. A plausible objective is that the government wants to maximize the sum of its firm's profit including the subsidy minus the cost of the subsidy, which is the firm's profit.[23] With this objective, the government wants to give a subsidy that leads to the Stackelberg outcome, which maximizes the firm's profit.

If both governments give identical subsidies that lower each firm's marginal cost, then both firms' best-response curves shift out as the figure shows. The intersection of the original best-response curves determine the original, unsubsidized equilibrium, e_1. The new, subsidized equilibrium, e_2, occurs where the new best-response curves intersect. Both firms produce more in the new equilibrium than in the original: $q_2 > q_1$.

Unlike the situation in which only one government subsidizes its firm, each subsidized firm increases its equilibrium output by the same amount so that the price falls.[24] Each government is essentially subsidizing final consumers in other countries without giving its own firm a strategic advantage over its rival.

Each government's welfare is the sum of its firm's profit including the subsidy minus the cost of the subsidy, which is the firm's profit ignoring the subsidy. Because the firms produce more than in the Nash-Cournot equilibrium, both firms earn less (ignoring the subsidies) so both countries suffer. It would be in both countries' best interests not to engage in a subsidy war. Nonetheless, both firms benefit from its subsidy, so both firms strongly lobby their governments for subsidies.

[23]The subsidy is a transfer from some citizens (taxpayers) to others (the owners of Airbus), so the government may not care about the size of the subsidy. The government does not care about consumers who live in another country.

[24]In 1992, the relevant governments signed a U.S.–EU agreement on trade in civil aircraft that limits government subsidies, including a maximum direct subsidy limit of 33% of development costs and various limits on variable costs. Irwin and Pavcnik (2004) found that aircraft prices increased by about 3.7% after the 1992 agreement. This price hike is consistent with ending a subsidy that was 5% of the firms' marginal costs.

Summary

1. **Market Structures.** Prices, profits, and quantities in a market equilibrium depend on the market's structure. Because profit-maximizing firms set marginal revenue equal to marginal cost, price is above marginal revenue—and hence marginal cost—only if firms face downward-sloping demand curves. In monopolistic, oligopolistic, and monopolistically competitive markets, firms face downward-sloping demand curves, in contrast to firms in a perfectly competitive market. When entry is blocked, as with a monopoly or an oligopoly, firms may earn positive profits; however, when entry is free, as in perfect competition or monopolistic competition, profits are driven toward zero. Noncooperative oligopolistic and monopolistically competitive firms, in contrast to competitive and monopolistic firms, must pay attention to their rivals.

2. **Cartels.** If firms successfully collude, they produce the monopoly output and collectively earn the monopoly level of profit. Although their collective profits rise if all firms collude, each individual firm has an incentive to cheat on a cartel arrangement to raise its own profit even higher. For cartel prices to remain high, cartel members must be able to detect and prevent cheating, and noncartel firms must not be able to supply very much output. When antitrust laws or competition policies prevent firms from colluding, firms may try to merge if permitted by law.

3. **Cournot Oligopoly.** If oligopolistic firms act independently, market output and the firms' profits lie between the competitive and monopoly levels. In a Cournot model, each oligopolistic firm sets its output at the same time. In the Cournot (Nash) equilibrium, each firm produces its best-response output—the output that maximizes its profit—given the output its rival produces. As the number of Cournot firms increases, the Nash-Cournot equilibrium price, quantity, and profits approach the price-taking levels.

4. **Stackelberg Oligopoly.** If one firm, the Stackelberg leader, chooses its output before its rivals, the Stackelberg followers, the leader produces more and earns a higher profit than each identical-cost follower firm. A government may subsidize a domestic oligopolistic firm so that it produces the Stackelberg leader quantity, which it sells in an international market. For a given number of firms, the Nash-Stackelberg equilibrium output exceeds that of the Nash-Cournot equilibrium, which exceeds that of the collusive equilibrium (which is the same as a monopoly produces). Correspondingly, the Stackelberg price is less than the Cournot price, which is less than the collusive or monopoly price.

5. **Bertrand Oligopoly.** In many oligopolistic or monopolistically competitive markets, firms set prices instead of quantities. If the product is homogeneous and firms set prices, the Nash-Bertrand equilibrium price equals marginal cost (which is lower than the Nash-Cournot quantity-setting equilibrium price). In markets with differentiated products, the Nash-Bertrand equilibrium price is above marginal cost. Typically, greater product differentiation increases the markup of price over marginal cost.

6. **Monopolistic Competition.** In monopolistically competitive markets, after all profitable entry occurs, the market has few enough firms that each firm faces a downward-sloping demand curve. Consequently, the firms charge prices above marginal cost. These markets are not perfectly competitive due to the relatively small number of firms. The number of firms may be small because of high fixed costs, economies of scale that are large relative to market demand, or because the firms sell differentiated products.

Questions

*Select questions are available on MyLab Economics; * = answer appears at the back of this book;* **A** *= algebra problem;* **C** *= calculus problem.*

1. Market Structures

1.1 Which market structure best describes (a) mobile network operators in a country, (b) plumbers in a small town, (c) farms that grow mangoes, and (d) city waste management? Why?

2. Cartels

2.1 At each Organization of Petroleum Exporting Countries (OPEC) meeting, Saudi Arabia, the largest oil producer, argues that the cartel should cut production. The Saudis complain that most OPEC member countries (including Saudi Arabia) produce more oil than the cartel agreement allots. Use a graph and words to explain why cartel members would produce more than the allotted amount given that they know that overproduction will drive down the price of their product.

2.2 As the aggregate profit of a cartel can exceed the combined profits of the same firms acting independently, firms have an incentive to form cartels and behave like a monopoly. However, cartels often fail for economic reasons. Why is this the case?

2.3 In April 2011, the European Commission found Procter & Gamble, Unilever, and Henkel, the leading producers of washing powder in Europe, guilty of operating a cartel from 2002 to 2005 in eight European Union countries to stabilize market positions and coordinate prices. Henkel received immunity for revealing the cartel to the Commission, and the other two companies were fined €315.2 million. In spring 2016, the Australian Competition and Consumer Commission (ACCC) found Colgate-Palmolive and Woolworths guilty of conspiring to control the price of laundry detergents from 2009 to 2013 for the purpose of substantially lessening competition in the market for the wholesale supply of powdered and liquid laundry detergent products for domestic use. Two other companies, Cussons and Unilever, were also involved. This time Unilever received immunity for revealing the cartel. Colgate-Palmolive and Woolworths were fined over \$35 million. The federal court case against Cussons began in June 2016. What are the main factors that increase the likelihood of a cartel being successful and which of them was weak in these two cases?

*2.4 A 2016 study on 14 Dutch cartels that had an average duration of about five years and were fined by the Netherlands Authority for Consumers and Markets between 2007 and 2012 found that these cartels implemented a number of informal internal rules. One was that compensation for violating agreed-upon market shares by over or under 5% was not needed. Another was that, where members divided regions and a member from outside a region received an order in violation of the allocation agreement, the receiving member was required to pay 2% of the contract amount to the cartelist who "owned" the region (Jaspers, 2017). Explain why cartel members might agree to implement informal coordinating rules like these.

2.5 The Federation of Quebec Maple Syrup Producers supplies over three-quarters of the world's maple syrup (see the Application "Cheating on the Maple Syrup Cartel"). Under government rules, the member firms jointly market their syrup through the federation, which sets quotas on how much each firm can produce. Show this cartel's price determination process using a graph similar to Figure 13.1. Show how much profit a firm would gain by cheating: by producing more than the cartel's quota.

2.6 On March 29, 2017, the European Commission blocked a £21 billion merger between the London Stock Exchange and the Deutsche Börse on the grounds that the deal would create a de facto monopoly in the area of fixed-income instruments (*The Guardian*, 2017). The London and Frankfurt exchanges had also been unsuccessful in merging in 2002 and 2005. When mergers that would presumably be profitable for both companies are blocked, it is often the case that share prices for both companies fall. Why?

*2.7 A market has an inverse demand curve $p = 100 - 2Q$ and four firms, each of which has a constant marginal cost of $MC = 20$. If the firms form a profit-maximizing cartel and agree to operate subject to the constraint that each firm will produce the same output level, how much does each firm produce? (*Hint*: See Chapter 11's treatment of monopoly.) **A**

2.8 A hub-and-spoke cartel is a collusive relationship between a group of competitors (the spokes) and one (or more) of their common suppliers or customers (the hub). In August 2011, the former Office of Fair Trading in the United Kingdom found that the country's big supermarkets and certain dairy processors had exchanged information to increase the prices consumers paid for milk, butter, and cheese in 2002 and 2003 (The National Archives, 2011). In late 2007 and early 2008, most of the parties involved had admitted their involvement in anti-competitive practices and had agreed to pay penalties amounting to over £120 million. In August 2011, additional penalties were imposed totaling £49.51 million. Why would retailers want to help suppliers raise the price of dairy products?

2.9 In May 2017, *The Weekend Australian* reported that New Zealand's two largest newspaper networks would appeal a decision by the country's Commerce Commission to block them from merging. The newspaper networks disagree that their merger would likely lessen competition for advertising and readers, saying that the Commission did not correctly account for other local or international providers of online news in New Zealand. The Commission did, however, accept that the merger would result in significant cost savings for the combined company over the next five years. Do you agree that a government may be less reluctant to prevent firms from merging the more firms there are in a market?

3. Cournot Oligopoly

3.1 Suppose there are n firms selling a homogeneous product at a constant marginal cost, $m = 5$, in a Cournot market. The inverse market demand curve is $p = 20 - Q$, where $Q = nq$. What is the output of each firm and the market quantity and price when the number of firms is 1, 2, 14, and 60? What are the competitive market quantity and price? As the number of firms increases, what happens to the firms' market power as measured by the Lerner Index? (*Hint*: See Appendix 13A.)

*3.2 What is the duopoly Nash-Cournot equilibrium if the market demand function is $Q = 1{,}000 - 1{,}000p$, and each firm's marginal cost is \$0.28 per unit? **A**

3.3 Duopoly quantity-setting firms face the market demand $p = 150 - q_1 - q_2$. Each firm has a marginal cost of \$60 per unit. What is the Nash-Cournot equilibrium? **A**

*3.4 Suppose there are n firms selling a homogeneous product at a constant marginal cost in a Cournot market. The market demand curve is $Q = 100/p$, where $Q = nq$. What is the market elasticity of demand? What happens to the equilibrium price and market power when either marginal cost rises or the number of firms falls? (*Hint*: See Equation 13.7.) **A**

3.5 Show how the Nash-Cournot equilibrium for n firms given in Appendix 13A changes if each firm faces a fixed cost of F as well as a constant marginal cost per unit. (*Hint*: You may need very little, if any, formal math, though you may use it.) **A**

3.6 Lee et al. (2010) estimated that a 2009 tax of 10 New Taiwan Dollars (\$0.30) per pack of cigarettes reduced Taiwanese cigarette consumption by 13.19%. Assuming that the market consists of two cigarette firms, show how this specific tax affects the Nash-Cournot equilibrium. (*Hint*: Show how the tax affects the firms' marginal costs and hence their best-response functions.)

3.7 How would the airlines' Nash-Cournot equilibrium (Figure 13.3) change if United's marginal cost was \$100 and American's was \$200? (*Hint*: See Solved Problem 13.1.) **A**

3.8 In 2015, Spirit reported that its "average cost per available seat mile excluding special items and fuel" was 5.7¢ compared to 8.5¢ for Southwest. Assuming that Spirit and Southwest compete on a single route, use a graph to show that their equilibrium quantities differ. (*Hint*: See Solved Problem 13.1.)

*3.9 If the inverse market demand function facing a duopoly is $p = a - bQ$, what are the Nash-Cournot equilibrium quantities if the marginal cost of Firm 1 is m and that of Firm 2 is $m + x$, where $x > 0$? Which firm produces more and which has the higher profit? **A**

*3.10 Why do prices increase if oligopolistic firms differentiate their products?

3.11 How would the Intel-AMD equilibrium in Solved Problem 13.2 change if AMD faced the same demand function as Intel, Equation 13.10? **A**

3.12 Firms 1 and 2 produce differentiated goods. Firm 1's inverse demand function is $p_1 = 260 - 2q_1 - q_2$, while Firm 2's inverse demand function is $p_2 = 260 - 2q_2 - q_1$. Each firm has a constant marginal cost of 20. What is the Nash-Cournot equilibrium in this market? (*Hint*: See Solved Problem 13.2.) **A**

3.13 Draw a figure to illustrate the Intel-AMD Nash-Cournot equilibrium in Solved Problem 13.2. At what quantities do the best-response functions hit the axes?

3.14 In a Nash-Cournot equilibrium, does an oligopolistic firm produce at less than full capacity, full capacity, or more than full capacity? Explain.

4. Stackelberg Oligopoly

*4.1 Duopoly quantity-setting firms face the inverse market demand curve, $p = 100 - 5Q$, where $Q = q_1 + q_2$. Each firm has a constant marginal cost of 10 per unit.

a. What is the Nash-Cournot equilibrium?

b. What is the Nash-Stackelberg equilibrium when Firm 1 moves first?

(*Hint*: See Equations 13.5 and 13.6, Solved Problem 13.3.) **A**

4.2 Instead of two firms, as in Question 4.1, three firms now face the inverse market demand curve, $p = 100 - 5Q$, where $Q = q_1 + q_2 + q_3$, and each of the firms has a constant marginal cost of 10 per unit. What is the Nash-Stackelberg equilibrium when Firm 1 moves first and both of the other two firms are followers? Compare your results to your answer to Question 4.1. (*Hint*: See Solved Problem 13.3 and Appendix 13B.) **C**

4.3 If two quantity-setting firms act simultaneously, is the Stackelberg outcome likely? Why or why not?

*4.4 Suppose you live on the second floor of a two-story building that you own. You have managed to get the ground floor zoned for commercial use and are presently advertising the space to prospective tenants. Both a convenience store and the retail outlet of a wireless communications company have expressed an interest in meeting with you to negotiate the terms of a lease. If you wish to maximize your rental income, which of these two businesses would you prefer to have as a tenant?

*4.5 The market demand function is $Q = 1{,}000 - 1{,}000p$. Each firm has a marginal cost of $m = 0.28$ (28¢ per unit). Firm 1, the leader, acts before Firm 2, the follower. Solve for the Nash-Stackelberg equilibrium quantities, prices, and profits. (*Hint*: See Appendix 13B and Solved Problem 13.3.) Compare your solution to the Nash-Cournot equilibrium (Question 3.2). **C**

*4.6 A country's competition regulator decides to allow the merger of the only two firms in a market reasoning that, while there would be a lessening of competition, the potential efficiency gains would position the combined company so that it can deal with the threat of multinationals and compete internationally. The inverse market demand curve is $p = 30 - 2Q$, where $Q = q_1 + q_2$. Each firm has a constant marginal and average cost of 6 per unit. What are the Nash-Cournot equilibrium quantities, price, profits, consumer surplus, and deadweight loss? If the firms merge, how do the equilibrium values change? **A**

4.7 Suppose there are initially two identical firms that serve a market with constant marginal and average costs, $MC = AC = 6$, per unit of output. Firm 1 gains a first-mover advantage from an opportunity to purchase an input at a lower cost than is available to Firm 2. This reduces Firm 1's constant marginal and average costs to 4.8 per unit and causes it to act like a monopoly with respect to its residual demand. If the inverse market demand curve is $p = 60 - 3Q$, where $Q = q_1 + q_2$, how do the market equilibrium quantities, price, profits, and market power change?

5. Bertrand Oligopoly

5.1 What happens to the homogeneous-good Nash-Bertrand equilibrium price if the number of firms increases? Why?

*5.2 Will the price be lower if duopoly firms set price or if they set quantity? Under what conditions can you give a definitive answer to this question?

5.3 In the Coke and Pepsi example, what is the effect of a specific tax, t, on the equilibrium prices? (*Hint*: What does the tax do to the firm's marginal cost? You do not have to use math to answer this problem.)

5.4 Suppose that a market for toothpaste has two sellers, each of which says its product has the best combination of ingredients to protect against tooth decay and to control plaque, tartar, and gum disease. Each of the profit-maximizing oligopolists faces the same constant marginal cost, $MC = 1$. The demand function for Firm 1 is $q_1 = 10 - 2p_1 + p_2$ and for Firm 2 is $q_2 = 21 - 3p_2 + 2p_1$, where p_1 is Firm 1's price and p_2 is Firm 2's price. What are the Nash-Bertrand equilibrium prices and quantities? If an advertising campaign featuring product endorsements by the dental association increases the demand for the first firm's product to $q_1 = 14 - 2p_1 + p_2$ but does not affect the second firm's demand function, what are the new equilibrium prices and quantities, and how do they compare to the pre-merger prices?

*5.5 Suppose that identical duopoly firms have constant marginal costs of \$10 per unit. Firm 1 faces a demand function of $q_1 = 100 - 2p_1 + p_2$, where q_1 is Firm 1's output, p_1 is Firm 1's price, and p_2 is Firm 2's price. Similarly, the demand Firm 2 faces is $q_2 = 100 - 2p_2 + p_1$. Solve for the Nash-Bertrand equilibrium. **C**

5.6 Solve for the Nash-Bertrand equilibrium for the firms described in Question 5.5 if both firms have a marginal cost of \$0 per unit. **A**

5.7 Solve for the Nash-Bertrand equilibrium for the firms described in Question 5.5 if Firm 1's marginal cost is \$30 per unit and Firm 2's marginal cost is \$10 per unit. **A**

5.8 Consider the two toothpaste producers in Question 5.4. If instead of undertaking an advertising campaign, Firm 1 convinced Firm 2 to form a cartel, would it have been better off? Assume that both firms also face a constant average cost, $AC = 1$.

5.9 Various firms use marketing to differentiate their bottles of water (see the Application "Bottled Water"). If the firms in this market engage in a Bertrand game, what is the effect of this differentiation on prices? What is the effect on welfare?

6. Monopolistic Competition

6.1 Solved Problem 13.4 shows that a monopolistically competitive firm maximizes its profit where it is operating at less than full capacity. Does this result depend upon whether firms produce identical or differentiated products? Why?

6.2 The national government and some state governments in India provide incentives to entice new business ventures for less developed areas of the country to increase employment. Subsidies may be offered to reduce the cost of transportation, investment in plant and machinery, and loans relating to new industrial units. What effect would these incentives have on the number of firms and the firms' profits in monopolistically competitive markets in those areas of the country? Explain.

6.3 In the monopolistically competitive airlines model, what is the equilibrium if firms face no fixed costs?

6.4 Does an oligopolistic or a monopolistically competitive firm have a supply curve? Why or why not? (*Hint*: See the discussion in Chapter 11 of why a monopoly does not have a supply curve.)

6.5 In a monopolistically competitive market, the government applies a specific tax of \$1 per unit of output. What happens to the profit of a typical firm in this market? Does the number of firms in the market change? Why? (*Hint*: See Solved Problem 13.5.)

*6.6 An incumbent firm, Firm 1, faces a potential entrant, Firm 2, with a lower marginal cost. The market demand curve is $p = 120 - q_1 - q_2$. Firm 1 has a constant marginal cost of \$20, while Firm 2's is \$10.

a. What are the Nash-Cournot equilibrium price, quantities, and profits if the government does not intervene?

b. To block entry, the incumbent appeals to the government to require that the entrant incur extra costs. What happens to the Nash-Cournot equilibrium if the legal requirement causes the marginal cost of the second firm to rise to that of the first firm, \$20?

c. Now suppose that the barrier leaves the marginal cost alone but imposes a fixed cost. What is the minimal fixed cost that will prevent entry? (*Hint*: See Solved Problem 13.5.) **A**

7. Challenge

*7.1 Many countries provide funding for small companies to start or expand their businesses. In the European Union, subsidies take the form of grants, loans, and guarantees. In Question 4.1, duopoly quantity-setting firms face the inverse market demand curve, $p = 100 - 5Q$, where $Q = q_1 + q_2$, and each has a constant marginal cost, $MC = 10$. How does the equilibrium change if only Firm 1 receives a subsidy of 3 per unit, and if both firms receive subsidies of 3 per unit? **A**

7.2 Suppose there are two firms, each headquartered in a different country, that are competing in the world market for business aircraft. The inverse market demand curve is $p = 1 - Q$, where $Q = q_1 + q_2$, and each firm has the same constant marginal and average cost, $MC = AC = 0.4$. Equilibrium in the market is initially Nash-Cournot. The government of the "domestic" country then decides to subsidize its aviation firm to give it a strategic cost advantage over the "foreign" firm. The reduction in its marginal (and average) cost of production allows the domestic firm to increase its output and profit at the expense of the foreign firm, and this increases domestic national income at the expense of the other country. How large a subsidy would be needed to increase the domestic firm's profit to the level of a Stackelberg leader? Show that with this subsidy, the foreign firm produces the Stackelberg follower quantity in equilibrium. **C**

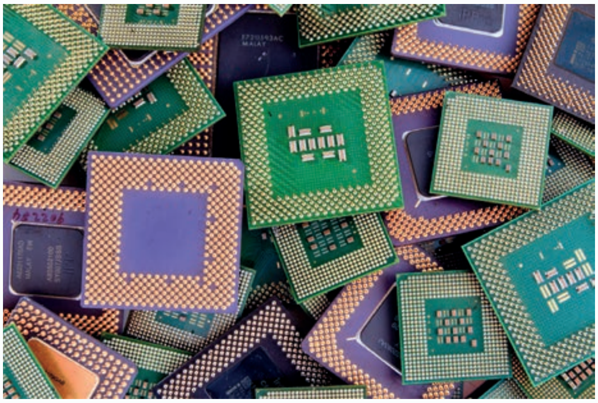
14 Game Theory

A camper awakens to the growl of a hungry bear and sees his friend putting on a pair of running shoes. "You can't outrun a bear," scoffs the camper. His friend coolly replies, "I don't have to. I only have to outrun you!"

Challenge

Intel and AMD's Advertising Strategies

Intel and Advanced Micro Devices (AMD) had essentially 100% of the market for central processing unit (CPU) for personal computers and 83% for graphic chip units (GPU) in the first quarter of 2016. Intel uses aggressive advertising—its very successful *Intel Inside* campaign—and charges relatively high prices, while traditionally AMD used little advertising and relied on the appeal of its lower prices. Intel controls more than 82% of the processor market and 70% of the graphic chip market.



According to Salgado's (2008) estimated demand functions, consumers were willing to pay a large premium for the Intel brand for processors. He found that, if Intel increased its advertising by 10% (holding prices constant), the total market demand would increase by 1%, while Intel's relative share would rise by more than 3%. Demand for AMD products would therefore fall. Salgado's work indicates that the two firms' shares would be roughly equal if they advertised equally (regardless of the level).

From the start of the personal computer era, Intel has been the 800-pound gorilla in the CPU market. Intel created the first commercial microprocessor chip in 1971. In 1991, Intel launched the Intel Inside® marketing and branding campaign. Intel offered to share costs for any manufacturer's PC print ads if they included the Intel logo. Not only did these funds reduce the computer manufacturers' costs, but also the logo assured consumers that their computers ran on the latest technology. Within six months, 300 computer manufacturers had agreed to support the campaign. After the manufacturers' ads started to appear, Intel advertised globally to explain the significance of the logo to consumers. The Intel Inside campaign was one of the first successful attempts at *ingredient branding*.

Advanced Micro Devices (AMD) entered the microchip market in 1975 when it started selling a reverse-engineered clone of the Intel 8080 microprocessor. In 1982, AMD and Intel signed a contract allowing AMD to be a licensed second-source manufacturer of Intel's 8086 and 8088 processors because IBM would use these chips in its PCs only if it had two microchip sources.

Why does Intel advertise aggressively while AMD engages in relatively little advertising? At the end of the chapter, we discuss a possible explanation: Intel was able to act first and thereby gain an advantage. (In contrast, in Solved Problem 14.1, we examine the possible outcomes if both firms acted simultaneously.)

When a small number of people or firms—such as ebook reader manufacturers—interact, they know that their actions significantly affect each other's welfare or profit, so they consider those actions carefully. Firms compete on many fronts beyond setting quantity or price. To gain an edge over rivals, a firm makes many decisions, such as which ebook standard to use, how much to advertise, whether to act to discourage a new firm from entering its market, how to differentiate its product, and whether to invest in new equipment.

An oligopolistic firm that ignores or inaccurately predicts the behavior of rivals is unlikely to do well. If Ford underestimates how many cars Toyota and Honda will produce, Ford may produce too many vehicles and lose money. These firms are aware of this strategic interdependence, recognizing that the plans and decisions of any one firm might significantly affect the profits of the other firms.

game theory
a set of tools that economists and others use to analyze decision making by players who use strategies

In this chapter, we use game theory (von Neumann and Morgenstern, 1944) to examine how a small number of firms or individuals interact. **Game theory** is a set of tools that economists, political scientists, military analysts, and others use to analyze players' strategic decision making. This chapter introduces the basic concepts of game theory.[1]

Game theory has many practical applications. It is particularly useful for analyzing how oligopolistic firms set prices, quantities, and advertising levels. Economists also use game theory to analyze bargaining between unions and management or between the buyer and seller of a car, interactions between polluters and those harmed by pollution, transactions between the buyers and sellers of homes, negotiations between parties with different amounts of information (such as between car owners and auto mechanics), bidding in auctions, and many other economic interactions. Beyond economics, game theory is used by political scientists and military planners to avoid or fight wars; by biologists to analyze evolutionary biology and ecology; and by philosophers, computer scientists, and many others.

game
any competition between players (firms) in which strategic behavior plays a major role

strategy
a battle plan that specifies the action that a player will make conditional on the information available at each move and for any possible contingency

action
a move that a player makes at a specified stage of a game, such as how much output a firm produces in the current period

payoffs
benefits received by players from the game's outcome, such as profits for firms, or incomes or utilities for individuals

rules of the game
regulations that include the timing of players' moves, the various actions that a player may make at each stage of the game, and other aspects of how the game is played

In this chapter, we concentrate on how oligopolistic firms interact within a *game*. A **game** is an interaction between players (such as individuals or firms) in which players use strategies. A **strategy** is a battle plan that specifies the *actions* that a player will make. An **action** is a move that a player makes at a specified stage of a game, such as how much output a firm produces in the current period.

For example, a firm may use a simple business strategy where it produces 100 units of output regardless of what any rival does. Here, the strategy consists of a single action—producing 100 units of output. However, some strategies consist of a combination of actions or moves, possibly contingent on what a rival does. For example, a firm might decide to produce a small quantity as long as its rival produced a small amount in the previous period, and a large quantity otherwise.

The **payoffs** of a game are the benefits received by players from the game's outcome, such as profits for firms, or incomes or utilities for individuals. A *payoff function* specifies each player's payoff as a function of the strategies chosen by all players. We normally assume that players seek to maximize their payoffs. In essence, this assumption simply defines what we mean by payoffs. Payoffs include all relevant benefits experienced by the players. Therefore, rational players should try to obtain the highest payoffs they can.

The **rules of the game** include the *timing* of players' moves (such as whether one player moves first), the various actions that are possible at a particular point in the game, and possibly other specific aspects of how the game is played. A full description of a game normally includes a statement of the players, the rules of the game (including the possible actions or strategies), and the payoff function, along with a statement regarding the information available to the players.

---

[1]For more details, see, for example, Fudenberg and Tirole (1991) or Gibbons (1992). For an interesting, brief history, see **http://www.econ.canterbury.ac.nz/personal_pages/paul_walker/gt/hist.htm**.

When analyzing a game, we usually have three objectives: to accurately describe and understand the game, to predict the likely outcome of the game, and to offer advice to managers as to how best to play the game.

In analyzing a game, it is crucial that we know how much information participants have. We start by assuming that all the relevant information is *common knowledge* to the players and then we relax that assumption. **Common knowledge** is a piece of information known by all players, and it is known by all players to be known by all players, and it is known to be known to be known, and so forth. In particular, we initially assume that players have **complete information**, a situation in which the strategies and payoffs of the game are *common knowledge*.

common knowledge a piece of information known by all players, that is known to be known by all players, and so forth

complete information a situation in which the strategies and payoffs of the game are common knowledge

The information that firms possess affects the outcome of a game. The outcome of a game in which a particular piece of information is known by all firms may differ from the outcome when some firms are uninformed. A firm may suffer a worse outcome if it does not know the potential payoffs of other firms. Similarly, a firm may do worse if it has limited ability to make calculations, as when its cost of making many calculations is prohibitively high or its managers have limited analytical abilities. We address nonrational behavior at the end of the chapter.

In this chapter, we examine five main topics

1. **Static Games.** A static game is played once by players who act simultaneously and hence, at the time they make a decision, do not know how other players will act.
2. **Repeated Dynamic Games.** If a static game is repeated over many periods, firms may use more complex strategies than in the static one-period game because a firm's action in one period may affect its rivals' actions in subsequent periods.
3. **Sequential Dynamic Games.** If one firm acts before its rival, it may gain an advantage by converting what would be an empty threat to its rival into a credible, observable action.
4. **Auctions.** An auction is a game where bidders have incomplete information about the value that other bidders place on the auctioned good or service.
5. **Behavioral Game Theory.** Some people make biased decisions based on psychological factors rather than using a rational strategy.

14.1 Static Games

In accordance with our principles of free enterprise and healthy competition, I'm going to ask you two to fight to the death for it. —Monty Python

We start by examining **static games**, in which each player acts only once and the players act simultaneously (or, at least, each player acts without knowing rivals' actions). In these games, firms have complete information about the payoff functions but imperfect information about rivals' moves.

static games games in which each player acts only once and the players act simultaneously

Examples include two rival firms making simultaneous one-time-only decisions on where to locate its new factory, teenagers' game of chicken in cars, a duel (such as between Aaron Burr and Alexander Hamilton in 1804), an employer's hiring negotiations with a potential employee, street vendors' choice of locations and prices outside a Super Bowl or World Cup game, and the Cournot and Bertrand models (Chapter 13).

We analyze a game between two firms in which each firm can each take one of only two possible actions and play each other only once. Our example is a simplified version of the real-world competition (as estimated by Brander and Zhang, 1990) between United and American Airlines.

The game has the following characteristics. The two *players* or firms are United and American. They play a *static game*—they compete only once. The *rules* of the game specify the possible actions or strategies that the firms can take and when they can take them. Each firm has only two possible *actions*: Each can fly either 48 thousand or 64 thousand passengers per quarter between Chicago and Los Angeles. Other than announcing their output levels, the firms cannot communicate, so they cannot make side deals or otherwise coordinate their actions. Each firm's *strategy* is to take one of the two actions, choosing either a low output (48 thousand passengers per quarter) or a high output (64 thousand). The firms announce their actions or strategies *simultaneously*.

The firms have *complete information*: They are aware of the possible strategies and the corresponding payoff (profit) to each firm. However, their information is imperfect in one important respect. Because they choose their output levels simultaneously, neither airline knows what action its rival will take when it makes its output decision.

In this section, we show how to represent these static games in a table and how to predict their outcomes.

Normal-Form Games

When you have eliminated the impossible, whatever remains, however improbable, must be the truth. —Sherlock Holmes (Sir Arthur Conan Doyle)

normal form
a representation of a static game with complete information specifying the players, their possible strategies, and their payoffs for each combination of strategies

We examine a **normal-form** representation of a static game with complete information that specifies the players in the game, their possible strategies, and the payoff function that identifies the players' payoffs for each combination of strategies. The normal-form representation of this static game is the *payoff matrix* or *profit matrix* in Table 14.1.

This payoff matrix shows the profits for each of the four possible combinations of the strategies that the firms may choose. For example, if American chooses a large quantity, $q_A = 64$ units per quarter, and United chooses a small quantity, $q_U = 48$ units per quarter, the firms' profits are in the cell in the lower-left corner of the profit matrix. That cell shows that American's profit (upper-right number) is \$5.1 million per quarter, and United's profit (bottom-left number) is \$3.8 million per quarter. We now have a full description of the game, including a statement of the players, the rules, a list of the allowable strategies, the payoffs, and the available information.

Because the firms choose their strategies simultaneously, each firm selects a strategy that maximizes its profit *given what it believes the other firm will do*. The firms

Table 14.1 Dominant Strategies in a Quantity Setting, Prisoners' Dilemma Game
MyLab Economics Video

		American Airlines	
		$q_A = 64$	$q_A = 48$
United Airlines	$q_U = 64$	4.1 4.1	3.8 5.1
	$q_U = 48$	5.1 3.8	4.6 4.6

are playing a *noncooperative game of imperfect information* in which each firm must choose an action before observing the simultaneous action by its rival. Thus, while the players have complete information about all players' payoffs, they have imperfect information about how the other will act.

We can predict the outcome of some games by using the insight that rational players will avoid strategies that are *dominated* by other strategies. First, we show that in some games we can predict a game's outcome if each firm has a single best strategy that dominates all others. Then, we show that in other games, by sequentially eliminating dominated strategies, we are left with a single outcome. Finally, we note that the outcome of a broader class of games can be precisely predicted based on each player's choosing a *best response* to the other players' actions—the response that produces the largest possible payoff.

dominant strategy
a strategy that produces a higher payoff than any other strategy the player can use for every possible combination of its rivals' strategies

Dominant Strategies We can precisely predict the outcome of any game in which every player has a **dominant strategy**: a strategy that produces a higher payoff than any other strategy the player can use for every possible combination of its rivals' strategies. When a firm has a dominant strategy, a firm can hold no belief about its rivals' choice of strategies that would cause it to choose one of its other, strictly dominated strategies.

Although firms do not always have dominant strategies, they have them in our airline game. American's managers can determine its dominant strategy using the following reasoning:

- *If United chooses the high-output strategy* ($q_U = 64$), *American's high-output strategy maximizes its profit*: Given United's strategy, American's profit is \$4.1 million (top-right number in the upper-left cell) with its high-output strategy ($q_A = 64$) and only \$3.8 million (top-right number in the upper-right cell) with its low-output strategy ($q_A = 48$). Thus, American is better off using a high-output strategy if United chooses its high-output strategy.
- *If United chooses the low-output strategy* ($q_U = 48$), *American's high-output strategy maximizes its profit*: Given United's strategy, American's profit is \$5.1 million with its high-output strategy and only \$4.6 million with its low-output strategy.
- *Thus, the high-output strategy is American's dominant strategy*: Whichever strategy United uses, American's profit is higher if it uses its high-output strategy. We show that American won't use its low-output strategy (because that strategy is dominated by the high-output strategy) by drawing a vertical, dark-red line through American's low-output cells in Table 14.1.

By the same type of reasoning, United's high-output strategy is also a dominant strategy. We draw a horizontal, light-red line through United's low-output strategy. Because the high-output strategy is a dominant strategy for both firms, we can predict that the outcome of this game is the pair of high-output strategies, $q_A = q_U = 64$. We show the resulting outcome—the cell in Table 14.1 where both firms use high-output strategies—by coloring that cell green.

A striking feature of this game is that the players choose strategies that do not maximize their joint profit. Each firm earns \$4.6 million if $q_A = q_U = 48$ rather than the \$4.1 million they actually earn by setting $q_A = q_U = 64$. In this type of game—called a **prisoners' dilemma** game—all players have dominant strategies that lead to a profit (or another payoff) that is inferior to what they could achieve if they cooperated and pursued alternative strategies.

prisoners' dilemma
a game in which all players have dominant strategies that result in profits (or other payoffs) that are inferior to what they could achieve if they used cooperative strategies

The prisoners' dilemma crops up in virtually every cops-and-robbers show. The cops arrest Larry and Duncan and put them in separate rooms so that they cannot talk to each other. An assistant district attorney tells Larry, "We have enough evidence to convict you both of a relatively minor crime for which you'll serve a year in prison. If you'll squeal on your partner and he stays silent, we can convict him of a major crime for which he'll serve five years and you'll go free. If you both confess, you'll each get two years." Meanwhile, another assistant district attorney is making Duncan the identical offer. By the same reasoning as in the airline example, both Larry and Duncan confess even though they are better off if they both keep quiet.

best response
the strategy that maximizes a player's payoff given its beliefs about its rivals' strategies

Nash equilibrium
a set of strategies such that, when all other players use these strategies, no player can obtain a higher payoff by choosing a different strategy

Best Response and Nash Equilibrium Many games do not have a dominant strategy solution. For these games, we use a more general approach. For any given set of strategies chosen by rivals, a player wants to use its **best response**—the strategy that maximizes a player's payoff given its beliefs about its rivals' strategies. We illustrated this idea in Chapter 13 when we derived the best-response curves for an oligopolistic firm.

A dominant strategy is a strategy that is a best response to *all possible* strategies that a rival might use. Thus, a dominant strategy is a best response. However, even if a dominant strategy does not exist, each firm can determine its best response to *any possible* strategies chosen by its rivals.

The idea that players use best responses is the basis for the Nash equilibrium, a solution concept for games formally introduced by John Nash (1951). A set of strategies is a **Nash equilibrium** if, when all other players use these strategies, no player can obtain a higher payoff by choosing a different strategy.[2] An appealing property of the Nash equilibrium is that it is self-enforcing. If the other players use a Nash equilibrium strategy, then no player wants to switch to another strategy. Each player believes that "Given the strategies chosen by my rivals, I made the best possible choice—I chose my best response."

The Nash equilibrium is the primary solution concept used by economists in analyzing games. It allows us to find solutions to more games than just those with a dominant strategy solution. If a game has a dominant strategy solution, then that solution must also be a Nash equilibrium. However, many games that do not have dominant strategy solutions have a Nash equilibrium.

To illustrate these points, we examine a more complex simultaneous-move game in which American and United can each produce an output of 48, 64, or 96 (thousand passengers per quarter). This game has nine possible output combinations, as the 3×3 profit matrix in Table 14.2 shows. Neither American nor United has a single, dominant strategy, but we can find a Nash equilibrium by using a two-step procedure. First, we determine each firm's best response to any given strategy of the other firm. Second, we check whether any pairs of strategies (cells in the profit matrix) are best responses for both firms. Each such pair of strategies is a Nash equilibrium.

We start by determining American's best response for each one of United's possible actions. If United chooses $q_U = 96$ (thousand passengers per quarter), the first row of the table, then American's profit is \$0 if it sets $q_A = 96$ (the first column), \$2.0 (million) if it chooses $q_A = 64$ (the second column), and \$2.3 if it selects $q_A = 48$ (third column). Thus, American's best response if United sets $q_U = 96$ is to select $q_A = 48$. We indicate American's best response by coloring the upper triangle in the last (third column) cell in this row dark green. Similarly, if United sets $q_U = 64$ (second row), American's best response is to set $q_A = 64$, where it earns \$4.1 million, so we color

[2]In Chapter 13, we used a special case of this definition of a Nash equilibrium in which we referred to actions instead of strategies. An action and a strategy are the same if the players can move only once; however, later in this chapter, we will consider games that last for many periods and hence need a definition based on strategies.

Table 14.2 Best Responses in a Quantity Setting Game MyLab Economics Video

		American Airlines		
		$q_A = 96$	$q_A = 64$	$q_A = 48$
United Airlines	$q_U = 96$	0 (United), 0 (American)	3.1 (United), 2.0 (American)	4.6 (United), 2.3 (American)
	$q_U = 64$	2.0 (United), 3.1 (American)	4.1 (United), 4.1 (American)	5.1 (United), 3.8 (American)
	$q_U = 48$	2.3 (United), 4.6 (American)	3.8 (United), 5.1 (American)	4.6 (United), 4.6 (American)

the upper triangle in the middle cell (second column) of the second row dark green. Finally, if United sets $q_U = 48$ (third row), American's best response is $q_A = 64$, where it earns \$5.1 million, so we color the upper triangle in the middle cell of the third row dark green.

We can use the same type of reasoning to determine United's best responses to each of American's strategies. If American chooses $q_A = 96$ (first column), then United maximizes its profit at \$2.3 million by setting $q_U = 48$, which we indicate by coloring the lower triangle light green in the lower left cell of the table. Similarly, we show that United's best response is $q_U = 64$, if American sets $q_A = 64$ or 48, which we show by coloring the relevant lower left triangles light green.

We now look for a Nash equilibrium, which is a pair of strategies where both firms are using a best-response strategy so that neither firm would want to change its strategy. This game has only one cell in which both the upper and lower triangles are green: $q_A = q_U = 64$. Given that its rival uses this strategy, neither firm wants to deviate from its strategy. For example, if United continued to set $q_U = 64$, but American raised its quantity to 96, American's profit would fall from \$4.1 to \$3.1. Or, if American lowered its quantity to 48, its profit would fall to \$3.8. Thus, American does not want to change its strategy.

Because no other cell has a pair of strategies that are best responses (green lower and upper triangles), at least one of the firms would want to change its strategy in each of these other cells. For example, at $q_A = q_U = 48$, either firm could raise its profit from \$4.6 to \$5.1 million by increasing its output to 64. At $q_A = 48$ and $q_U = 64$, American can raise its profit from \$3.8 to \$4.1 million by increasing its quantity to $q_A = 64$. Similarly, United would want to increase its output when $q_A = 64$ and $q_U = 48$. None of the other strategy combinations is a Nash equilibrium because at least one firm would want to deviate. Thus, we were able to find the single Nash equilibrium to this game by determining each firm's best responses.

In these airline examples, we have assumed that the firms can pick only from a small number of output levels. However, we can use game theory to find the Nash equilibrium in games in which the firms can choose any output level. We showed such a generalization for the airline example in Chapter 13. In Figure 13.3, we determined the best-response curves for each of these airlines, found that these best-response curves intersected only once, and identified the set of outputs at that intersection as the Nash-Cournot equilibrium. Indeed, that equilibrium is the same as the equilibria in Tables 14.1 and 14.2.

Failure to Maximize Joint Profits

The dominant-strategy analysis in Table 14.1 and the best-response analysis in Table 14.2 show that noncooperative firms may not reach the joint profit-maximizing outcome. Whether players achieve the outcome that maximizes joint profit depends on the profit matrix.

We illustrate this idea using an advertising example.

Common Confusion: Firms necessarily raise their profits by advertising.

We'll show that, for some profit matrices, all the firms would benefit if they could agree not to advertise.

Table 14.3 shows two possible profit matrices for an advertising game in which each firm can choose whether to advertise. In the Nash equilibrium, collective profit is not maximized in the first game but is maximized in the second game.

The game in panel a is a prisoners' dilemma game similar to the airline game in Table 14.1. In this game, a firm's advertising does not bring new customers into the market but only has the effect of stealing business from the rival firm. Because each firm must decide whether or not to advertise at the same time, neither firm knows the strategy of its rival when it chooses its strategy.

If neither firm advertises, then each firm makes a profit of 2 (say, $2 million), as the upper-left cell of the profit matrix in panel a shows. If Firm 1 advertises but Firm 2

Table 14.3 Advertising Game MyLab Economics Video

(a) Advertising Only Takes Customers from Rivals

		Firm 1	
		Do Not Advertise	*Advertise*
Firm 2	*Do Not Advertise*	Firm 1: 2 Firm 2: 2	Firm 1: 3 Firm 2: 0
	Advertise	Firm 1: 0 Firm 2: 3	Firm 1: 1 Firm 2: 1

(b) Advertising Attracts New Customers to the Market

		Firm 1	
		Do Not Advertise	*Advertise*
Firm 2	*Do Not Advertise*	Firm 1: 2 Firm 2: 2	Firm 1: 4 Firm 2: 3
	Advertise	Firm 1: 3 Firm 2: 4	Firm 1: 5 Firm 2: 5

does not, then Firm 1 takes business from Firm 2 and raises its profit to 3, while the profit of Firm 2 is reduced to 0. The gain to Firm 1 is less than the loss to Firm 2 because the revenue that is transferred from Firm 2 to Firm 1 as customers shift is partially offset by the cost of Firm 2's advertising. If both firms advertise, then each firm gets a profit of 1, as the cell on the lower right shows.

Advertising is a dominant strategy for both firms.[3] We use red lines to show that the firms do not use the dominated do-not-advertise strategies. The outcome in which both firms advertise is therefore a dominant strategy solution. Advertising for both firms is also a Nash equilibrium because each firm is choosing its best response to the other firm's strategy.

In this Nash equilibrium, each firm earns 1, which is less than the 2 it would make if neither firm advertised. Thus, *the sum of the firms' profits is not maximized in this simultaneous-choice one-period game*.

Many people are surprised when they see this result. Why don't the firms cooperate, refrain from advertising, and earn 2 instead of 1? This game is an example of a prisoners' dilemma: the game has a dominant strategy solution in which the players receive lower profits than they would get if the firms could cooperate. Each firm makes more money by advertising regardless of the strategy used by the other firm, even though their joint profit is maximized if neither advertises.

The reason they don't cooperate is a lack of trust. Each firm uses the no-advertising strategy only if the firms have a binding (enforceable) agreement. The reason they do not trust each other is that each firm knows it is in the other firm's best interest to deviate from the actions that would maximize joint profits.[4]

In contrast, in panel b, when either firm advertises, the promotion attracts new customers to both firms. That is, each firm's advertising has a market expansion effect. If neither firm advertises, both earn \$2 (million). If only one firm advertises, its profit rises to \$4, which is more than the \$3 that the other firm makes.

If both advertise, they are collectively better off than if only one advertises or neither advertises. Again, advertising is a dominant strategy for a firm because it earns more by advertising regardless of the strategy the other firm uses. This dominant strategy solution is a Nash equilibrium, but this game is not a prisoners' dilemma. This Nash equilibrium maximizes the firms' combined profits, which is the same outcome that would arise if the firms could cooperate.

Thus, whether the Nash equilibrium maximizes the combined profit for the players depends on the properties of the game summarized by the profit matrix. The distinction in these two games is that the Nash equilibrium in which both advertise is the same as the collusive equilibrium in panel b where advertising increases the market size, but it is not the collusive equilibrium in panel a where advertising cannibalizes the sales of the other firm.

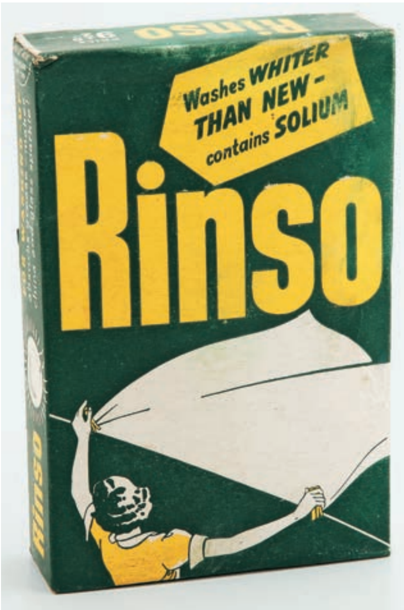
---

[3]Firm 1 goes through the following reasoning. "If my rival does not advertise, I get 2 if I do not advertise and I get 3 if I do advertise, so advertising is better. If my rival does advertise, I get 0 if I do not advertise and I get 1 if I do advertise, so advertising is still better." Regardless of what Firm 2 does, advertising is better for Firm 1, so advertising is a dominant strategy for Firm 1 (and not advertising is a dominated strategy). Firm 2 faces a symmetric problem, so advertising is a dominant strategy.

[4]Suppose the two firms meet in advance and agree not to advertise. However, if the firms are going to engage in this game only once, each has an incentive to cheat on the agreement. For example, if Firm 1 believes that Firm 2 will stick to the agreement and not advertise, Firm 1 can increase its profit from \$2 million to \$3 million by violating the agreement and advertising.

Application

Strategic Advertising



Firms with market power, such as oligopolies, often advertise.[5] In 2015, the largest advertiser in the United States was Procter & Gamble, the producer of Crest toothpaste, Pampers diapers, and various other household products. Procter & Gamble spent \$4.3 billion in the United States on advertising and \$8.3 billion globally (11% of its net sales worldwide).

In oligopoly markets, firms consider the likely actions of their rivals when deciding how much to advertise. How much a firm should spend on advertising depends critically on whether the advertising helps or harms its rival.

For example, when a firm advertises to inform consumers about a new use for its product, its advertising may cause the quantity demanded for its own *and* rival brands to rise, as happened with toothpaste ads. Before World War I, only 26% of Americans brushed their teeth. By 1926, in part because of ads like those in Ipana's "pink toothbrush" campaign, which detailed the perils of bleeding gums, the share of Americans who brushed rose to 40%. Ipana's advertising helped all manufacturers of toothbrushes and toothpaste.[6]

Alternatively, a firm can raise its profit if it can convince consumers that its product is superior to other brands. For example, from the 1930s through the early 1970s, *secret ingredients* were a mainstay of consumer advertising. These ingredients were named using combinations of letters and numbers to give the impression they were developed in laboratories rather than by Madison Avenue. Dial soap boasted that it contained AT-7. Rinso detergent had solium, Comet included Chlorinol, and Bufferin had di-alminate. Among the toothpastes, Colgate had Gardol, Gleem had GL-70, Crest had fluoristan, and Ipana had hexachlorophene and Durenamel.

Empirical evidence indicates that the impact of a firm's advertising on other firms varies across industries. The cola market is an example of the extreme case in which a firm's advertising brings few new customers into the market and primarily serves to steal business from rivals. Gasmi, Laffont, and Vuong (1992) reported that Coke's or Pepsi's gain from advertising comes at the expense of its rivals; however, cola advertising has almost no effect on total market demand, as in panel a. Similarly, advertising by one brand of an erectile dysfunction drug increases its share and decreases that of its rivals (David and Markowitz, 2011). A more extreme example is comparative over-the-counter analgesics advertising, which harms rivals more than it helps the advertising firm, so total market profit falls (Anderson, 2016).

At the other extreme is cigarette and beer advertising. Roberts and Samuelson (1988) found that cigarette advertising increases the size of the market but does not change market shares substantially, as in panel b.[7] Similarly, Chandra and Weinberg (2015) find a positive spillover of a beer manufacturer's advertising on its rivals.

Intermediate results include saltine crackers (Slade, 1995) and Canadian fast-foods, where advertising primarily increases general demand but has a small effect on market share (Richards and Padilla, 2009), and CPUs where Intel's advertising has a smaller effect on total market demand than on Intel's share (Salgado, 2008).

[5] Under perfect competition, there is no reason for an individual firm to advertise as a firm can sell as much as it wants at the market price.

[6] Although it's difficult to believe, starting in the 1970s, Wisk liquid detergent claimed that it solved a major social problem: ring around the collar (**https://www.youtube.com/watch?v=e3N_skYSGoY**). Presumably, some consumers—even among those who were gullible enough to find this ad compelling—could generalize that applying other liquid detergents would work equally well.

[7] However, the Centers for Disease Control and Prevention's evidence suggests that advertising may shift the brand loyalty of youths.

Multiple Equilibria

Many oligopoly games have more than one Nash equilibrium. We illustrate this possibility with an entry game. Two firms are each considering opening a gas station at a highway rest stop that has no gas stations. The rest stop has enough physical space for at most two gas stations. The profit matrix in Table 14.4 shows that the demand for gasoline is adequate for only one station to operate profitably. If both firms enter, each loses \$1 (hundred thousand). Neither firm has a dominant strategy. Each firm's best action depends on what the other firm does.

By examining the firm's best responses, we can identify two Nash equilibria: Firm 1 *enters* and Firm 2 *does not enter*, or Firm 2 *enters* and Firm 1 *does not enter*. Each is a Nash equilibrium because neither firm wants to change its behavior. Given that Firm 2 does not enter, Firm 1 does not want to change its strategy from entering to staying out of the market. If it changed its behavior, it would go from earning \$1 to earning \$0. Similarly, given that Firm 1 enters, Firm 2 does not want to switch its behavior and enter because it would lose \$1 instead of making \$0. The outcome in which only Firm 2 enters is also a Nash equilibrium by the same type of reasoning.

How do the players know which (if any) Nash equilibrium will result? They *don't* know. It is difficult to see how the firms choose strategies unless they collude and can enforce their agreement. For example, the firm that enters could pay the other firm to stay out of the market. Without an enforceable collusive agreement, even discussions between the firms before they make decisions are unlikely to help. These pure Nash equilibria are unappealing because they call for identical firms to use different strategies.

Table 14.4 Nash Equilibria in an Entry Game

		Firm 1	
		Do Not Enter	*Enter*
Firm 2	*Do Not Enter*	0 0	1 0
	Enter	0 1	–1 –1

Solved Problem 14.1

MyLab Economics Solved Problem

Intel and AMD are the dominant central processing unit manufacturers. Assume they play the following game once and act simultaneously. Their profits are symmetric (which is consistent with the estimates of Salgado, 2008). If both choose low levels of advertising, Intel's profit, π_I, and AMD's profit, π_A, are each 2. If both choose high, each earns 3. If Intel's advertising is high and AMD's is low, $\pi_I = 8$ and $\pi_A = 4$. If Intel's advertising is low and AMD's is high, $\pi_I = 4$ and $\pi_A = 8$. Describe how each firm chooses its strategy? Describe the Nash equilibrium or equilibria.

Answer

1. *Use a profit matrix to show the firms' best responses.* The payoff matrix shows the four possible pairs of strategies and the associated profits. If Intel chooses a low level of advertising (top row), AMD's profit is 2 if its advertising is low and

8 if it is high, so its best response is high, as indicated by the dark green triangle in the upper right of the top right cell. If Intel's advertising is high (bottom row), AMD's profit is 4 if its advertising is low and 3 if it is high, so its best response is low, as indicated by the dark green triangle in the upper right of the lower-left cell. Similarly, we use light green triangles in the lower left of cells to show Intel's best responses.

		AMD	
		Low	*High*
Intel	*Low*	2 (AMD: upper right) / 2 (Intel: lower left)	8 / 4
	High	4 / 8	3 / 3

2. *Identify the Nash equilibria using the best responses*. For a pair of strategies to be a Nash equilibrium, both firms must be using a best response. Thus, this game has two Nash equilibria: Intel's advertising is high and AMD's is low (lower-left cell) and Intel's advertising is low and AMD's is high (upper-right cell), where the cells are green.

pure strategy
each player chooses an action with certainty

mixed strategy
a firm (player) chooses among possible actions according to probabilities it assigns

Mixed Strategies

In each of the games that we have considered so far, including the simultaneous-move entry game in Table 14.4, we have assumed that the firms use a **pure strategy**: Each player chooses a single action. In addition to using a pure strategy, a firm in the entry game may employ a **mixed strategy** in which the player chooses among possible actions according to probabilities it assigns. A pure strategy assigns a probability of 1 to a single action, whereas a mixed strategy is a probability distribution over actions. That is, a pure strategy is a rule telling the player what action to take, whereas a mixed strategy is a rule telling the player which dice to throw, coin to flip, or other device to use to choose an action.

How business decisions are made.

In the entry game, both firms may use the same mixed strategy: Both firms enter with a probability of one-half—say, if a flipped coin comes up heads. This pair of mixed strategies is a Nash equilibrium because neither firm wants to change its strategy, given that the other firm uses its Nash equilibrium mixed strategy.

If both firms use this mixed strategy, each of the four outcomes in the payoff matrix in Table 14.4 is equally likely. The probability that the outcome in a particular cell of the matrix occurs is the product of the probabilities that each player chooses the relevant action.

The probability that a player chooses a given action is $\frac{1}{2}$, so the probability that both players will choose a given pair of actions is $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$, because their actions are independent. Consequently, Firm 1 has a one-fourth chance of earning \$1 (upper-right cell), a one-fourth chance of losing \$1 (lower-right cell), and a one-half chance of earning \$0 (upper-left and lower-left cells). Thus, Firm 1's expected profit—the firm's profit in each possible outcome times the probability of that outcome—is

$$\left(\$1 \times \tfrac{1}{4}\right) + \left(-\$1 \times \tfrac{1}{4}\right) + \left(\$0 \times \tfrac{1}{2}\right) = \$0.$$

Given that Firm 1 uses this mixed strategy, Firm 2 cannot achieve a higher expected profit by using a pure strategy. If Firm 2 uses the pure strategy of entering with probability 1, it earns \$1 half the time and loses \$1 the other half, so its expected profit is \$0. If Firm 2 stays out with certainty, it earns \$0 with certainty.

If Firm 2 believes that Firm 1 will use its equilibrium mixed strategy, Firm 2 is indifferent as to which pure strategy it uses (of the strategies that have a positive probability in that firm's mixed strategy). In contrast, if one of the actions in the equilibrium mixed strategy has a higher expected payoff than some other action, it would pay to increase the probability that Firm 2 takes the action with the higher expected payoff. However, if all of the pure strategies that have positive probability in a mixed strategy have the same expected payoff, then the expected payoff of the mixed strategy must also have that expected payoff. Thus, Firm 2 is indifferent as to whether it uses any of these pure strategies or any mixed strategy over these pure strategies.

In our example, why would a firm pick a mixed strategy where its probability of entering is one-half? In a symmetric game such as this one, we know that both players have the same probability of entering, θ. Moreover, for Firm 2 to use a mixed strategy it must be indifferent between entering or not entering if Firm 1 enters with probability θ. Firm 2's payoff from entering is $[\theta \times (-1)] + [(1 - \theta) \times 1] = 1 - 2\theta$. Its payoff from not entering is $[\theta \times 0] + [(1 - \theta) \times 0] = 0$. Equating these two expected profits, $1 - 2\theta = 0$, and solving, we find that $\theta = \frac{1}{2}$ Thus, both firms using a mixed strategy where they enter with a probability of one-half is a Nash equilibrium.

The entry game has two pure-strategy Nash equilibria—one firm employing the pure strategy of entering and the other firm pursuing the pure strategy of not entering—and a mixed-strategy Nash equilibrium. If Firm 1 decides to *enter* with a probability of $\frac{1}{2}$, Firm 2 is indifferent between choosing to enter with a probability of 1 (the pure strategy of *enter*), 0 (the pure strategy of *do not enter*), or any fraction in between these extremes. However, for the firms' strategies to constitute a mixed-strategy Nash equilibrium, both firms must choose to enter with a probability of one-half.

An important reason for introducing the concept of a mixed strategy is that some games have no pure-strategy Nash equilibria. However, Nash (1950) proved that every static game with a finite number of players and a finite number of actions has at least one Nash equilibrium, which may involve mixed strategies.

Some game theorists argue that mixed strategies are implausible because firms do not flip coins to choose strategies. One response is that firms may only appear to be unpredictable. In this game with no dominant strategies, neither firm has a strong reason to believe that the other will choose a pure strategy. It may think about its rival's behavior as random. However, in actual games, a firm may use some information or reasoning that its rival does not observe so that it chooses a pure strategy. Another response is that a mixed strategy may be appealing in some games, such as the entry game, where a random strategy and symmetry between players are plausible.

Application
Tough Love

We can use game theory to explain many interactions between parents and their kids. In the United States, the term *boomerang generation* refers to young adults who return home after college, a first job, or the military to live with their parents. (The Japanese call them *parasite singles*.)

The Great Recession hit young people particularly hard. The U.S. unemployment rate for 20- to 24-year-olds went from 8.5% in 2007 to 11% in 2008, then rose to 16% in 2009, and has stayed above 13% through 2012, but fell to 10% by 2015.

As a result, more adult offsprings moved back to live with their parents after college. The share of 25- to 34-year-olds living in multigenerational households rose from 11% in 1980 to 22% in 2007 and 24% in 2010 (the highest it has been since the 1950s), and continued to rise to 26% in 2015.[8] In the European Union, over half (56%) of young adults, 20–29 years old, lived with their parents in 2013, ranging from 10% in Denmark to 84% in Slovakia. In 2015, 48% of U.K. 20–24-year-olds and 25% of 25–29-year-olds lived with their parents.

In many parents' minds the question arises whether by supporting their kids, they discourage them from working. Rather than unconditionally supporting their children, would they help their kids more by engaging in tough love: kicking their kids out and making them support themselves? The following Solved Problem addresses this question.

Solved Problem 14.2

Mimi wants to support her son Jeff if he looks for work but not otherwise. Jeff (unlike most young people) wants to look for a job only if his mother will not support his life of indolence. Their payoff matrix is

		Jeff	
		Look for Work	*Loaf*
Mimi	*Support*	2 4	4 –1
	No Support	1 –1	0 0

If Jeff and Mimi choose actions simultaneously, what are the pure- or mixed-strategy equilibria?

Answer

1. *Use a best-response analysis to determine whether any of the four possible pairs of pure strategies is a Nash equilibrium.* We start by determining each player's best responses. If Jeff loafs, Mimi's best response is to cut off support because her payoff is 0 without support and −1 with support. We show her best response by a light-green triangle in the lower left of the no support/loaf cell.

[8]According to a 2015 Pew report, the share of parents financially helping an adult child was 61% in the United States, 60% in Italy, and 48% in Germany.

Similarly, her best response if he works is to support him (the payoff is 4 with support and −1 with no support), so we shade the lower-left triangle in the support/look cell. Jeff's best responses are the dark-green triangles in the upper right of the support/loaf and no support/look cells. In none of the four cells are both the triangles shaded. That is, none of these pairs of pure strategies is a Nash equilibrium because one or the other player would want to change his or her strategy.

2. *By equating expected payoffs, determine the mixed-strategy equilibrium.* If Mimi provides support with probability θ_M, Jeff's expected payoff from looking for work is $2\theta_M + [1 \times (1 - \theta_M)] = 1 + \theta_M$, and his expected payoff from loafing is $4\theta_M + [0 \times (1 - \theta_M)] = 4\theta_M$. His expected payoffs are equal if $1 + \theta_M = 4\theta_M$, or $\theta_M = \frac{1}{3}$. Similarly, if Jeff looks for work with probability θ_J, then Mimi's expected payoff from supporting him is $4\theta_J + [(-1) \times (1 - \theta_J)] = 5\theta_J - 1$, and her expected payoff from not supporting him is $-\theta_J + [0 \times (1 - \theta_J)] = -\theta_J$. By equating her expected payoffs, $5\theta_J - 1 = -\theta_J$, we determine that his mixed-strategy probability is $\theta_J = \frac{1}{6}$. Thus, although this game has no pure-strategy Nash equilibria, it has a mixed-strategy Nash equilibrium.

14.2 Repeated Dynamic Games

dynamic games
games in which players move either sequentially or repeatedly

In static, normal-form games, players have imperfect information about how other players will act because everyone moves simultaneously and only once. In contrast, in a **dynamic game**, players move either repeatedly or sequentially. Consequently, a player has perfect information about other players' previous moves.

We consider two types of dynamic games. We start with a *repeated* or *multiperiod* game in which a single-period, simultaneous-move game, such as the airline prisoners' dilemma game, is played at least twice and possibly many times. Although the players move simultaneously in each period, they know about their rivals' moves in previous periods, so a rival's previous move may affect a player's current action. As a result, it is a dynamic game.

In the next section, we turn to *sequential games*. We examine a *two-stage game*, which is played once and hence can be said to occur in a "single period." In the first stage, Player 1 moves. In the second stage, Player 2 moves and the game ends with the players receiving payoffs based on their actions.

Strategies and Actions in Dynamic Games

A major difference between static and dynamic games is that dynamic games require us to distinguish between strategies and actions. An *action* is a single move that a player makes at a specified time, such as choosing an output level or a price. A *strategy* is a battle plan that specifies the full set of actions that a player will make throughout the game and may involve actions that are conditional on prior actions of other players or on additional information available at a given time.

For example, American's strategy might state that it will fly 64 thousand passengers between Chicago and Los Angeles this quarter if United flew 64 thousand last quarter, but that it will fly only 48 thousand this quarter if United flew 48 thousand last quarter. This distinction between an action and a strategy is moot in a simultaneous-move static game, where an action and a strategy are effectively the same.

Cooperation in a Repeated Prisoners' Dilemma Game

To illustrate the difference between a static game and a repeated game, we consider a repeated prisoners' dilemma game. Each period has a single stage in which both players move simultaneously. However, these are dynamic games because Player 1's move in period t precedes Player 2's move in period $t + 1$; hence, the earlier action may affect the later one. The players know all the moves from previous periods, but they do not know each other's moves within this period because they will act simultaneously.

We showed that if American and United engage in a single-period prisoners' dilemma game, Table 14.1, the two firms produce more than they would if they colluded. Yet cartels do form. What's wrong with this theory, which says that cartels won't occur? One explanation is that markets last for many periods, and collusion is more likely to occur in a multi-period game than in a single-period game.

In a single-period game, one firm cannot punish the other firm for cheating on a cartel agreement. But if the firms play period after period, a wayward firm can be punished by the other.

Suppose now that the airlines' single-period prisoners' dilemma game Table 14.1 is repeated quarter after quarter. If they play a single-period game, each firm takes its rival's strategy as a given and assumes that it cannot affect that strategy. In a repeated game, a firm may devise a strategy for this period that depends on its rival's previous actions. For example, a firm may set a low output level for this period only if its rival set a low output level in the previous period. In a repeated game, a firm can influence its rival's behavior by *signaling* and by *threatening to punish*.

Signaling When antitrust laws make a firm hesitant to contact its rival directly, it may try to influence its rival's behavior by *signaling*. For example, American could use a low-quantity strategy for a couple of periods to signal United that it desires that the two firms cooperate and produce that low quantity in the future. If United does not respond by lowering its output in future periods, then American suffers lower profits for only a couple of periods. However, if United responds to this signal and lowers its quantity, both firms can profitably produce at the low quantity thereafter.

If the low-output strategy is so lucrative for everyone, why don't firms always cooperate when engaging in such indefinitely repeated games? One reason is that the cooperative outcome is not the only possible Nash equilibrium. This game has another Nash equilibrium in which each firm chooses the high output every period. If United believes that American will produce the high output in every period, then its best response is to produce the high output every period. This same reasoning also applies to American. Each firm's belief about its rival will be confirmed by experience and neither firm will have an incentive to change its strategy.

Threatening to Punish To induce cooperation, a firm can threaten to punish its rival for not restricting output. We use the profit matrix in Table 14.1 to illustrate how the airlines could threaten to punish rivals to ensure collusion. Suppose that American announces or somehow indicates to United that it will use the following two-part strategy:

- American will produce the smaller quantity each period as long as United does the same.
- If United produces the larger quantity in period t, American will produce the larger quantity in period $t + 1$ and all subsequent periods.

If United believes that American will follow this strategy, United knows that it will make \$4.6 million each period if it produces the smaller quantity. If United produces the large quantity its profit in that period rises to \$5.1 million, but it lowers its potential profit to \$4.1 million in each following period even if it continues to produce the high quantity. Thus, United gains half a million dollars relative to the cooperative payoff ($\$0.5 = \$5.1 - \$4.6$) in the period when it first defects from the cooperative output, but it loses half a million dollars relative to cooperation ($-\$0.5 = \$4.1 - \$4.6$) in each subsequent period. After only two punishment periods, the loss would be much larger in magnitude than the initial gain.[9]

American is using a *trigger strategy*: a strategy in which a rival's defection from a collusive outcome triggers a punishment. In this case, the trigger strategy is extreme because a single defection calls for a firm to punish its rival forever by producing the high output in all subsequent periods. However, if both firms adopt this trigger strategy, the outcome is a Nash equilibrium in which both firms choose the low output and obtain the collusive profit in every period: Defection and punishment need not occur. The firms may use less extreme trigger strategies. For example, a strategy that involved just two periods of punishment for a defection may make defection unattractive in this example.

Solved Problem 14.3

Show that if American and United Airlines know that they will play the game just described repeatedly for exactly T periods that the firms are unlikely to cooperate.

Answer

Start with the last period and work backward. In the last period, T, the firms know that they're not going to play again, so they know they can cheat—produce a large quantity—without fear of punishment. As a result, the last period is like a single-period game, and both firms produce the large quantity. That makes the $T - 1$ period the last interesting period. By the same reasoning, the firms will cheat in $T - 1$ because they know that they will both cheat in the last period and hence no additional punishment can be imposed. Continuing this type of argument, we conclude that maintaining an agreement to produce the small quantity will be difficult if the game has a known stopping point.

Comment: Playing the same game many times does not necessarily help the firms cooperate. With a known end period, cooperating is difficult. However, if the players know that the game will end but aren't sure when, cheating is less likely to occur. Cooperation is therefore more likely in a game that will continue forever or will end at an unknown period than in a game with a known final period.

14.3 Sequential Dynamic Games

We now turn to sequential dynamic games, in which one firm moves before another. We show how to represent these games diagrammatically and predict their outcomes.

extensive form
specifies the players, the sequence in which they move, the actions they can take, the information they have about other players' previous moves, and the payoffs over all possible strategies

Game Tree

Rather than use the normal form, economists analyze sequential dynamic games in their **extensive form**, which specifies the n players, the sequence in which they make their moves, the actions they can take at each move, the information that each player has about players' previous moves, and the payoff function over all possible

[9]Presumably a firm discounts future gains or losses (Chapter 16) because a dollar today is worth more than a dollar in the future. However, the effect of such discounting over a period as short as a few quarters is small.

strategies. In this section, we assume that players not only have complete information about the payoff function but also have perfect information about the play of the game to this point.

To illustrate a sequential-move, two-stage game, we assume that American Airlines can choose its output level before United does. This sequential-move game is called a *Stackelberg game* (see Chapter 13). The striking result of this analysis is that when one player can move before the other, the outcome is different from that in a game where they have to move simultaneously. For simplicity, we assume that American and United Airlines can choose only output levels of 96, 64, and 48 million passengers per quarter.

Using Table 14.2, the normal-form representation of this game, we derived the Nash equilibrium when both firms moved simultaneously. To demonstrate the role of sequential moves, we use an *extensive-form diagram* or *game tree*, Figure 14.1, which shows the order of the firms' moves, each firm's possible actions at the time of its move, and the resulting profits at the end of the game.

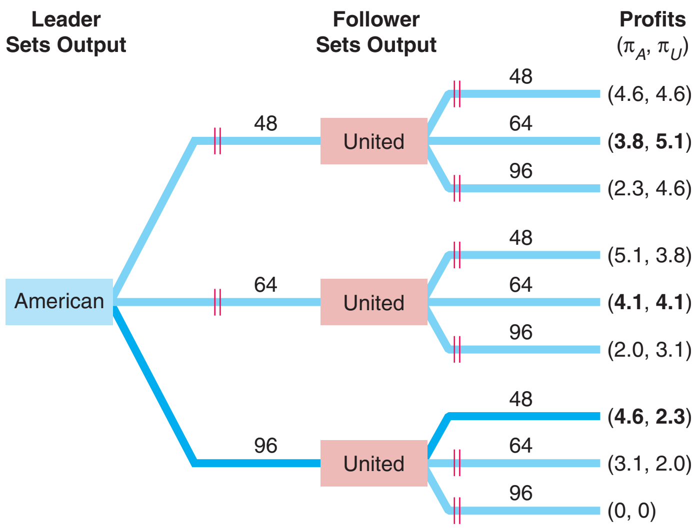
In the figure, each box is a point of decision by one of the firms, called a *decision node*. The name in the decision node box indicates that it is that player's turn to move. The lines or *branches* extending out of the box represent a complete list of the possible actions that the player can make at that point of the game. On the left side of the figure, American, the leader, starts by picking one of the three output levels. In the middle of the figure, United, the follower, chooses one of the three quantities after learning the output level American chose. The right side of the figure shows the profits that American and United earn, given that they sequentially took the actions to reach this final branch. For instance, if American selects 64 and then United chooses 96, American earns \$2.0 million profit per quarter and United earns \$3.1 million.

subgame
all the subsequent decisions that players may make given the actions already taken and corresponding payoffs

Within this game are *subgames*. At a given stage, a **subgame** consists of all the subsequent decisions that players may make given the actions already taken and corresponding payoffs. The second stage, where United makes a choice, has three possible subgames. In Figure 14.1, if in the first stage American chooses $q_A = 48$, the relevant subgame is the top node in the second stage and its three branches. This game has four subgames. The second stage has three subgames in which United makes a decision given each of American's three possible first-stage actions. The fourth subgame is the entire game at the time of the first-stage decision.

Figure 14.1 Stackelberg Game Tree MyLab Economics Video

American, the leader firm, chooses its output level first. Given American's choice, United, the follower, picks an output level. The right side of the figure shows the firms' profits that result from these decisions. Two red lines through an action line indicate that the firm rejects that action.



Subgame Perfect Nash Equilibrium

In solving a problem of this sort, the grand thing is to be able to reason backward. —Sherlock Holmes (Sir Arthur Conan Doyle)

subgame perfect Nash equilibrium
players' strategies are a Nash equilibrium in every subgame

To predict the outcome of this sequential game, we introduce a stronger version of the Nash equilibrium concept. A set of strategies forms a **subgame perfect Nash equilibrium** if the players' strategies are a Nash equilibrium in every subgame. As the entire dynamic game is a subgame, a subgame perfect Nash equilibrium is also a Nash equilibrium. In contrast, in a simultaneous-move game such as the static prisoners' dilemma, the only subgame is the game itself, so the Nash equilibrium and the subgame perfect Nash equilibrium are the same.

Table 14.2 shows the normal-form representation of this game in which the Nash equilibrium for the simultaneous-move game is for each firm to choose 64. However, if the firms move sequentially, the subgame perfect Nash equilibrium results in a different outcome.

backward induction
first determine the best response by the last player to move, next determine the best response for the player who made the next-to-last move, then repeat the process back to the move at the beginning of the game

We can solve for the subgame perfect Nash equilibrium using **backward induction**, where we first determine the best response by the last player to move, next determine the best response for the player who made the next-to-last move, and then repeat the process until we reach the move at the beginning of the game. In our example, we work backward from the decision by the follower, United, to the decision by the leader, American, moving from the right to the left side of the game tree.

How should American, the leader, select its output in the first stage? For each possible quantity it can produce, American predicts what United will do and picks the output level that maximizes its own profit. Thus, to predict American's action in the first stage, American determines what United, the follower, will do in the second stage, given each possible output choice by American in the first stage. Using its conclusions about United's second-stage reaction, American makes its first-stage decision.

United, the follower, does not have a dominant strategy. The amount it chooses to produce depends on the quantity that American chose. If American chose 96, United's profit is \$2.3 million if its output is 48, \$2.0 million if it produces 64, and \$0 if it picks a quantity of 96. Thus, if American chose 96, United's best response is 48. The double lines through the other two action lines show that United will not choose those actions.

Using the same reasoning, American determines how United will respond to each of American's possible actions, as the right side of the figure illustrates. American predicts that

- If American chooses 48, United will sell 64, and American's profit will be \$3.8 million.
- If American chooses 64, United will sell 64, and American's profit will be \$4.1 million.
- If American chooses 96, United will sell 48, and American's profit will be \$4.6 million.

Thus, to maximize its profit, American chooses 96 in the first stage. United's strategy is to make its best response to American's first-stage action: United selects 64 if American chooses 48 or 64, and United picks 48 if American chooses 96. Thus, United responds in the second stage by selecting 48. In this subgame perfect Nash equilibrium, neither firm wants to change its strategy. Given that American Airlines sets its output at 96, United is using a strategy that maximizes its profit, $q_U = 48$, so it doesn't want to change. Similarly, given how United will respond to each possible American output level, American cannot make more profit than if it sells 96.

The subgame perfect Nash equilibrium requires players to believe that their opponents will act optimally—in their own best interests. No player has an incentive to deviate from the equilibrium strategies. The reason for adding the requirement of subgame perfection is that we want to explain what will happen if a player does not follow the equilibrium path. For example, if American does not choose its equilibrium output in the first stage, subgame perfection requires that United will still follow the strategy that maximizes its profit in the second stage conditional on American's actual output choice.

Not all Nash equilibria are subgame perfect Nash equilibria. For example, suppose that American's strategy is to pick 96 in the first stage, and United's strategy is to choose 96 if American selects 48 or 64, and 48 if American chooses 96. The outcome is the same as the subgame perfect Nash equilibrium we just derived because American selects 96, United chooses 48, and neither firm wants to deviate.[10] Thus, this set of strategies is a Nash equilibrium. However, this set of strategies is not a subgame perfect Nash equilibrium. Although this Nash equilibrium has the same equilibrium path as the subgame perfect Nash equilibrium, United's strategy differs from the equilibrium path. If American had selected 48 (or 64), United's strategy would not result in a Nash equilibrium. United would receive a higher profit if it produced 64 rather than the 96 that this strategy requires. Therefore, this Nash equilibrium is not subgame perfect.

This subgame perfect Nash equilibrium, or Stackelberg equilibrium, differs from the simultaneous-move, Nash-Cournot equilibrium. American, the Stackelberg leader, sells 50% more than the Cournot quantity, 64, and earns \$4.6 million, which is 15% more than the Cournot level of profit, \$4.1 million. United, the Stackelberg follower, sells a quantity, 48, and earns a profit, \$2.3 million, both of which are less than the Cournot levels. Thus, although United has more information in the Stackelberg equilibrium than it does in the Cournot model—it knows American's output level—it is worse off than if both firms chose their actions simultaneously.

Credibility

Why do the simultaneous-move and sequential-move games have different outcomes? Given the option to act first, American chooses a large output, 96, so that United's profit-maximizing action is to pick a relatively small output, 48. American benefits from moving first and choosing the Stackelberg leader quantity.

In the simultaneous-move game, why doesn't American announce that it will produce the Stackelberg leader's output to induce United to produce the Stackelberg follower's output level? The answer is that when the firms move simultaneously, United doesn't believe American's warning that it will produce a large quantity, because it is not in American's best interest to produce that large a quantity of output.

credible threat
an announcement that a firm will use a strategy that will harm its rival and that the rival believes because the firm benefits from using that strategy

For a firm's announced strategy to be a **credible threat**, rivals must believe that the firm's strategy is rational in the sense that it is in the firm's best interest to use it.[11] If American produced the leader's level of output and United produced the Cournot level, American's profit would be lower than if it too produced the Cournot level. Because American cannot be sure that United will believe its threat and reduce its output in the

[10]Given United's strategy, American does not have any incentive to deviate. If American chooses 48 it will get \$2.3 million and if it chooses 64 it will get \$2.0 million, both of which are less than the \$4.6 million if it chooses 96. And given American's strategy, no change in United's strategy would raise its profit.

[11]No doubt you've been in a restaurant and listened to an exasperated father trying to control his brat with such extreme threats as "If you don't behave, you'll have to sit in the car while we eat dinner" or "If you don't behave, you'll never see television again." The kid, of course, does not view such threats as credible and continues to terrorize the restaurant—proving that the kid is a better game theorist than the father.


Come here! Don't make me run after you!

simultaneous-move game, American produces the Cournot output level. In contrast, in the sequential-move game, because American moves first, its commitment to produce a large quantity is credible.

The intuition for why commitment makes a threat credible is that of "burning bridges." If the general burns the bridge behind the army so that the troops can only advance and not retreat, the army becomes a more fearsome foe—like a cornered animal.[12] Similarly, by limiting its future options, a firm makes itself stronger.[13]

Not all firms can make credible threats, however, because not all firms can make commitments. Typically, for a threat to succeed, a firm must have an advantage that allows it to harm the other firm before that firm can retaliate. Identical firms that act simultaneously cannot credibly threaten each other. However, a firm may be able to make its threatened behavior believable if firms differ. An important difference is the ability of one firm to act before the other. For example, an incumbent firm could lobby for the passage of a law that forbids further entry.

Dynamic Entry Game

We can illustrate the use of laws as a form of commitment by using the entry game. In some markets, by moving first, a firm can act strategically to prevent potential rivals from entering the market. How can an *incumbent*, monopoly firm deter a (potential) *rival* from entering that market? Does the incumbent gain by deterring entry?

The incumbent may prevent entry if it can make a creditable threat by acting first. However, a manager cannot deter entry merely by telling a potential rival, "Don't enter! This market ain't big enough for the two of us." The potential rival would merely laugh and suggest that the manager's firm exit if it doesn't want to share the market.

Exclusion Contracts We consider an example where the incumbent can pay a third party to prevent entry. A mall has a single shoe store, the incumbent firm. If the incumbent pays the mall's owner b, a clause is added to their rental agreement that guarantees the incumbent the *exclusive right* to be the only shoe store in the mall.[14] If this payment is made, the landlord agrees to rent the remaining space only to a restaurant, a toy store, or some other business that does not sell shoes. Should the shoe store pay?

---

[12]"On hemmed-in ground, I would block any way of retreat. On desperate ground, I would proclaim to my soldiers the hopelessness of saving their lives." Sun Tzu, *On the Art of War*.

[13]Some psychologists use the idea of commitment to treat behavioral problems. A psychologist may advise an author with writer's block to set up an irreversible procedure: If the author's book is not finished by a certain date, the author's check for $10,000 will be sent to the group the author hates most in the world—be it the Nazi Party, the Ku Klux Klan, or the National Save the Skeets Foundation. Such an irreversible commitment helps the author get the project done by raising the cost of failure. (We can imagine the author playing a game against the author's own better self.)

[14]Dallas-Fort Worth, Fort-Lauderdale-Hollywood, San Francisco, and other airports sell the exclusive rights to provide a particular product or service at their airports. For example, San Francisco International Airport put out a call for bids for the exclusive right to provide a bookstore with the minimum acceptable bid of $400,000 for one year.

The game tree in Figure 14.2 shows the two stages of the game involving the incumbent and its potential rival, another shoe store. In the first stage, the incumbent decides whether to pay b to prevent entry. In the second stage, the potential rival decides whether to enter. If it enters, it incurs a fixed fee of F to build its store in the mall.

The right side of the figure shows the incumbent's and the potential rival's profits (π_i, π_r) for each of the three possible outcomes. The outcome at the top of the figure shows that if the incumbent does not buy exclusivity and the potential rival does not enter, the incumbent earns the "monopoly" profit of $\pi_i = 10$ ($10 thousand) per month and its potential rival earns nothing, $\pi_r = 0$. The middle outcome shows that if the incumbent does not pay the exclusivity fee and the potential rival enters, the incumbent earns a duopoly profit of $\pi_i = 4$ and the rival earns the duopoly profit less its fixed cost, F, of entering, $\pi_r = 4 - F$. In the bottom outcome, the incumbent pays b for the exclusivity right so that it earns the monopoly profit less the exclusivity fee, $\pi_i = 10 - b$, and its potential rival earns nothing, $\pi_r = 0$.

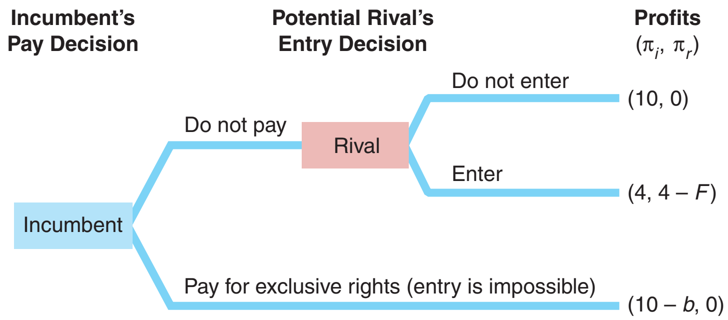
To solve for the subgame perfect Nash equilibrium, we work backward, starting with the last decision, the potential rival's entry decision. The top portion of the game tree shows what happens if the incumbent does not pay the landlord to prevent entry. The potential rival enters if it earns more from entering, $\pi_r = 4 - F$, than if it stays out of the market, $\pi_r = 0$. That is, the potential rival enters if $F \leq 4$. In the bottom portion of the game tree, where the incumbent pays b for an exclusive contract that prevents entry, the potential rival has no possible action.

Which of the three possible outcomes occurs depends on the incumbent's exclusivity fee, b, and the potential rival's fixed cost of entering the market, F:

- **Blockaded entry** ($F > 4$): The potential rival chooses not to enter even if the incumbent does not pay to have an exclusive contract, so $\pi_r = 0$. The incumbent avoids spending b and still earns the monopoly profit, $\pi_i = 10$.
- **Deterred entry** ($F \leq 4$, $b \leq 6$): Because $F \leq 4$, entry will occur unless the incumbent pays the exclusivity fee. The incumbent chooses to pay the exclusivity fee, b, because its profit from doing so, $\pi_i = 10 - b \geq 4$, is at least as large as what it earns if it permits entry and earns the duopoly profit, $\pi_i = 4$. Because the rival does not enter, it earns nothing: $\pi_r = 0$.

Figure 14.2 Game Tree: Whether an Incumbent Pays to Prevent Entry MyLab Economics Video

If the potential rival stays out of the market, it makes no profit, $\pi_r = 0$, and the incumbent firm makes the monopoly profit, $\pi_i = 10$. If the potential rival enters, the incumbent earns the duopoly profit of 4 and the rival makes $4 - F$, where F is its fixed cost of entry. If the duopoly profit, 4, is less than F, entry does not occur. Otherwise, entry occurs unless the incumbent acts to deter entry by paying for exclusive rights to be the only shoe store in the mall. The incumbent pays the landlord only if $10 - b > 4$.



- **Accommodated entry** ($F \leq 4$, $b > 6$): Entry will occur unless the incumbent pays the fee because the rival's fixed costs are less than or equal to 4. The incumbent does not pay for an exclusive contract. The exclusivity fee is so high that the incumbent earns more by allowing entry, $\pi_i = 4$, than it earns if it pays for exclusivity, $\pi_i = 10 - b < 4$. Thus, the incumbent earns the duopoly profit, $\pi_i = 4$, and the rival makes $\pi_r = 4 - F$.

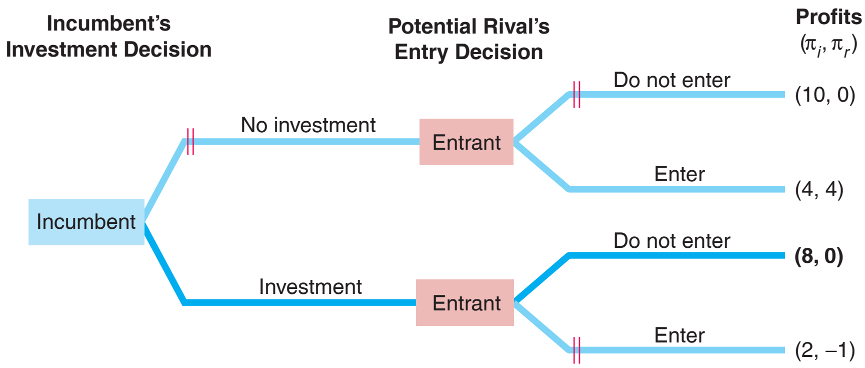
In short, the incumbent does not pay for an exclusive contract if the potential rival's cost of entry is prohibitively high ($F > 4$) or if the cost of the exclusive contract is too high ($b > 6$).

The next Solved Problem uses dynamic game theory to reject the following false belief:

Common Confusion: A firm invests in new equipment that lowers its marginal cost only if the savings from lower production costs more than offsets the cost of the investment.

Solved Problem 14.4

Use the following game tree to demonstrate that an incumbent faced with potential entry may invest in new equipment even when the investment does not lower its cost. In the first stage of the game, the incumbent firm decides whether to invest in new robotic equipment, which lowers its marginal cost of production. In the second stage, a potential rival decides whether to enter the market.



Answer

1. *Determine the potential rival's response in the second stage to each possible action taken by the incumbent in the first stage.* To solve for the subgame perfect Nash equilibrium, we work backward from the potential rival's entry decision in the second stage of the game. If the incumbent does not invest, its rival enters because its profit from entering, $\pi_r = 4$, exceeds its zero profit if it does not enter. If the incumbent does invest, its potential rival stays out of the market because entry would be unprofitable: $\pi_r = -1 < 0$.

2. *Determine the incumbent's decision given its potential rival's responses.* If the incumbent does not invest and the rival enters, the incumbent earns $\pi_i = 4$. If it invests and the potential rival does not enter, the incumbent earns $\pi_i = 8$. Thus, the incumbent invests.

Comment: This investment would not pay without a threat of entry. Investing would cause the incumbent's profit to fall from $\pi_i = 10$ to 8. However, because the investment deters entry, the incumbent benefits from investing.

Application

Keeping Out Casinos

Suppose you own the only casino in town. You know that the larger your casino, the more games and entertainment you can provide, so the more customers your casino will attract. You face a trade-off because as the casino becomes larger, your costs increase. Thus, you must pick the optimal size to maximize your monopoly profit.

A complication arises if, at the monopoly-profit-maximizing size, it is profitable for a new firm to enter. If so, it may pay to expand your casino. You can credibly commit to maintaining a larger casino because such investments are costly to reverse. Given a large enough casino, a potential entrant will lose money by building a competing casino and hence won't enter.

A casino owner learns about entry plans when a potential entrant begins contacting a network of vendors and suppliers. Cookson (2015) estimated that incumbent casinos expand their floor space by 13 to 16% in response to an entry threat. He also found that entry is half as likely (33% versus 66%) if the incumbent invests in a large expansion rather than one that is 20,000 square feet smaller.

limit pricing
a firm sets its price (or its output) so that another firm cannot enter the market profitably

Limit Pricing A firm is **limit pricing** if it sets its price (or, equivalently, its output) so that another firm cannot enter the market profitably. For example, the incumbent could set a price below the potential rival's marginal cost so entry would be unprofitable. Or, the incumbent could produce so much output that the price is very low and too few customers remain for the potential rival to make a profit. However, to limit price successfully, a firm must have an advantage over its rivals, as the following example illustrates.

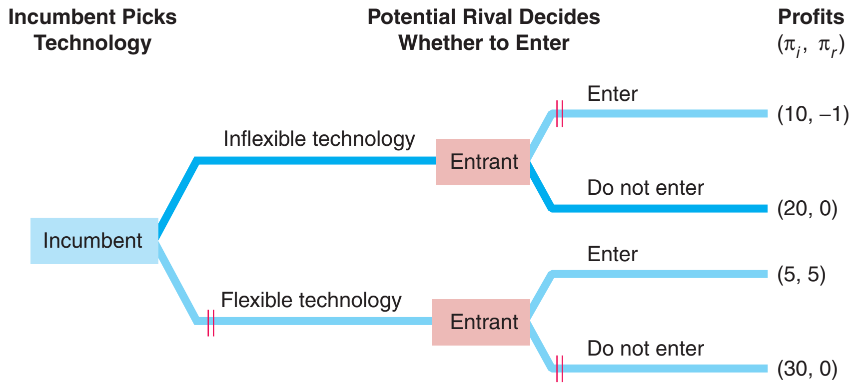
An incumbent firm is making a large, monopoly profit, which attracts the interest of a potential rival. The incumbent could threaten that it will limit price if entry occurs. It could announce that, after entry, it will charge a price so low that the other firm will make a loss. This threat will only work if the threat is credible. It is not credible if the two firms have identical costs and market demand is adequate to support both firms. Once entry occurs, it is in the incumbent's best interest to charge the duopoly price and make a profit rather than charge such a low price that everyone loses money. Realizing that the incumbent won't actually limit price, the potential rival ignores the threat and enters.

For the threat of limit pricing to be credible, the incumbent must have an advantage over its rival. For example, if the incumbent's costs are lower than those of the potential rival, the incumbent can charge a price so low that the rival would lose money while the incumbent earns a higher profit than if it allows entry.

Another example is an extreme form of the Stackelberg oligopoly example. The Stackelberg leader acts first and produces a large quantity so that the follower produces a smaller quantity. Depending on the demand curve and the firms' costs, it may be even more profitable for the leader to produce such a large quantity that the follower cannot earn a profit. That is, the leader makes limit pricing credible by committing to provide a very large output level.

Solved Problem 14.5

In the first stage of a game between an incumbent and a potential rival, the incumbent builds its plant using either an inflexible technology that allows it to produce only a (large) fixed quantity, or a flexible technology that allows it to produce small or large quantities. In the second stage, the potential rival decides whether to enter. With the inflexible technology, the incumbent makes so much output that its threat to limit price is credible, as the following game tree illustrates. What strategy (technology) maximizes the incumbent's profit?



Answer

1. *Work backward by determining the potential rival's best strategy conditional on each possible action by the incumbent*. This game has two proper subgames. The upper right subgame shows the profits if the potential rival enters or if it does not enter given that the incumbent uses the inflexible technology. The potential rival loses money ($\pi_r = -1$) if it enters, but breaks even ($\pi_r = 0$) if it doesn't, so it does not enter. In the lower right subgame, the potential rival decides whether to enter given that the incumbent is using the flexible technology. Here, the potential rival prefers to enter and earn a profit of $\pi_r = 5$ rather than stay out and earn nothing.

2. *Given the responses by the potential rival to each of the incumbent's strategies, determine the incumbent's best strategy*. If the incumbent uses the flexible technology, entry occurs, and the incumbent earns $\pi_i = 5$. However, if the incumbent uses the inflexible technology, the other firm does not enter, and the incumbent's profit is $\pi_i = 20$. Thus, the incumbent chooses the inflexible technology.

Comment: The incumbent would earn even a higher profit with the flexible technology if no entry occurs. However, if the incumbent chooses the flexible technology, its rival will enter, so the incumbent is better off committing to the inflexible technology. The inflexible technology serves as a credible threat that the incumbent will limit price.

14.4 Auctions

To this point, we have examined games in which players have complete information about payoff functions. We now turn to an important game, the *auction*, in which players devise bidding strategies without knowing other players' payoff functions.

auction
a sale in which property or a service is sold to the highest bidder

An **auction** is a sale in which a good or service is sold to the highest bidder. A substantial amount of exchange takes place through auctions. Typically, governments award contracts using procurement auctions. In recent years, governments have auctioned portions of the airwaves for radio stations, mobile phones, and wireless Internet access and have used auctions to set up electricity and transport markets. Other goods commonly sold at auction are natural resources such as timber, as well as houses, cars, agricultural produce, horses, antiques, and art. In this section, we first consider the various types of auctions and then investigate how the rules of the auction influence buyers' strategies.

Elements of Auctions

Before deciding what strategy to use when bidding in an auction, one needs to know the rules of the game. Auctions have three key components: the number of units being sold, the format of the bidding, and the value that potential bidders place on the good.

Number of Units Sellers may use auctions to sell one or many units of a good. In 2004, Google auctioned its initial public offering of many identical shares of stock at one time. Many auctions are for a single good, such as an original painting. For simplicity in this discussion, we concentrate on auctions for a single, indivisible item.

Format Virtually all auctions are variants of the *English auction*, the *Dutch auction*, the *sealed-bid auction*, *or* the *double auction*.

- **English auction.** In the United States and Britain, almost everyone has seen an *English* or *ascending-bid auction*, at least in the movies. The auctioneer starts the bidding at the lowest price that is acceptable to the seller and then repeatedly encourages potential buyers to bid more than the previous highest bidder. The auction ends when no one is willing to bid more than the current highest bid: "Going, going, gone!" The last bidder, who has the highest bid, wins the auction, buying the good. Sotheby's and Christie's use English auctions to sell art and antiques.
- **Dutch auction.** A *Dutch auction* or *descending-bid auction* ends dramatically with the first "bid." The seller starts by asking if anyone wants to buy at a relatively high price. The seller reduces the price by given increments until someone accepts the offered price and buys at that price. Variants of Dutch auctions are often used to sell multiple goods at once, such as in Google's initial public offering auction and the U.S. Treasury's sales of Treasury bills.
- **Sealed-bid auction.** In a *sealed-bid auction*, everyone submits a bid simultaneously without seeing other bids (for example, by submitting each bid in a sealed envelope), and the highest bidder wins. The price the winner pays depends on whether it is a first-price auction or a second-price auction. In a *first-price auction*, the winner pays its own, highest bid. Governments often use this type of auction. In a *second-price auction*, the winner pays the amount bid by the second-highest bidder.

  Many online auction houses use a variant of the second-price auction. For example, you bid on eBay by specifying the maximum amount you are willing to bid. If your maximum is greater than the maximum bid of other participants, eBay's computer places a bid on your behalf that is a small increment above the maximum bid of the second-highest bidder. This system differs from the traditional sealed-bid auction in that people can continue to bid until the official end of the auction, and potential bidders know the current bid price (but not the maximum that the highest bidder is willing to pay). Thus, eBay has some of the characteristics of an English auction.
- **Double auction.** All potential buyers and sellers in a *double auction* may make public offers stating prices at which they are willing buy or sell. They may accept another participant's offer to buy or sell. Traditionally, most financial exchanges in which people trade stocks, options, or other securities were *oral* double auctions. Traders stood in *open pits* and would shout, wave cards in the air, or use hand signals to convey their offers or to signal agreements to trade. In recent years, almost all of these exchanges have switched to *electronic* double-auction systems.

Value Auctioned goods may have a *private value* or a *common value*. Typically, this distinction turns on whether the good is unique.

- **Private value.** If each potential bidder places a different personal value on a good, we say that the good has a *private value*. Individual bidders know how much the good is worth to them but not how much other bidders value it. An archetypical example is an original work of art, which people may value very differently.
- **Common value.** Many auctions involve a good that has the same fundamental value to everyone, but no buyer knows exactly what that *common value* is. For example, in a timber auction, firms bid on all the trees in a given area. All firms know what the current price of lumber is; however, they do not know exactly how many board feet of lumber are contained in the trees.

In many actual auctions, goods have both private value and common value. For example, in the tree auction, bidding firms may differ not only in their estimates of the amount of lumber in the trees (common value), but also in their costs of harvesting (private value).

Bidding Strategies in Private-Value Auctions

A potential buyer's optimal strategy depends on the number of units, the format, and the type of values in an auction. For specificity, we examine auctions in which each bidder places a different private value on a single, indivisible good.

Second-Price Auction Strategies According to eBay, if you choose to bid on an item in its second-price auction, you should "enter the maximum amount you are willing to pay for the item" (**pages.ebay.com/education/gettingstarted/bidding.html**). Is eBay's advice correct?

In a traditional sealed-bid, second-price auction, bidding your highest value *weakly dominates* all other bidding strategies: The strategy of bidding your maximum value leaves you *as well off* as, or *better off* than, bidding any other value. The amount that you bid affects whether you win, but it does not affect how much you pay if you win, which equals the second-highest bid.

Suppose that you value a folk art carving at \$100. If the highest amount that any other participant is willing to bid is \$85 and you place a bid greater than \$85, you will buy the carving for \$85 and receive \$15 (= \$100 − \$85) of consumer surplus. Other bidders pay nothing and gain no consumer surplus.

Should you ever bid more than your value? Suppose that you bid \$120. You face three possibilities. First, if the highest bid of your rivals is greater than \$120, then you do not buy the good and receive no consumer surplus. This outcome is the same as what you would have received if you had bid \$100, so bidding higher than \$100 does not benefit you.

Second, if the highest alternative bid is less than \$100, then you win and receive the same consumer surplus that you would have received had you bid \$100. Again, bidding higher does not affect the outcome.

Third, if the highest bid by a rival were an amount between \$100 and \$120—say, \$110—then bidding more than your maximum value causes you to win, but you purchase the good for more than you value it, so you receive negative consumer surplus: −\$10 (= \$100 − \$110). In contrast, if you had bid your maximum value, you would not have won, and your consumer surplus would have been zero—which is better than losing \$10. Thus, bidding more than your maximum value can never make you better off than bidding your maximum value and you may suffer.

Should you ever bid less than your maximum value, say, \$90? No, because you only lower the odds of winning without affecting the price that you pay if you do win. If the highest alternative bid is less than \$90 or greater than your value, you receive the same consumer surplus by bidding \$90 as you would by bidding \$100. However, if the highest alternative bid lies between \$90 and \$100, you will lose the auction and give up positive consumer surplus by underbidding.

Thus, you do as well or better by bidding your value than by over- or underbidding. This argument does not turn on whether or not you know other bidders' valuation. If you know your own value but not other bidders' values, bidding your value is your best strategy. If everyone follows this strategy, the person who places the highest value on the good will win and will pay the second-highest value.

English Auction Strategy Suppose instead that the seller uses an English auction to sell the carving to bidders with various private values. Your best strategy is to raise the current highest bid as long as your bid is less than the value you place on the good, \$100. If the current bid is \$85, you should increase your bid by the smallest permitted amount, say, \$86, which is less than your value. If no one raises the bid further, you win and receive a positive surplus of \$14. By the same reasoning, it always pays to increase your bid up to \$100, where you receive zero surplus if you win.

However, it never pays to bid more than \$100. The best outcome that you can hope for is to lose and receive zero surplus. Were you to win, you would have negative surplus.

If all participants bid up to their value, the winner will pay slightly more than the value of the second-highest bidder. Thus, the outcome is essentially the same as in the sealed-bid, second-price auction.

Equivalence of Auction Outcomes For Dutch or first-price, sealed-bid auctions, one can show that participants will shave their bids to less than their value. The basic intuition is that you do not know the values of the other bidders. Reducing your bid decreases the probability that you win but increases your consumer surplus if you win. Your optimal bid, which balances these two effects, is lower than your actual value. Your bid depends on your beliefs about the strategies of your rivals. The best strategy is to bid an amount that is equal to or slightly greater than what you expect will be the second-highest bid, given that your value is the highest.

Thus, the expected outcome is the same under each format for private-value auctions: The winner is the person with the highest value, and the winner pays roughly the second-highest value. According to the Revenue Equivalence Theorem (Klemperer, 2004), under certain plausible conditions we would expect the same revenue from any auction in which the winner is the person who places the highest value on the good.

Winner's Curse

winner's curse
auction winner's bid exceeds an item's common value

A phenomenon occurs in common-value auctions that does not occur in private-value auctions. The **winner's curse** is that the auction winner's bid exceeds the item's common value. The overbidding occurs when bidders are uncertain about the true value of the good.[15]

---

[15]Mike Shor has a clever website, **www.gametheory.net/mike/applets/winnercurse**, which demonstrates the winner's curse. The site asks you, "How much should you offer for a company of uncertain valuation?" You can try various bidding strategies to see which works best.

When the government auctions off timber on a plot of land, potential bidders may differ in their estimates of how many board feet of lumber are available on that land. The higher one's estimate, the more likely that one will make the winning bid. If the average bid is accurate, then the high bid is probably excessive. Thus, the winner's curse is paying too much.

Rational bidders adjust their bids to avoid the winner's curse. Each bidder reasons that, "I can reduce the likelihood of falling prey to the winner's curse by *shading* or reducing my bid below my estimate. I know that if I win, I am probably overestimating the value of the good. The amount by which I should shade my bid depends on the number of other bidders, because the more people who bid, the more likely it is that the winning bid is an overestimate."

Because intelligent bidders shade their bids, sellers can do better with an English auction than with a sealed-bid auction.[16] In an English auction, bidders revise their views about the object's value as they watch others bid.

Because intelligent bidders shade their bids, sellers of common-value goods can do better with an English auction than with a sealed-bid auction. In an English auction, bidders may revise their views about the object's value as they watch others bid.

Many sellers have learned this lesson. For example, online auction sites such as eBay do not use sealed bid auctions, choosing modified English auctions or other types of auctions instead. This lesson is one that many governments have not learned. They continue to rely heavily on sealed-bid auctions even though it is likely they would earn more money if they used an English auction to sell lumber, rights to airwaves, and other property.

Application

Bidder's Curse

What's the maximum you would bid for an item that you know that you can buy for a fixed price of p? No matter how much you value the good, it doesn't make sense to bid more than p. Yet, people commonly do that on eBay. Lee and Malmendier (2011) call bidding more than what should be one's valuation—here, the fixed price—*bidder's curse*.

They examined eBay auctions of a board game, Cashflow 101, a game that is supposed to help people better understand their finances. A search on eBay for Cashflow 101 not only listed the auctions but also the availability of the game for a fixed price. During the period studied, the game was continuously available for a fixed price on the eBay site (by sellers with identical or better reputations and with lower shipping costs).

Even if only a few buyers overbid, their bids may affect the auction price and who wins. The auction price exceeded the fixed price in 42% of the auctions. The average overpayment was 10% of the fixed price. Only a small number of bidders, 17%, bid above the fixed price. However, people who bid too much are disproportionately likely to win the auction and, hence, determine the winning price.

One possible behavioral economics explanation is that bidders paid limited attention to the fixed-price option. Lee and Malmendier found that overbidding was less likely the closer the fixed price appeared on the same screen to the auction and hence the more likely that bidders would notice the fixed-price listing.

Another explanation is lack of bidding experience. Garratt et al. (2012) and Feng et al. (2016) found that inexperienced bidders were more likely to overbid than were experienced bidders.

[16]Although rational bidders should avoid the winner's curse, economists have observed the winner's curse in many situations. An important example is the takeover market: the market for corporate acquisitions. See Thaler (1994) for many examples of the winner's curse, including a discussion of corporate takeovers.

14.5 Behavioral Game Theory

We normally assume that people are rational in the sense that they optimize using all available information. However, they may be subject to psychological biases and may have limited powers of calculation that cause them to act irrationally, as described in the Application "Bidder's Curse." Such possibilities are the domain of behavioral economics (Chapters 4 and 11), which seeks to augment the rational economic model to better understand and predict economic decision making.

Another example of nonoptimal strategies occurs in *ultimatum games*. People often face an *ultimatum*, where one person (the *proposer*) makes a "take it or leave it" offer to another (the *responder*). No matter how long the parties have negotiated, once the proposer issues an ultimatum, the responder has to accept or reject the offer with no opportunity to make a counter-offer. The ultimatum game is a sequential game in which the proposer moves first and the responder moves second (see Camerer, 2003).

Application

GM's Ultimatum

In 2009, General Motors (GM), facing bankruptcy, planned to shut down about one-fourth of its dealerships in the United States and one-third in Canada. Because GM was concerned that dealer opposition could cause delays and impose other costs, it offered dealers slated for termination an ultimatum. They would receive a (small) payment from GM if they did not oppose the restructuring plan.

Dealers could accept the ultimatum and get something, or they could reject the offer, oppose the reorganization, and receive nothing. Although it was irrational, some dealers rejected the ultimatum and loudly complained that GM was "high-handed, oppressive, and patently unfair." In 2011, some terminated Canadian dealerships filed a class-action suit against GM of Canada. They lost their case in 2015.

An Experiment

The possibility that someone might turn down an offer even at some personal cost is important in business and personal negotiations. To gain insight into real decisions, Camerer (2004) conducted an ultimatum experiment.

A group of student participants met in a computer lab. Each person was either a proposer or a responder. A computer anonymously matched each proposer with a responder. In this game, the two people divided \$10. Each proposer made an ultimatum offer to the responder of a particular amount. A responder who accepted received the amount offered and the proposer got the rest of the \$10. If the responder rejected the offer, both players received nothing.

To find the rational, subgame perfect solution, we use backward induction. In the second stage, the responder should accept if the offer is positive. Thus in the first stage, the proposer should offer the lowest possible positive amount.

However, such rational behavior is not a good predictor of actual outcomes. Most offers are between \$3 and \$4—far more than the "rational" minimum offer. Offers of less than \$2 were rare, and responders turned them down about half the time.

One concern about such experiments is that the payoffs are small enough that not all participants take the game seriously. However, when the total amount to be divided was increased to \$100, the results were essentially unchanged: The typical offer remained between 30% and 40% of the total. If anything, responders are even more likely to turn down lowball offers when the stakes were higher.

Reciprocity

Some responders who reject lowball offers feel the proposer is being greedy and would prefer to make a small sacrifice rather than reward such behavior. Some responders are angered by low offers, some feel insulted, and some feel that they should oppose "unfair" behavior. Most proposers anticipate such feelings and offer a significant amount to the responder, but almost always less than 50%.

Apparently, most people accept that the advantage of moving first should provide some extra benefit to proposers, but not too much. Moreover, they believe in *reciprocity*. If others treat us well, we want to return the favor. If they treat us badly, we want to "get even" and will retaliate if the cost does not seem excessive. Thus, if a proposer makes a low offer, many responders are willing to give up something to punish the proposer, using "an eye for an eye" philosophy.

Eckel and Grossman (1996) found that men are more likely than women to punish if the personal cost is high in an ultimatum game. They speculate that this difference may explain gender patterns in wages and unemployment during downturns, where men are more likely to insist on a given wage than are women, who are more flexible.

Challenge Solution

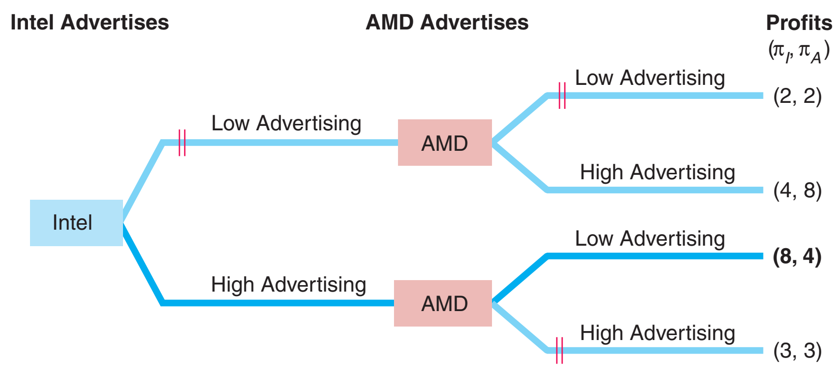
Intel and AMD's Advertising Strategies

MyLab Economics Solved Problem

As we've seen, when one firm in a market acts before another, the first mover may gain an advantage large enough to discourage the second firm from entering the market. In a less extreme case, the original firm may gain a smaller advantage so that the second firm enters, but it produces less than the original firm (as in the airlines' Stackelberg model). We can use this insight to provide a possible explanation for the Challenge: In the market for CPUs for personal computers, why does Intel advertise substantially while AMD does not?

In Solved Problem 14.1, we examined a game where Intel and AMD act simultaneously and have symmetric profits. That game has two pure strategy equilibria. In each, one firm advertises at a low level while the other firm sets a high level. Exercise 6.2 asks you to show this game also has a mixed strategy equilibrium in which each firm sets its advertising low with probability $\frac{1}{7}$ (and has an expected profit of about 3.71).

In contrast, the game has a clear outcome given that Intel acted first, as actually happened. The game tree shows that Intel decides on how much to advertise before AMD can act. AMD then decides how much to advertise. We solve for the subgame-perfect, Nash equilibrium by working backward. For the profits in this game, if Intel were to have a minimal advertising campaign, AMD makes more if it advertises a lot ($\pi_A = 8$) than if it too has a low level of advertising ($\pi_A = 2$). If Intel advertises heavily, AMD makes more with a low-level advertising campaign ($\pi_A = 4$) than with a high-level campaign ($\pi_A = 3$). Given how it expects AMD to behave, Intel intensively advertises because doing so produces a higher profit ($\pi_I = 8$) than does the lower level of advertising ($\pi_I = 4$).



Thus, because Intel acts first and can commit to advertising aggressively, it can place AMD in a position where it makes more with a low-key advertising campaign. Of course, the results might vary if the profits in the game tree differ, but this example provides a plausible explanation for why the firms use different strategies.

Summary

Economists use game theory to analyze conflict and cooperation among players (such as firms). Each player adopts a strategy or battle plan to compete with other firms. Economists typically assume that players have *common knowledge* about the rules of the game, the payoff functions, and other players' knowledge about these issues. In many games, players have *complete information* about how payoffs depend on the strategies of all players. In some games, players have *perfect information* about players' previous moves.

1. **Static Games.** In a static game, such as in the Cournot model or the prisoners' dilemma game, players each make one move simultaneously. Economists use a normal-form representation or payoff matrix to analyze a static game. Typically, economists study static games in which players have complete information about the payoff function—the payoff to any player conditional on the actions all players take—but imperfect information about how their rivals behave because they act simultaneously. The set of players' strategies is a Nash equilibrium if, given that all other players use these strategies, no player can obtain a higher payoff by choosing a different strategy. A given game may have multiple Nash equilibria including pure-strategy and mixed-strategy Nash equilibria. A Nash equilibrium in static games does not necessarily maximize the joint payoffs of all the players.
2. **Repeated Dynamic Games.** In some dynamic games, a static game is repeated, such as when firms make price or quantity decisions every quarter. Therefore, a firm may use a strategy in which it makes a particular move contingent on its rival's actions in previous periods. By using contingent strategies, such as a tit-for-tat strategy or another trigger strategy, it is often easier for firms to maximize their joint payoff—achieve a collusive solution—in a repeated game than in a single-period game.
3. **Sequential Dynamic Games.** In other dynamic games, firms move sequentially, with one player acting before another. By moving first, a firm is able to make a *commitment* or *credible threat*. Consequently, the first mover may receive a higher profit than if the firms act simultaneously. For example, in the Stackelberg oligopoly model, one firm is a *leader* in a sequential game and therefore chooses its output level before rival firms (followers) choose theirs. Applying backward induction, the leader anticipates a follower's reaction and chooses its best output accordingly in the first stage. This first-stage output is a commitment that allows the leader to gain a first-mover advantage. The leader produces more output and earns higher profits than does a follower firm with the same costs.
4. **Auctions.** Auctions are games of incomplete information because bidders do not know the valuation others place on a good. Buyers' optimal strategies depend on the characteristics of an auction. Under fairly general conditions, if the auction rules result in a win by the person placing the highest value on a good that various bidders value differently, the expected price is the same in all auctions. For example, the expected price in various types of private-value auctions is the value of the good to the person who values it second highest. In auctions where everyone values the good the same, though they may differ in their estimates of that value, the successful bidder may suffer from the winner's curse—paying too much—unless bidders shade their bids to compensate for their overoptimistic estimation of the good's value.
5. **Behavioral Game Theory.** People may not use rational strategies because of psychological bias, lack of reasoning ability, or their belief that other managers will not use rational strategies. The ultimatum game illustrates that people commonly use irrational strategies in certain circumstances.

Questions

*Select questions are available on MyLab Economics; * = answer appears at the back of this book;* **A** *= algebra problem;* **C** *= calculus problem.*

1. Static Games

*1.1 Show the payoff matrix and explain the reasoning in the prisoners' dilemma example where Larry and Duncan, possible criminals, will get one year in prison if neither talks; if one talks, one goes free and the other gets five years; and if both talk, both get two years. (*Note*: The payoffs are negative because they represent years in jail, which is a bad.)

1.2 Two large firms, one domestic and one foreign, are competing in the world market for aluminum. If either obtains a subsidy from its respective national government, it can shift profits away from its competitor and increase national income. The firms have the following payoff matrix. Does either firm have a dominant strategy? What is the Nash equilibrium?

		Domestic	
		Do not lobby	*Lobby*
Foreign	*Do not lobby*	0, 0	–1, 4
	Lobby	4, –1	2, 2

1.3 Two manufacturers of toothpaste are debating whether or not to increase their advertising campaigns to promote the special attributes of their products. They have the following payoff matrix. Does either firm have a dominant strategy? What is the Nash equilibrium in this game? Explain.

		Firm 1	
		No change in advertising	*Advertise more*
Firm 2	*No change in advertising*	3, 1	4, 4
	Advertise more	5, 3	6, 2

*1.4 Divit and Sara are each considering starting a new small business selling textiles in the same less-developed region in India. If they have the following profit matrix, does either person enter the market? Does your answer change if the state government provides a subsidy of ₹250,000 rupees to entice new business ventures in the region?

		Divit	
		Start	*Do not start*
Sara	*Start*	–200, 750	500, 0
	Do not start	0, 1,000	0, 0

1.5 Two track stars—the 100-meter gold medalist and the 200-meter gold medalist—agree to a 150-meter duel. Before the race, each athlete decides whether to improve his performance by taking anabolic steroids. If one athlete takes steroids and the other doesn't, the athlete who uses steroids will win. Each athlete's utility payoff is 20 from winning the race, 10 from tying, and 0 from losing. Each athlete's utility falls by 6 if she takes steroids. Model this scenario as a game in which the players simultaneously decide whether to take steroids.

a. What is the Nash equilibrium? Is the game a prisoners' dilemma? Explain.

b. Suppose that one athlete's utility of taking steroids is -12, while the other's remains -6. What is the Nash equilibrium? Is the game a prisoners' dilemma?

1.6 Suppose Procter & Gamble (PG) and Johnson & Johnson (JNJ) are simultaneously considering new advertising campaigns. Each firm may choose a high, medium, or low level of advertising. What are each firm's best responses to its rival's strategies? Does either firm have a dominant strategy? What is the Nash equilibrium in this game?

		PG		
		High	Medium	Low
JNJ	High	1, 1	3, 2	5, 3
	Medium	2, 3	4, 4	6, 5
	Low	3, 5	5, 6	7, 5

1.7 Two firms supply a market. If they have the following profit matrix for low, medium, and high levels of output, identify each firm's best response to its rival's actions. What is the Nash equilibrium?

		Firm 1		
		Low	Medium	High
Firm 2	Low	1, 1	2, 2	5, 3
	Medium	3, 3	4, 4	3, 8
	High	2, 6	5, 6	4, 5

1.8 Modify the previous question so that if Firm 1 chooses High and Firm 2 chooses Low (the upper-right corner), Firm 1 receives 1 rather than 3. How does that change your answer?

1.9 Suppose two firms sell a homogeneous product at constant marginal cost, $MC = 2$, in a market. The inverse market demand curve is $p = 50 - 2Q$, where $Q = q_1 + q_2$. What is each firm's market share in a Nash-Cournot equilibrium? If advertising by either or both firms increases the market demand to $p = 62 - 2Q$, is the market share of either firm (q_1/Q, q_2/Q) affected? **C**

1.10 Takashi Hashiyama, president of the Japanese electronics firm Maspro Denkoh Corporation, was torn between having Christie's or Sotheby's auction the company's $20 million art collection, which included a van Gogh, a Cézanne, and an early Picasso (Carol Vogel, "Rock, Paper, Payoff," *New York Times,* April 29, 2005, A1, A24). He resolved the issue by having the two auction houses' representatives compete in the playground game of rock-paper-scissors. A rock (fist) breaks scissors (two fingers sticking out), scissors cut paper (flat hand), and paper smothers rock. At stake were several million dollars in commissions. Christie's won: scissors beat paper.

a. Show the profit or payoff matrix for this rock-paper-scissors game where the payoff is -1 if you lose, 0 if you tie, and 1 if you win.

b. Sotheby's expert in impressionist and modern art said, "[T]his is a game of chance, so we didn't really give it much thought. We had no strategy in mind." In contrast, the president of Christie's in Japan researched the psychology of the game and consulted with the 11-year-old twin daughters of the director of their Impressionist & Modern Art department. One of these girls said, "Everybody knows you always start with scissors. Rock is way too obvious, and scissors beats paper." The other opined, "Since they were beginners, scissors was definitely the safest." Evaluate these comments on strategy. What strategy would you recommend if you knew that your rival was consulting with 11-year-old girls? In general, what pure or mixed strategy would you have recommended, and why?

1.11 You and a good friend have just finished having a drink together at a local bar and are both insisting on paying the bill. The two of you agree to each flip a coin to resolve this "dispute." You pay (win) if both coins show either heads or tails; your friend pays (wins) if one coin shows heads and the other shows tails. What is the payoff matrix for this game? Is there a pure-strategy Nash equilibrium? What is the mixed-strategy Nash equilibrium? **A**

*1.12 Two Internet providers supply a local market. If they have the following profit matrix for charging low and high prices, does either firm have a dominant strategy? Is there a pure-strategy Nash equilibrium? What is the mixed-strategy Nash equilibrium? (*Hint*: See Solved Problems 14.1 and 14.2.) **A**

		Firm 2	
		Low Price	High Price
Firm 1	Low Price	49, 47	48, 52
	High Price	47, 57	52, 55

1.13 Two couples love eating at restaurants, particularly when they can go together. However, one couple prefers organic food while the other would rather eat non-organic food. If they go together, the payoff is higher for the couple who favors the type of food more (payoff = 4), and less for the other couple (payoff = 3). If they dine out alone, the payoff is smaller and the same for each (payoff = 2). Show the payoff matrix for these two couples and determine the Nash equilibria. How do the payoffs from the Nash equilibria compare? (*Hint*: See Solved Problems 14.1 and 14.2.) **A**

1.14 Suppose that the two couples in Question 1.13 only like to eat at restaurants when they can go together so that the payoff is negative and the same for each (payoff = −1) if they dine out alone. How does the equilibrium change? (*Hint*: See Solved Problems 14.1 and 14.2.) **A**

1.15 Lori employs Max. She wants him to work hard rather than to loaf. She considers offering him a bonus or not giving him one. All else the same, Max prefers to loaf.

		Max	
		Work	*Loaf*
Lori	*Bonus*	2 1	3 −1
	No Bonus	−1 3	0 0

If they choose actions simultaneously, what are their strategies? Why does this game have a different type of equilibrium than the game in Solved Problem 14.2? **A**

1.16 Show that advertising is a dominant strategy for both firms in both panels of Table 14.4. Explain why that set of strategies is a Nash equilibrium.

1.17 In the Application "Strategic Advertising," would the cola advertising or cigarette advertising game be an example of a prisoners' dilemma game?

1.18 In the *battle of the sexes* game, the husband likes to go to the mountains on vacation, and the wife prefers the ocean, but they both prefer to take their vacations together.

		Husband	
		Mountains	*Beach*
Wife	*Mountains*	1 2	−1 −1
	Beach	−1 −1	2 1

What are the Nash equilibria? Discuss whether this game and equilibrium concept make sense for analyzing a couple's decisions. How might you change the game's rules so that it makes more sense? **A**

2. Repeated Dynamic Games

2.1 In a repeated game, how does the outcome differ if firms know that the game will be (a) repeated indefinitely, (b) repeated a known, finite number of times, and (c) repeated a finite number of times but the firms are always unsure whether the current period will be the last? (*Hint*: See Solved Problem 14.3.)

*2.2 In the repeated-game airline example in Solved Problem 14.3, what happens if the players know the game will last only five periods? What happens if the game is played forever but the managers of one or both firms care only about current profit?

3. Sequential Dynamic Games

*3.1 Two firms supply a market. They have the following profit matrix for producing 10 or 20 units of output.

		Firm 2	
		10	20
Firm 1	10	52 54	64 58
	20	56 68	24 28

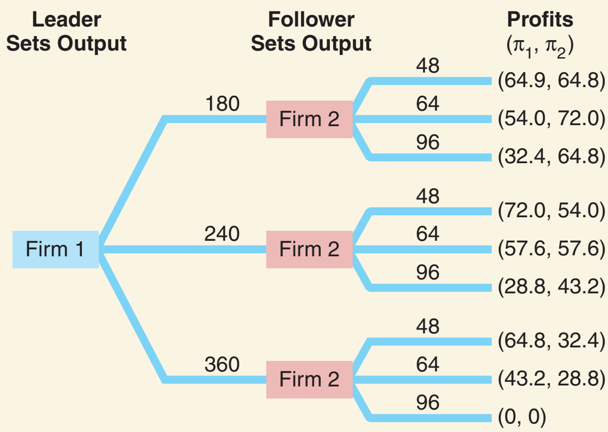
a. What is the Nash equilibrium if both firms make their decisions simultaneously?

b. Draw the game tree if Firm 1 can decide first. What is the outcome? Why?

c. Draw the game tree if Firm 2 can decide first. What is the outcome? Why?

3.2 Does your answer to Question 3.1 change if the government imposes a lump-sum franchise tax of 54 on each firm (that is, the payoffs in the matrix are all reduced by 54)? What would happen if the lump-sum tax were 65 instead?

3.3 Solve for the Stackelberg subgame perfect Nash equilibrium for the following game tree. What is the joint-profit maximizing outcome? Why is that not the outcome of this game?



3.4 In Solved Problem 14.2, suppose that Mimi can move first. What are the equilibria, and why (use an extensive-form diagram)? Now repeat your analysis if Jeff can move first.

3.5 Suppose that Question 1.4 were modified so that the state government does not provide an incentive for new business startups, but Sara has a head start over Divit and can move first. What is the Nash equilibrium? Explain.

3.6 Levi Strauss and Wrangler are planning new-generation jeans and must decide on the colors for their products. The possible colors are white, black, and violet. The payoff to each firm depends on the color it chooses and the color chosen by its rival, as the profit matrix shows:

a. Given that the firms move simultaneously, identify any dominant strategies in this game and find any Nash equilibria.

		Levi Strauss		
		White	*Black*	*Violet*
Wrangler	*White*	10 / 10	30 / 20	40 / 30
	Black	20 / 30	0 / 0	35 / 15
	Violet	15 / 40	20 / 35	0 / 0

b. Now suppose the firms move sequentially, with Wrangler moving first. Draw a game tree and identify any subgame perfect Nash equilibria in this sequential-move game.

3.7 Two firms are considering which of two markets to enter. The markets are located in different cities. They have the following profit matrix:

		Firm 2	
		Market 1	*Market 2*
Firm 1	*Market 1*	11 / 10	13 / 13
	Market 2	12 / 15	10 / 9

If both firms move simultaneously, does either firm have a dominant strategy? Is there a Nash equilibrium? How does your answer change if the local government for Market 1 provides a subsidy of 3 for firms that set up a new business in the city?

3.8 To increase its market share, one firm in a duopoly has been charging a price less than its competitor. This has reduced the other firm's profits, and that second firm is now threatening the first with a price war. If the first firm does not raise its price, the second one will lower its price to undercut the price presently being charged by the first firm. Should the first firm consider the threat credible and raise its price? The table below shows the profits each firm would make on the various possible outcomes.

a. Draw the game tree. Who moves first?

b. What is the equilibrium?

c. Should Firm 1 believe Firm 2's threat is credible?

d. What action does Firm 1 take?

	Firm 1		Firm 2
	Profit	*Market Share*	*Profit*
Raise price, no retaliatory price cut	90	50%	90
Raise price, retaliatory price cut	77	43%	86
Do not raise price, no retaliatory price cut	88	55%	81
Do not raise price, retaliatory price cut	76	48%	79

3.9 In 2007, Italy announced that an Italian journalist, Daniel Mastrogiacomo, who had been held hostage for 15 days by the Taliban in Afghanistan, had been ransomed for 5 Taliban prisoners. Governments in many nations denounced the act as a bad idea because it rewarded terrorism and encouraged more abductions. Consequently, the Afghanistan government announced that it would no longer make such trades ("Afghanistan: Government Pledges End to Hostage Deals," Radio Free Europe, April 16, 2007). Use an extensive-form game tree to analyze the basic arguments. Can you draw any hard-and-fast conclusions about whether the Italians' actions were a good or bad idea? (*Hint*: Does your answer depend on the relative weight one puts on future costs and benefits relative to those today?)

3.10 Firm 1 has a first-mover advantage from research that resulted in a new cost-saving technology and lower costs of production than potential rivals. This monopolist could even sell at a price so low that other potential competitors with higher costs would lose money and not enter the market. If a potential entrant expects the incumbent to produce a large quantity, it does not enter. The table below shows the profits each firm would make on the various possible outcomes. Draw a game tree to determine if the incumbent might produce even more than the monopoly quantity in the first period. What is the equilibrium? (*Hint*: See Solved Problems 14.4 and 14.5.)

	Incumbent	Rival
Monopoly output, no entry	300	0
Monopoly output, entry	169	50
Large output, no entry	232	0
Large output, entry	51	−1

*3.11 A monopoly manufacturing plant currently uses many workers to pack its product into boxes. It can replace these workers with an expensive set of robotic arms. Although the robotic arms raise the monopoly's fixed cost substantially, they lower its marginal cost because it no longer has to hire as many workers. Buying the robotic arms raises its total cost: The monopoly can't sell enough boxes to make the machine pay for itself, given the market demand curve. Suppose the incumbent does not invest. If its rival does not enter, it earns \$0 and the incumbent earns \$900. If the rival enters, it earns \$300 and the incumbent earns \$400. Alternatively, the incumbent invests. If the rival does not enter, it earns \$0 and the incumbent earns \$500. If the rival enters, the rival loses \$36 and the incumbent makes \$132. Show the game tree. Should the monopoly buy the machine anyway? (*Hint*: See Solved Problem 14.4.)

*3.12 Suppose that an incumbent can commit to producing a large quantity of output before the potential entrant decides whether to enter. The incumbent chooses whether to commit to produce a small quantity, q_i, or a large quantity. The rival then decides whether to enter. If the incumbent commits to the small output level and if the rival does not enter, the rival makes \$0 and the incumbent makes \$900. If it does enter, the rival makes \$125 and the incumbent earns \$450. If the incumbent commits to producing the large quantity, and the potential entrant stays out of the market, the potential entrant makes \$0 and the incumbent makes \$800. If the rival enters, the best the entrant can make is \$0, the same amount it would earn if it didn't enter, but the incumbent earns only \$400. Show the game tree. What is the subgame perfect Nash equilibrium? (*Hint*: See Solved Problems 14.4 and 14.5.)

*3.13 Before entry, the incumbent earns a monopoly profit of $\pi_m = \$10$ (million). If entry occurs, the incumbent and entrant each earn the duopoly profit, $\pi_d = \$3$. Suppose that the incumbent can induce the government to require all firms to install pollution-control devices that cost each firm \$4. Show the game tree. Should the incumbent urge the government to require pollution-control devices? Why or why not?

3.14 A gas station at a rest stop along the highway can pay the owner of the rest stop \$40,000 to prevent a second station from opening. Without entry, the incumbent gas station's profit is $\pi_i = \$100{,}000$. With entry, its duopoly profit would be \$45,000 and the entrant would earn a profit of \$30,000. Will the incumbent pay for exclusivity? Will entry occur? Use a game-tree diagram to answer these questions.

4. Auctions

4.1 Charity events often use silent auctions. A donated item, such as a date with a movie star (Colin Firth and Scarlett Johansson in 2008) or a former president (Bill Clinton in 2013), is put up for bid. In a silent auction, bidders write down bids and submit them. Some silent auctions use secret bids submitted in sealed envelopes, which are kept confidential. Other silent auctions are open: the bidder writes down a bid on a bulletin board that everyone present can see. Which kind of auction would you expect to raise more revenue for the charity?

4.2 At the end of performances of his Broadway play "Cyrano de Bergerac," Kevin Kline, who starred as Cyrano, the cavalier poet with a huge nose, auctioned his prosthetic proboscis, which he and his co-star, Jennifer Garner, autographed (Dan Mitchell, "This Time, Santa Has Been Too Naughty," *New York Times*, December 9, 2007) to benefit Broadway Cares in its fight against AIDS. An English auction was used. One night, a television producer grabbed the nose for \$1,400, while the next night it fetched \$1,600. On other nights, it sold for \$3,000 and \$900. Why did the value fluctuate substantially from night to night? Which bidder's bid determined the sales price? How did the audience's knowledge that the proceeds would go to charity affect the auction price?

4.3 Suppose that Firm 1, Firm 2, and Firm 3 are the only three firms interested in the lot at the corner of First Street and Glendon Way. The seller auctions the lot using a second-price sealed-bid auction. Suppose Firm 1 values the lot at $v_1 = \$20{,}000$, Firm 2 at \$18,500, and Firm 3 at \$16,800. Each bidding firm's surplus is $v_i - p$ if it wins the auction and 0 if it loses. The values are private. What is each bidder's optimal bid? Which firm wins the auction, and what price does that firm pay?

5. Behavioral Game Theory

5.1 Draw a game tree that represents the ultimatum game in which the proposer is a first mover who decides how much to offer a responder and the responder then decides to accept or reject the offer. The total amount available is \$50 if the responder accepts the offer, but both players get nothing if the responder rejects the offer. Offers must be in whole dollars. What is the subgame perfect Nash equilibrium? What would you expect to happen in practice?

5.2 People sometimes turn down an offer even at some personal cost to themselves. Suppose, for example, that a firm decides to close a money-losing plant. To avoid additional costs that opposition to the closure might entail, the firm offers each employee a small payment in addition to the legally required minimum of one week of severance pay for each year of employment. If the employees do not accept the offer, they receive only the legal minimum. Why might the employees vote to reject the ultimatum?

6. Challenge

6.1 In the game between Intel and AMD in the Challenge Solution, suppose that each firm earns a profit of 9 if both firms advertise. What is the new subgame perfect Nash equilibrium outcome? Show in a game tree.

6.2 Derive the mixed strategy equilibrium if both Intel and AMD act simultaneously in the game in the Challenge Solution. What is the expected profit of each firm? (*Hint*: see Solved Problems 14.1 and 14.2 and the Challenge Solution.) **A**

6.3 Most major electric car manufacturers belong to two rival camps. Each group uses one of two incompatible technologies to charge their cars at recharging stations (similar to gas stations). Both technologies use direct current to charge car batteries to 80% of capacity in less than 20 minutes. The Japanese auto manufacturers (Honda, Mazda, Mitsubishi, Nissan, Subaru, and Toyota) and PSA Peugeot Citroen back the CHAdeMO technology. Most German and U.S. automakers (Audi, BMW, Chrysler, Daimler, Ford, General Motors, Porsche, and Volkswagen) back the Combo technology. The CHAdeMO technology was introduced first and is in much wider use (Julia Pyper, "Charger Standards Fight Confuses Electric Vehicle Buyers, Puts Car Company Investments at Risk," **www.eenews.net,** July 24, 2013). Relabel the extensive-form diagram in the Challenge Solution (keeping the current payoffs) to illustrate the decision making of these firms. Now, change your analysis. Some industry analysts believed that these firms introduced the Combo standard to slow sales of the Nissan Leaf so that rival manufacturers could catch up. Change the payoffs to illustrate why the German-American group might choose that standard even though the Japanese-Peugeot group acted first.

15 Factor Markets

The laborer is worthy of his reward. —Timothy 5:18

Challenge

Athletes' Salaries and Ticket Prices

In what baseball team owners think fondly of as the "good old days," teams successfully colluded to keep athletes' salaries low. The "reserve clause" in standard player contracts stated that even after the contract expired, the player could only negotiate with his current team. Through a series of court cases and collective bargaining, this clause became ineffective starting in 1976.

Since then, top players' salaries have skyrocketed in baseball (as well as in other sports such as Canadian and U.S. football and hockey and European soccer). With free agency—where star players can negotiate with any team—the average real salary rose from $214,500 (in 2016 dollars) in 1976 to $4.4 million in 2016. The highest 2016 salary was $33 million for Clayton Kershaw of the Dodgers, and 127 major-league players earned $10 million or more.

In 1999, Kevin Brown received the first contract for over $100 million. The Los Angeles Dodgers held a press conference to announce that they had re-signed Brown, their star pitcher, to a new seven-year $105 million contract. When reporters asked Mr. Brown what effect his contract would have on ticket prices, he responded, "I have never believed that players' salaries are directly related to ticket prices." The reporters snickered. Several of these newspaper pundits wrote that Mr. Brown's salary hike would drive up ticket prices to cover the expense.

Since then, salaries continued to rise. Giancarlo Stanton inked a 2015–2027 contract with the Marlins for $325 million.

After each such signing, a debate rings out about the effect of salaries on price:

Common Confusion: Higher salaries on a sports team causes its ticket prices to rise.

For example, a *St. Louis Today* columnist confidently predicted that the St. Louis Cardinals would raise ticket prices if the team re-signed Albert Pujols to a Rodriguez-like salary in 2012, and a *Washington Examiner* columnist blamed skyrocketing ticket prices over the last couple of decades on free spending on players by owners such as George Steinbrenner. Yet some writers in the *Wall Street Journal* and other newspapers now accept Mr. Brown's argument—that no link exists between salaries and ticket prices.

In Chapter 8, we showed that if both marginal and average cost curves shift up, the competitive market price rises. Does it follow that when a team re-signs a star athlete to a higher salary that it will raise ticket prices? If not, how does the sports example differ from the case where both marginal and average cost curves rise? We answer these questions in the Challenge Solution at the end of this chapter.

Sports teams, like other firms, hire labor and buy other inputs or factors, which they use to produce their outputs (goods and services). In this chapter, we show that a factor's market equilibrium price depends on the structure of factor markets and the output market. We first look at competitive factor and output markets, derive a competitive firm's demand curves for inputs, and determine the market equilibrium. Then, we examine the effect of a monopoly in either or both markets. Next, we consider markets in which a firm is a **monopsony**: the only buyer of a good in a market. A monopsony is the mirror image of a monopoly. Whereas a monopoly sells at a price higher than a competitive industry would charge, a monopsony buys at a lower price than a competitive industry would.

monopsony
the only buyer of a good in a given market

In this chapter, we examine three main topics

1. **Competitive Factor Market.** The intersection of the factor supply curve and factor demand curve (which depends on firms' production functions and the market price for output) determines the equilibrium in a competitive factor market.
2. **Effects of Monopolies on Factor Markets.** If firms exercise market power in either factor or output markets, the quantities of inputs and outputs sold fall.
3. **Monopsony.** A monopsony maximizes its profit by paying a price below the competitive level, which creates a deadweight loss for society.

15.1 Competitive Factor Market

Virtually all firms rely on factor markets for at least some inputs. The firms that buy factors may be competitive price takers or noncompetitive price setters, such as a monopsony. Perfectly competitive, monopolistically competitive, oligopolistic, and monopolistic firms sell factors. Here we examine factor markets in which buying and selling firms are competitive price takers. In the next section, we consider noncompetitive factor markets.

Factor markets with many small buyers and sellers are competitive. The FloraHolland flower auction in Amsterdam (Chapter 8) with 7,000 suppliers and 4,500 buyers typifies such a competitive market. The sellers supply inputs (flowers in bulk) to buyers, who sell outputs (trimmed flowers in vases, wrapped bouquets) at retail to final customers.

Chapter 5 derives the supply curve of labor by examining how individuals' choices between labor and leisure depend on tastes and the wage rate. The standard derivation of a competitive supply curve in output markets applies to factor markets as well. Chapter 8 determines the competitive supply curves of firms in general, including those that produce factors for other firms. Given that we know the supply curve, all we need to do to analyze a competitive factor market is to determine the factor's demand curve.

Short-Run Factor Demand of a Firm

A profit-maximizing firm's demand for a factor of production is downward sloping: The higher the price of an input, the less the firm wants to buy. To understand what is behind a firm's factor demand, we examine a firm that uses capital and labor to produce output from factors. Using the theory of how firms behave (Chapters 6 and 7), we show how the amount of an input the firm demands depends on the prices of the factors and the price of the final output.

We start by considering the short-run factor demand for labor of a firm that can vary labor but not capital. Then we examine long-run factor demands when both inputs are variable.

In the short run, a firm has a fixed amount of capital, $\overline{K}$, and can vary the number of workers, L, it employs. Will the firm's profit rise if it hires one more worker? The answer depends on whether its revenue or labor costs rise more when output expands.

marginal revenue product of labor (MRP_L) the extra revenue from hiring one more worker

An extra worker per hour raises the firm's output per hour, q, by the marginal product of labor, $MP_L = \Delta q/\Delta L$ (Chapter 6). How much is that extra output worth to the firm? The extra revenue, R, from the last unit of output is the firm's marginal revenue, $MR = \Delta R/\Delta q$. As a result, the **marginal revenue product of labor** (MRP_L), the extra revenue from hiring one more worker, is[1]

$$MRP_L = MR \times MP_L.$$

For a firm that is a competitive employer of labor, the marginal cost of hiring one more worker per hour is the wage, w. Hiring an extra worker raises the firm's profit if the marginal benefit—the marginal revenue product of labor—is greater than the marginal cost—the wage—from one more worker: $MRP_L > w$. If the marginal revenue product of labor is less than the wage, $MRP_L < w$, the firm can raise its profit by reducing the number of workers it employs. Thus, *the firm maximizes its profit by hiring workers until the marginal revenue product of the last worker exactly equals the marginal cost of employing that worker, which is the wage*:

$$MRP_L = w.$$

For now, we restrict our attention to competitive firms. A competitive firm faces an infinitely elastic demand for its output at the market price, p, so its marginal revenue is p (Chapter 8), and its marginal revenue product of labor is

$$MRP_L = p \times MP_L.$$

A competitive firm's *marginal revenue product* for any input is also called the *value of the marginal product* because it equals the market price times the marginal product of labor: the market value of the extra output.

The competitive firm hires labor to the point at which its marginal revenue product of labor equals the wage:

$$MRP_L = p \times MP_L = w. \tag{15.1}$$

Table 15.1 illustrates the relationship in Equation 15.1. If the firm hires $L = 3$ workers per hour, the marginal product from the third worker is 5 units of output per hour. Because the firm can sell the output at the market price $p = \$3$ per unit, the extra revenue from hiring the third worker is $MRP_L = p \times MP_L = \$3 \times 5 = \$15$. By hiring this worker, the firm increases its profit because the wage of this worker is only $w = \$12$. If the firm hires a fourth worker, the marginal product of labor from this last worker falls to 4, and the marginal revenue product of labor falls to \$12.

---

[1]In the short run, output is a function of only labor, $q(L)$. The price the firm receives from selling q units of output is given by its demand function, $p(q)$. Thus, the revenue that the firm receives is $R(L) = p[q(L)]q(L)$. The extra revenue that the firm obtains from using an extra amount of labor services is derived using the chain rule of differentiation:

$$MRP_L \equiv \frac{dR}{dL} = \frac{dR}{dq} \times \frac{dq}{dL} \equiv MR \times MP_L.$$

Table 15.1 Marginal Product of Labor, Marginal Revenue Product of Labor, and Marginal Cost

Labor, L	Marginal Product of Labor, MP_L	Marginal Revenue Product of Labor, $MRP_L = 3MP_L$	Change in Profit	Output, q	Marginal Cost, $MC = 12/MP_L$
2	6	\$18	6	13	\$2
3	5	\$15	3	18	\$2.4
4	4	\$12	0	22	\$3
5	3	\$9	−3	25	\$4
6	2	\$6	−6	27	\$6

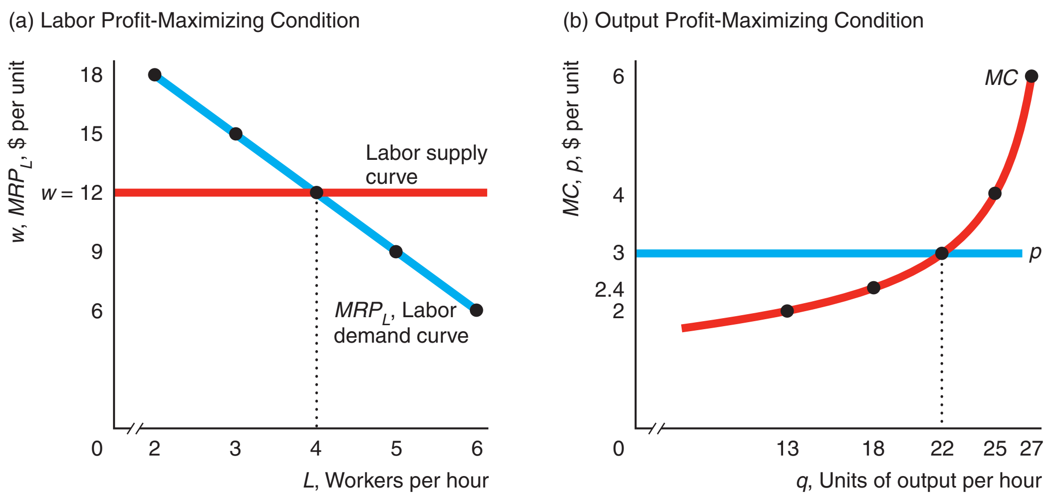
Notes: Wage, w, is \$12 per hour of work. Price, p, is \$3 per unit of output. Labor is variable, and capital is fixed.

Thus, the extra revenue from the last worker exactly equals that worker's wage, so the firm's profit is unchanged. Were the firm to hire a fifth worker, the $MRP_L = \$9$ is less than the wage of \$12, so its profit would fall.

Panel a of Figure 15.1 shows the same relationship. The wage line, $w = \$12$, intersects the MRP_L curve at $L = 4$ workers per hour. *The wage line is the supply of labor the firm faces*. As a competitive buyer of labor services, the firm can hire as many workers as it wants at a constant wage of \$12. *The marginal revenue product of labor curve, MRP_L, is the firm's demand curve for labor* when other inputs are fixed. It shows the maximum wage a firm is willing to pay to hire a given number of workers. Thus, the intersection of the supply curve of labor facing the firm and the firm's demand curve for labor, Equation 15.1, determines the profit-maximizing number of workers.

Figure 15.1 The Relationship Between Labor Market and Output Market Equilibria MyLab Economics Video

(a) The firm's profit is maximized at $L = 4$ workers per hour where the wage line, $w = \$12$, crosses the marginal revenue product of labor, MRP_L, curve, which is also the demand curve for labor. (b) The firm's profit is maximized at 22 units of output (produced by 4 workers), for which its marginal cost curve, $MC = w/MP_L$, equals the market price, $p = \$3$.



A firm's labor demand curve is usually downward sloping because of the law of diminishing marginal returns (Chapter 6). The marginal product from extra workers, MP_L, of a firm with fixed capital eventually falls as the firm increases the amount of labor it uses. Table 15.1 illustrates that the marginal product of labor falls from 6 for the second worker to 2 for the sixth worker. Because the marginal product of labor declines as more workers are hired, the marginal revenue product of labor (which equals a constant price times the marginal product of labor) or demand curve must slope downward as well.

Solved Problem 15.1

MyLab Economics Solved Problem

A paper firm's short-run Cobb-Douglas production function is estimated to be[2]

$$q = L^{0.6}\overline{K}^{0.2}, \tag{15.2}$$

where $\overline{K}$ is the fixed amount of capital. In the short-run, the firm's capital is fixed at $\overline{K} = 32$. The price of a unit of paper is $p = 50$. What is the paper firm's short-run production function and labor demand function? How many workers does the firm hire if the wage is $w = 15$?

Answer

1. *Determine the paper firm's short-run production function by substituting $\overline{K}$ into* Equation 15.2. Setting $\overline{K} = 32$ in Equation 15.2, we find that $q = L^{0.6}32^{0.2} = 2L^{0.6}$.

2. *Use the short-run production function to obtain the short-run marginal product of labor, then multiply that by price to determine the marginal product of labor.* The marginal product of labor of the short-run Cobb-Douglas production function is $MP_L = 0.6(2L^{0.6})/L = 1.2L^{-0.4}$ (Appendix 6C). Thus, given that $p = 50$, the short-run marginal revenue product of labor (or the labor demand function) is $MRP_L = 1.2pL^{-0.4} = 60L^{-0.4}$.

3. *Determine the number of workers per hour that the firm hires using Equation 15.1.* According to Equation 15.1, $MRP_L = w$, so for the paper firm, $60L^{-0.4} = 15$, or $L = 32$.

Profit Maximization Using Labor or Output Chapter 8 presents another profit-maximization condition: A competitive firm maximizes its profit by operating where the market price, p, equals the marginal cost of an extra unit of output, MC: $p = MC$. This output profit-maximizing condition is equivalent to the labor profit-maximizing condition in Equation 15.1. Dividing Equation 15.1 by MP_L, we find that

$$p = \frac{w}{MP_L} = MC.$$

As Chapter 7 shows, the marginal cost equals the wage, w, times 1 over the marginal product of labor, which is the extra labor, $\Delta L/\Delta q$, necessary to produce one more unit of output. The marginal cost is the cost of the extra labor, $w\Delta L$, needed to produce the extra output, Δq.

---

[2]This production function is based on the estimates of Hossain et al. (2012) for a Bangladesh paper firm. I chose the units of output so that the constant multiplier A in the general Cobb-Douglas, $q = AL^aK^b$, equals 1.

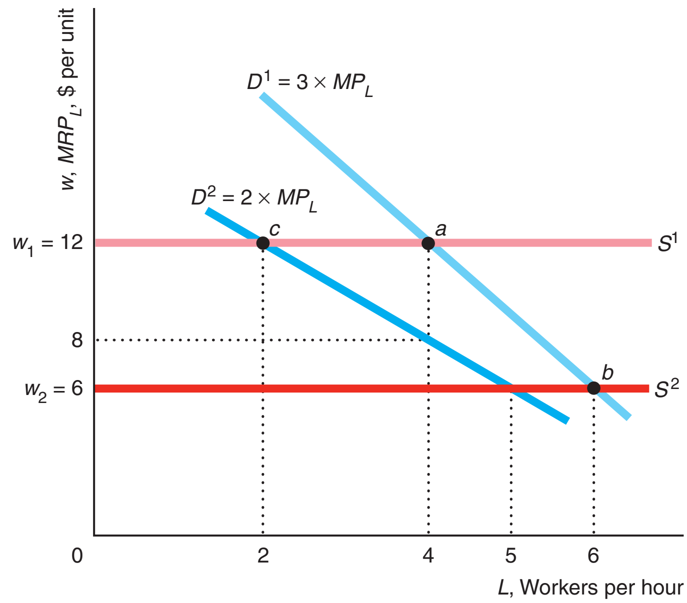
Table 15.1 illustrates this relationship. The fourth column shows how the amount of output produced varies with the number of workers. Because 3 workers produce 18 units of output and 4 workers produce 22 units of output, the marginal product of the fourth worker is 4 units of output. With a wage of \$12, the marginal cost for the last unit of output is $MC = w/MP_L = \$12/4 = \$3$. The market price is also \$3, so the firm maximizes its profit by producing 22 units of output, as panel b of Figure 15.1 illustrates.

In summary, the two profit-maximizing equilibria in Figure 15.1 give the same answer: The firm maximizes its profit by hiring 4 workers to produce 22 units of output. Panel a shows that the firm maximizes its profit by hiring 4 workers, for which the marginal benefit or marginal revenue product from the last worker, MRP_L, equals the marginal cost of that worker, w. Panel b shows that the firm maximizes its profit by producing 22 units of output, for which the marginal benefit or marginal revenue from the last unit of output, $p = \$3$, equals the marginal cost of the last unit of output, MC.

How Changes in Wages and Prices Affect Factor Demand The number of workers a firm hires depends on the wage and the price of the final good, as Equation 15.1 shows. Suppose that the supply of labor shifts so that the wage falls from $w_1 = \$12$ to $w_2 = \$6$ while the market price remains constant at \$3. The firm hires more workers because the cost of more labor falls while the incremental revenue from additional output is unchanged. Figure 15.2 shows that a drop in the wage due to a downward shift of the labor supply curve from S^1 to S^2 causes a shift along the labor demand curve D^1 from point *a*, where the firm hires 4 workers, to point *b*, where the firm hires 6 workers per hour.

Figure 15.2 Shift of and Movement Along the Labor Demand Curve

If the market price is \$3, the firm's labor demand curve is D^1. A fall in the wage causes a *shift of the supply curve* from S^1 to S^2 and a *movement along the demand curve for labor*. If the wage is $w_1 = \$12$, the firm hires 4 workers per hour, equilibrium point *a*. If the wage falls to $w_2 = \$6$, the firm hires 6 workers, point *b*. A fall in the market price to \$2 causes a *shift of the firm's demand curve for labor* from D^1 to D^2. If the market wage stays constant at $w_1 = \$12$, the fall in the market price causes a *movement along the supply curve* S^1: The number of workers the firm hires falls from 4, point *a* on D^1 and S^1, to 2, point *c* on D^2 and S^1.



If the market price falls from \$3 to \$2, the demand curve for labor shifts downward from D^1 to D^2. Demand D^2 is only $\frac{2}{3} = (2MP_L)/(3MP_L)$ as high as D^1 at any given quantity of labor. If the wage stays constant at $w_1 = \$12$, the firm reduces its demand for workers from 4, point *a*, to 2, point *c*. Thus, a shift in either the market wage or the market price affects the amount of labor that a firm employs.

Solved Problem 15.2

MyLab Economics Solved Problem

How does a competitive firm adjust its demand for labor when the government imposes a specific tax of t on each unit of output?

Answer

1. *Give intuition*. The specific tax lowers the price per unit the firm receives, so we can apply the same type of analysis we just used for a fall in the market price.

2. *Show how the tax affects the marginal revenue product of labor*. The marginal revenue product of labor for a competitive firm is the price the firm receives for the good times the marginal product of labor. The tax reduces the price the firm receives. The tax does not affect the relative prices of labor and capital, so it does not affect the marginal product of labor for a given amount of labor, $MP_L(L)$. For a given amount of labor, the marginal revenue product of labor falls from $p \times MP_L(L)$ to $(p - t) \times MP_L(L)$. The marginal revenue product of labor curve—the labor demand curve—shifts downward until it is only $(p - t)/p$ as high as the original labor demand curve at any quantity of labor.

Long-Run Factor Demand

In the long run, the firm may vary all of its inputs. Now if the wage of labor rises, the firm adjusts both labor and capital. As a result, the short-run marginal revenue product of labor curve that holds capital fixed is not the firm's long-run labor demand curve. The long-run labor demand curve takes account of changes in the firm's use of capital as the wage rises.

In both the short and long run, the labor demand curve is the marginal revenue product curve of labor. In the short run, the firm cannot vary capital, so the short-run MP_L curve, and hence the short-run MRP_L curve, are both relatively steep. In the long run, when the firm can vary all inputs, its long-run MP_L curve and MRP_L curve are flatter.

Figure 15.3 shows the relationship between the long- and short-run labor demand curves for the paper firm in Solved Problem 15.1.[3] In the short run, capital is fixed at $\overline{K} = 32$, the wage is $w = \$15$, the rental rate of capital is $r = \$5$, and the price is $p = \$50$. The firm hires 32 workers per hour, point *a* on its short-run labor demand curve, where $\overline{K} = 32$. Using 32 workers and 32 units of capital is profit maximizing in the long run, so point *a* is also on the firm's long-run labor demand curve.

[3]Appendix 15A formally shows that the long-run labor demand and capital demand functions for a Cobb-Douglas production function are functions of the market price, p; the wage rate, w; and the rental rate of capital, r. Substituting the parameters for the paper firm, $a = 0.6$, $b = 0.2$, and $A = 1$, into Equation 15A.4, we find that the firm's long-run labor demand curve is

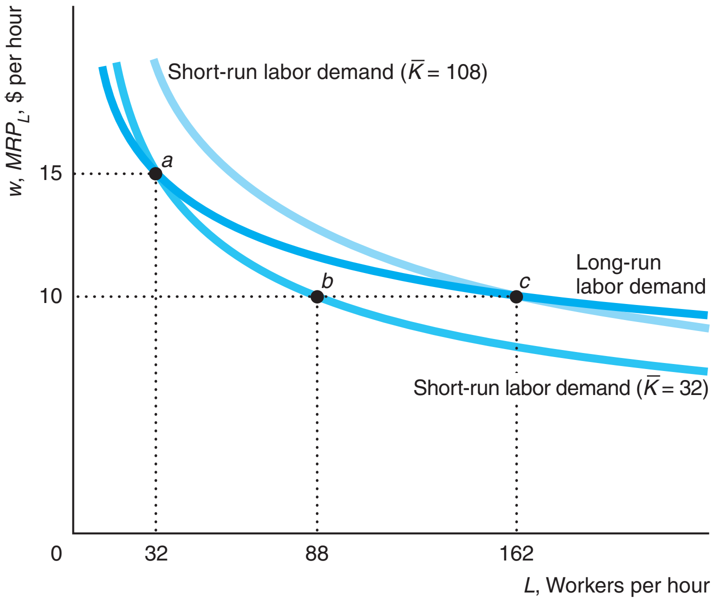
$$L = (0.6/w)^4(0.2/r)p^5.$$

Its long-run capital demand curve, Equation 15A.5, is

$$K = (0.6/w)^3(0.2/r)^2p^5.$$

Figure 15.3 Labor Demand of a Paper Firm

If the long-run market price is \$50 per unit, the rental rate of capital services is $r = \$5$, and the wage is $w = \$15$ per hour, a paper firm hires 32 workers (and uses 32 units of capital) at point *a* on its long-run labor demand curve. In the short run, if capital is fixed at $\bar{K} = 32$, the firm still hires 32 workers per hour at point *a* on its short-run labor demand curve. If the wage drops to \$10 and capital remains fixed at $\bar{K} = 32$, the firm would hire 88 workers, point *b* on the short-run labor demand curve. In the long run, however, it would increase its capital to $K = 108$ and hire 162 workers, point *c* on the long-run labor demand curve and on the short-run labor demand curve with $K = 108$.



In the short run, if the wage fell to \$10, the firm could not increase its capital, so it would hire 88 workers, point *b* on the short-run labor demand curve, where $\bar{K} = 32$. In the long run, however, the firm would employ more capital and even more labor (because it can sell as much output as it wants at the market price). It would hire 162 workers and use 108 units of capital, which is point *c* on both the long-run labor demand curve and the short-run labor demand curve for $K = 108$.

Factor Market Demand

A factor market demand curve is the sum of the factor demand curves of the various firms that use the input. Determining a factor market demand curve is more difficult than deriving consumers' market demand for a final good. In a single market, we derive the market demand curve by horizontally summing the demand curves for individual consumers (Chapter 2). However, inputs such as labor and capital are used in many output markets. Thus, to derive the labor market demand curve, we first determine the labor demand curve for each output market and then sum across output markets to obtain the factor market demand curve.

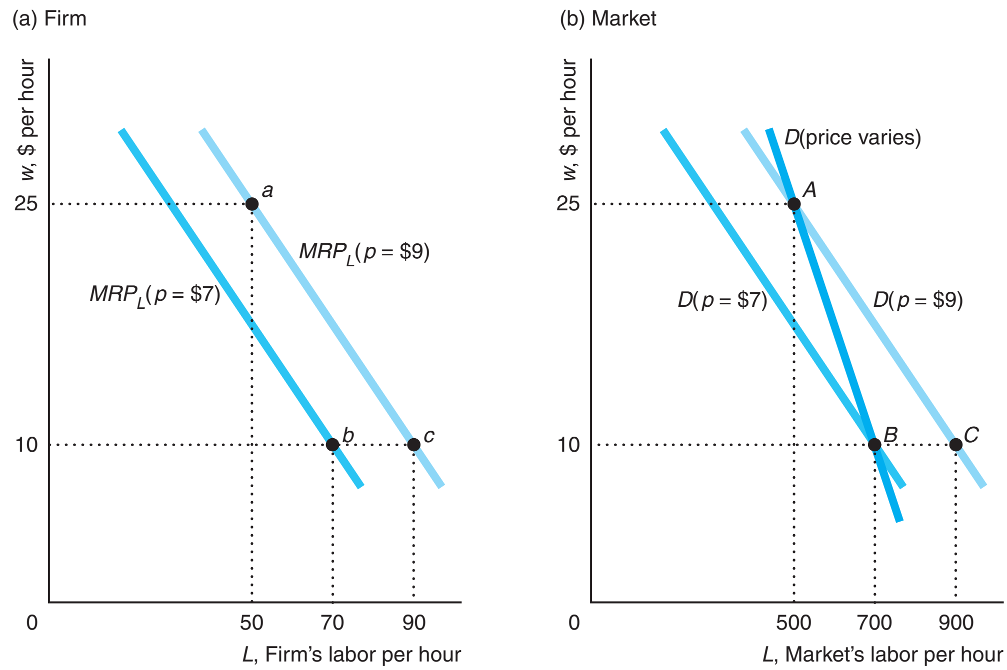
The Marginal Revenue Product Approach Earlier we derived the factor demand of a competitive firm that took the output market price as given. The problem we now face is that the output market price depends on the factor's price. As the factor's price falls, each firm, taking the original market price as given, uses more of the factor to produce more output. This extra production by all the firms in the market causes the market price to fall. As the market price falls, each firm reduces its output and hence its demand for the input. Thus, a fall in an input price causes less of an increase in factor demand than would occur if the market price remained constant, as Figure 15.4 illustrates.

At the initial output market price of \$9 per unit, the competitive firm's labor demand curve (panel a) is $MRP_L(p = \$9) = \$9 \times MP_L$. When the wage is \$25 per hour, the firm hires 50 workers: point *a*. The 10 firms in the market (panel b) demand 500 hours of work: point *A* on the demand curve $D(p = \$9) = 10 \times \$9 \times MP_L$. If the wage falls to \$10 while the market price remains fixed at \$9, each firm hires 90 workers, point *c*, and all the firms in the market would hire 900 workers, point *C*. However, the extra output drives the price down to \$7, so each firm hires 70 workers, point *b*, and the firms collectively demand 700 workers, point *B*. The market labor demand curve for this output market that takes price adjustments into account, *D* (price varies), goes through points *A* and *B*. Thus, the market's demand for labor is steeper than it would be if output prices were fixed.

An Alternative Approach For certain types of production functions, it is easier to determine the market demand curve by using the output profit-maximizing equation rather than the marginal revenue product approach. Suppose that calculator

Figure 15.4 Firm and Market Demand for Labor MyLab Economics Video

When the output price is $p = \$9$, the individual competitive firm's labor demand curve is $MRP_L(p = \$9)$. If $w = \$25$ per hour, the firm hires 50 workers, point *a* in panel a, and the 10 firms in the market demand 500 workers, point *A* on the labor demand curve $D(p = \$9)$ in panel *b*. If the wage falls to \$10, each firm would hire 90 workers, point *c*, if the market price stayed fixed at \$9. The extra output, however, drives the price down to \$7, so each firm hires 70 workers, point *b*. The market's demand for labor that takes price adjustments into account, *D*(price varies), goes through points *A* and *B*.



manufacturers are competitive and use a fixed-proportions production function, producing each calculator using one microchip and one plastic case. Each plastic case costs p_p, and each microchip costs p_m. What is the calculator market's demand for microchips?

Figure 15.5 shows the demand both for calculators, Q, and microchips, M. Because the numbers of chips and calculators are equal, $Q = M$, the horizontal axes for chips and calculators are the same.

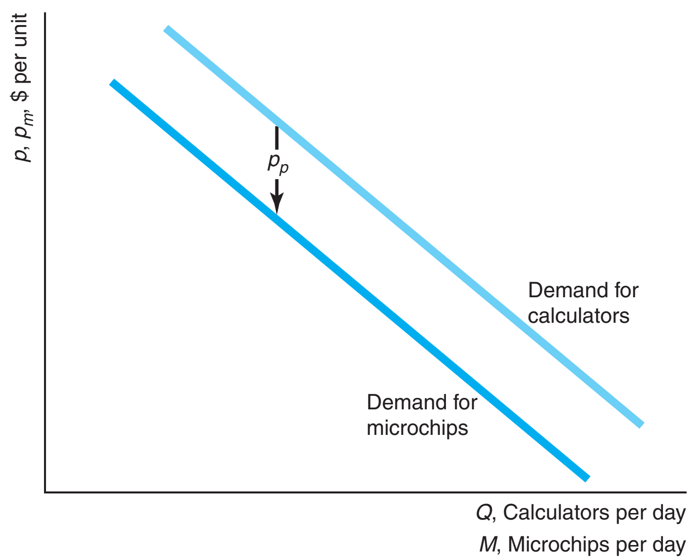
Because each calculator requires one chip and one case, the marginal cost of producing a calculator is $MC = p_p + p_m$. Each competitive firm operates where the market price equals the marginal cost: $p = p_p + p_m = MC$. As a result, the most that any firm would pay for a silicon chip is $p_m = p - p_p$, the amount left over from selling a calculator after paying for the plastic case. Thus, the calculator market's demand curve for microchips lies p_p below the demand curve for calculators, as the figure shows.[4]

Competitive Factor Market Equilibrium

The intersection of the factor market demand curve and the factor market supply curve determines the competitive factor market equilibrium. We've just derived the factor market demand. The long-run factor supply curve for each firm is its marginal cost curve above the minimum of its average cost curve, and the factor market supply

Figure 15.5 Demand for Microchips in Calculators

It takes one microchip, which costs p_m, and one plastic case, which costs p_p, to produce a calculator, so the marginal cost of a calculator is $MC = p_m + p_p$. Competitive firms operate where the price of a calculator is $p = p_m + p_p$. Thus, the demand curve for a microchip lies p_p below that of a calculator.



[4]The inverse demand function for calculators is a decreasing function of quantity, $p(Q)$. Similarly, the inverse demand function for microchips is $p_m(M)$. Because $Q = M$, we can write the profit-maximization condition as $p(Q) = p_m(M) + p_p$. Thus, the demand for chips lies p_p below the demand for calculators: $p_m(M) = p(Q) - p_p$.

curve is the horizontal sum of the firms' supply curves (Chapter 8). As we've already analyzed competitive market equilibria for markets in general in Chapters 2, 3, 8, and 9, we won't repeat the analysis. (Been there. Done that.)

Factor prices are equalized across markets (Chapter 10). For example, if wages were higher in one industry than in another, workers would shift from the low-wage industry to the high-wage industry until the wages were equalized.

15.2 Effects of Monopolies on Factor Markets

Having examined the factor market equilibrium where competitive firms sell a factor to a competitive output market, we now survey the effects of market power on factor market equilibrium. If firms in the output market *or* the factor market exercise market power by setting price above marginal cost, fewer factors are sold than would be sold if all firms were competitive.

Market Structure and Factor Demands

Factor demand curves vary with market power. As we saw in Chapters 11 and 12, the marginal revenue of a profit-maximizing firm, $MR = p(1 + 1/\ \ )$, is a function of the elasticity, , of its output demand curve and the market price, p. Thus, the firm's marginal revenue product of labor function is

$$MRP_L = p\left(1 + \frac{1}{\ }\right)MP_L.$$

The labor demand curve is $p \times MP_L$ for a competitive firm because it faces an infinitely elastic demand at the market price, so its marginal revenue equals the market price.

The marginal revenue product of labor or labor demand curve for a competitive market is above that of a monopoly or oligopoly firm. Figure 15.6 shows the short-run market factor demand for a paper firm if it is a competitive firm, one of two identical Cournot quantity-setting firms, or a monopoly.[5]

A monopoly operates in the elastic section of its downward-sloping demand curve (Chapter 11), so its demand elasticity is less than -1 and finite: $-\infty < \ \ \leq -1$. As a result, at any given price, the monopoly's labor demand, $p(1 + 1/\ \ )MP_L$, lies below the labor demand curve, pMP_L, of a competitive firm with an identical marginal product of labor curve.

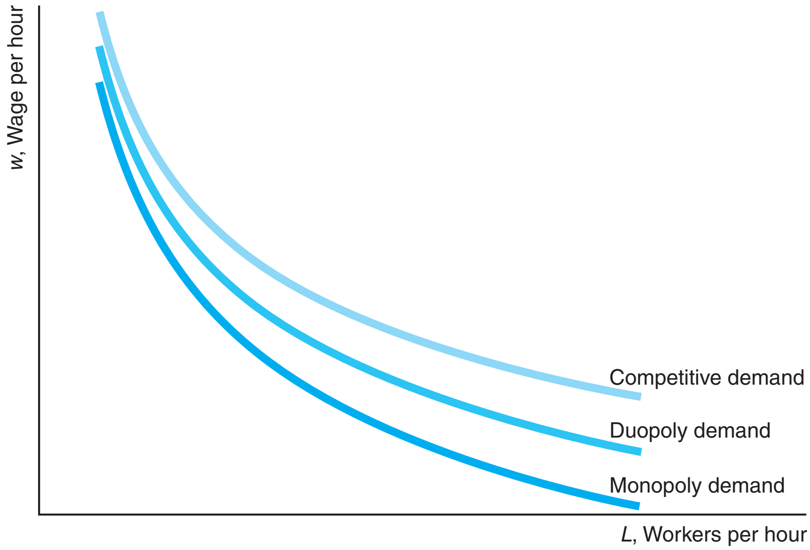
The elasticity of demand a Cournot firm faces is $n\ \ $, where n is the number of identical firms and is the market elasticity of demand (Chapter 13). Given that they have the same market demand curve, a duopoly Cournot firm faces a demand curve that is twice as elastic as the demand curve a monopoly faces. Consequently, a Cournot duopoly firm's labor demand curve, $p[1 + 1/(2\ \ )]MP_L$, lies above that of a monopoly but below that of a competitive firm. From now on, we concentrate on the competitive and monopoly equilibria because the oligopoly and monopolistically competitive equilibria lie between these polar cases.

---

[5]In the short run, the paper firm's marginal product function is $MP_L = 1.2L^{-0.4}$. The labor demand is $p \times 1.2L^{-0.4}$ for a competitive firm, $p[1 + 1/(2\ \ )] \times 1.2L^{-0.4}$ for one of two identical Cournot duopoly firms, and $p(1 + 1/\ \ ) \times 1.2L^{-0.4}$ for a monopoly.

Figure 15.6 How Paper Firm Labor Demand Varies with Market Structure

For all profit-maximizing firms, the labor demand curve is the marginal revenue product of labor: $MRP_L = MR \times MP_L$. Because marginal revenue differs with market structure, so does the MRP_L. At a given wage, a competitive paper firm demands more workers than a Cournot duopoly firm, which demands more workers than a monopoly.



A Model of Market Power in Input and Output Markets

When a firm with market power in either the factor or the output market raises its price, the price to final consumers rises. As a result, consumers buy fewer units, so the firm demands fewer units of the input. We use a linear example to illustrate how monopolies affect factor market equilibrium. The inverse demand function, $p(Q)$, for the final good is

$$p = 80 - Q. \tag{15.3}$$

Figure 15.7 plots this demand curve. The firm can hire an unlimited number of workers at \$20 an hour. Each unit of output, Q, requires one unit of labor, L, and no other factor, so the marginal product of labor is 1.

As a benchmark, we start our analysis with competitive factor and output markets. Then we ask how the factor market equilibrium changes for a monopoly. Next, we examine a monopolized factor market and a competitive output market. Finally, we investigate the effect of market power in both markets.

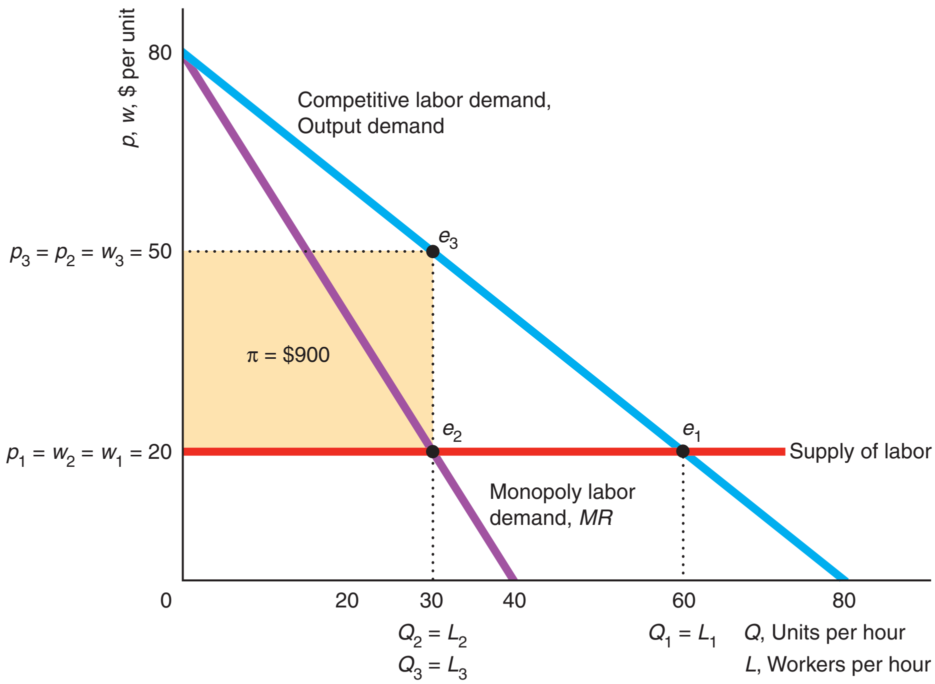
Competitive Factor and Output Markets The intersection of the relevant supply and demand curves determines the competitive equilibria in both input and output markets in Figure 15.7. Because $Q = L$, the figure measures both output and labor on the same horizontal axis.

The marginal product of labor is 1 because one extra worker produces one more unit of output. Thus, the competitive market's demand for labor, $MRP_L = p \times MP_L = p$, is identical to the output demand curve. The labor demand function is the same as the output demand function, where we replace p with w and Q with L:

$$w = 80 - L. \tag{15.4}$$

Figure 15.7 Effect of Output Market Structure on Labor Market Equilibrium

Because one unit of output is produced with one unit of labor, the marginal product of labor is 1, so the competitive labor demand curve is the same as the output demand curve. If both markets are competitive, the labor market equilibrium is e_1. A monopoly's labor demand curve is identical to its marginal revenue curve. An output monopoly charges final consumers a higher price, so it buys less labor. The new labor equilibrium is e_2. With a labor monopoly (union), the equilibrium is e_3.



The competitive supply of labor is a horizontal line at \$20. Given a competitive output market, the intersection of this supply curve of labor and the competitive demand for labor (Equation 15.4) determines the labor market equilibrium, e_1, where $20 = 80 - L$. Thus, the competitive equilibrium amount of labor services is $L_1 = 60$, and the equilibrium wage is $w_1 = \$20$.

The cost of producing a unit of output equals the wage, so the supply curve of output is also horizontal at \$20. The intersection of this output supply curve and the output demand curve, Equation 15.3, occurs at $Q_1 = 60$ and $p_1 = \$20$. A competitive firm's average cost, w_1, exactly equals the price at which it sells its good, p_1, so the competitive firm breaks even.

Competitive Factor Market and Monopolized Output Market Because a monopoly in the output market charges a higher price than a competitive market would, it sells fewer units of output and hires fewer workers. The monopoly faces a competitive labor supply curve that is horizontal at the wage $w_2 = \$20$. Thus, the output monopoly's marginal cost is \$20 per unit.

The monopoly's marginal revenue curve is twice as steep as the linear output demand curve it faces (Chapter 11):

$$MR_Q = 80 - 2Q.$$

The monopoly maximizes its profit where its marginal revenue equals its marginal cost:

$$MR_Q = 80 - 2Q = 20 = MC.$$

Thus, the equilibrium quantity is $Q_2 = 30$. Substituting this quantity into the output demand function, Equation 15.3, we find that the equilibrium price is \$50. The monopoly makes $p_2 - w_2 = \$50 - \$20 = \$30$ per unit. Its profit is $\pi = \$900$, as the shaded rectangle in Figure 15.7 shows.

Because the monopoly's marginal product of labor is 1, its demand curve for labor equals its marginal revenue curve:

$$MRP_L = MR_Q \times MP_L = MR_Q.$$

We obtain its labor demand function by replacing Q with L and MR_Q with w in its marginal revenue function:

$$w = 80 - 2L.$$

The intersection of the competitive labor supply curve, $w_2 = \$20$, and the monopoly's demand for labor curve determines the labor market equilibrium, e_2, where $80 - 2L = 20$. Thus, the equilibrium amount of labor is $L_2 = 30$.

This example illustrates that a monopoly hurts final consumers and drives some sellers of the factor (workers) out of this market. Final consumers pay \$30 more per unit than they would pay if the market were competitive. Because of the higher price, consumers buy less output, $Q_2 = 30 < 60 = Q_1$. Consequently, the monopoly demands less labor than a competitive market does: $L_2 = 30 < 60 = L_1$. If the supply curve of labor were upward sloping, this reduction in demand would also reduce workers' wages.

Monopolized Factor Market and Competitive Output Market Now suppose that the output market is competitive and that a labor monopoly exists. One possibility is that the workers form a union that acts as a monopoly. Instead, for simplicity, we'll assume that the labor monopoly is the only firm that can supply the workers employed in the output market.[6]

The labor monopoly sets its marginal revenue equal to its marginal cost, which is \$20. Because the competitive output market's labor demand curve is the same as the output demand curve, the marginal revenue curve this labor monopoly faces is the same as the marginal revenue curve of an output monopoly, where we replace Q with L:

$$MR_L = 80 - 2L.$$

The labor monopoly operates at e_3 in Figure 15.7, where its marginal revenue equals its marginal cost of \$20:

$$80 - 2L = 20.$$

The labor monopoly sells $L_3 = 30$ hours of labor services. Substituting this quantity into the labor demand curve, Equation 15.4, we find that the monopoly wage is $w_3 = \$50$. Because the labor monopoly makes $w_3 - \$20 = \$30$ per hour of labor services and it sells 30 hours, its profit is $\pi = \$900$.

[6]Many markets have firms that only supply labor to other firms. Manpower, Kelly Services, and Accountemps provide temporary office workers and other employees. Many construction firms supply only skilled craftspeople. Still other firms specialize in providing computer programmers.

The competitive supply to the output market is horizontal at $w_3 = \$50$. The output equilibrium occurs where this supply curve hits the output demand curve, Equation 15.3: $50 = 80 - Q$. Thus, the equilibrium quantity is $Q_3 = 30$. The equilibrium price is the same as the wage, $p_3 = w_3 = \$50$. As a result, the output firms break even.

In our example, in which one unit of labor produces one unit of output, consumers fare the same whether the labor market or the output market is monopolized. Consumers pay $p_2 = p_3 = \$50$ and buy $Q_2 = Q_3 = 30$ units of output. The labor market equilibria are different: The wage is higher if the monopoly is in the labor market rather than the output market. The profit goes to the monopoly regardless of which market is monopolized.

Application

Unions and Profits

Workers acting collectively within a union can raise the wage much in the same manner as any other monopoly. A union's success in raising the wage depends on the elasticity of demand it faces, members' ability to act collectively, laws, and the share of the labor market that is unionized.

In the United States, if the majority of workers in a firm vote to unionize, the union contract will cover all workers. Through the union's negotiations with the firm, workers' wages may rise. Consequently, following unionization, the value of the stock of that firm—which reflects the profitability of the firm—may fall. Lee and Mas (2012) estimated that the average decrease in the value of a unionized firm is \$40,500 (in \$1998) per worker eligible to vote, or about a 10% drop in the value of the firm.

Lee and Mas noted that this drop in value following unionization is due to workers capturing some of the firm's former profit through higher wages and the rest results from inefficiency because the firm cannot use the optimal mix of labor and capital. Based on estimates from other studies, Lee and Mas calculated that 8% of the loss in value is due to higher wages and 2% stems from inefficiency.

Similarly, Stanfield and Tumarkin (2016) found that a 2012 law in the state of New South Wales, Australia that restricts union political activity weakened unions. It allowed firms to negotiate more favorable contract terms and raised their values.

Monopoly in Successive Markets If the labor and output markets are both monopolized, consumers get hit with a double monopoly markup. The labor monopoly raises the wage, in turn raising the cost of producing the final output. The output monopoly then increases the final price even further.[7]

Figure 15.8 illustrates this double markup. The output monopoly's marginal revenue curve, $MR_Q = 80 - 2Q$, is the same as its labor demand curve, $w = 80 - 2L$. Because the labor demand curve is linear, the labor monopoly's marginal revenue curve is twice as steeply sloped:

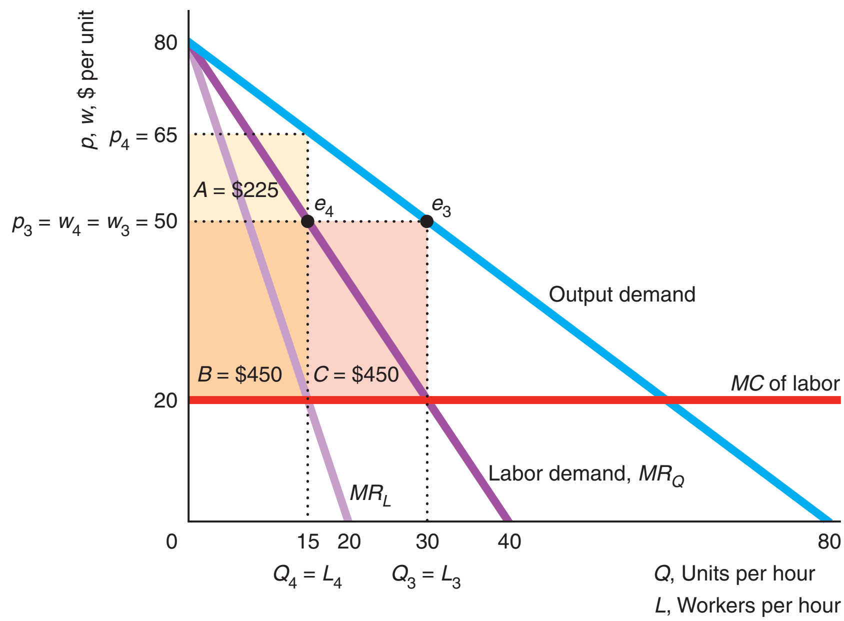
$$MR_L = 80 - 4L.$$

The labor monopoly maximizes its profit by setting its marginal revenue equal to its marginal cost: $80 - 4L = 20$. Thus, at the labor market equilibrium, e_4, the labor monopoly provides $L_4 = 15$ workers. Substituting this quantity into the

[7]In our example, the labor monopoly has a constant marginal cost of $m = \$20$. It operates where its marginal cost equals its marginal revenue, $w(1 + 1/\ {}_L)$, where ${}_L$ is the elasticity of labor demand. Thus, the wage is greater than marginal cost: $w = m\mu_L$, where $\mu_L = 1/(1 + 1/\ {}_L) > 1$ is the multiplicative labor monopoly markup. The wage is the output monopoly's marginal cost. The output monopoly further marks up the price: $p = w\mu_Q = m\mu_L\mu_Q$, where $\mu_Q = 1/(1 + 1/\ {}_Q) > 1$ is the multiplicative output monopoly markup and ${}_Q$ is the output demand elasticity.

Figure 15.8 Double Monopoly Markup

If a market has two successive monopolies, consumers are hit with a double monopoly markup. The labor market equilibrium is e_4, where the wage, w_4, is \$30 above the labor market's marginal and average cost of \$20. The product market monopoly's price, p_4, is \$15 above its marginal cost, w_4. If the labor monopoly integrates vertically, consumers gain ($p_3 < p_4$), and total profit increases from $A + B$ to $B + C$.



labor demand curve, $w = 80 - 2L$, we find that the labor monopoly's equilibrium wage is $w_4 = \$50$. Thus, the labor monopoly marks up its wage \$30 above its marginal cost. Its profit is area $B = \$30 \times 15 = \$450$ in the figure.

To maximize its profit, the output monopoly sets its marginal revenue, $MR_Q = 80 - 2Q$, equal to its marginal cost, $w_4 = \$50$. It sells $Q_4 = 15$ units of output. Substituting this quantity into the output demand curve, we learn that the output monopoly's equilibrium price is $p_4 = \$65$. The output monopoly's markup is \$15 above its marginal cost. Its profit is area $A = \$225$.

This double markup harms consumers. They pay a higher price—\$65 rather than \$50—than they would pay if a monopoly was in just one market or the other.

Solved Problem 15.3

How are consumers affected and how do profits change in the example if the labor monopoly buys the monopoly producer, which is called *vertical integration*?[8]

Answer

1. *Solve for the postmerger equilibrium.* The new merged monopoly's output demand is the market demand, and its marginal revenue from extra output is $MR_Q = 80 - 2Q$, as Figure 15.8 shows. Now that the firms are one, the former labor monopoly no longer marks up the labor to its production unit. Its marginal

[8]See the Supplemental Material "Vertical Integration" in MyLab Economics, Chapter 15, for more details on vertical integration.

cost of an extra unit of output is \$20. The monopoly maximizes its profit by setting its marginal cost equal to MR_Q. The resulting output equilibrium is the same as it would be with a single labor monopoly. Equilibrium output is $Q_3 = 30$ and $p_3 = \$50$. The integrated monopoly's profit is $\$30 \times 30 = \$900$, area $B + C$.

2. *Compare the premerger and postmerger equilibria*. Consumers benefit from this merger. Because the price they pay falls from $p_4 = \$65$ to $p_3 = \$50$, they buy 15 extra units of output. The firms also benefit. The combined profit with two monopolies is areas $A + B = \$675$, which is less than the profit of the integrated firm, areas $B + C = \$900$. The labor monopoly can offer the output monopoly more than it earns as a separate firm and still increase its own profit: The firms can split the extra \$225. Thus, everyone may gain from a vertical merger that eliminates one of the two monopoly markups.

Comment: The potential of a double markup occurs in many markets. For example, for the first five years after the Apple iPhone was introduced, it was sold only through AT&T in the United States. Apparently, Apple and AT&T worked together to avoid a double markup by *quasi-integrating:* signing a contract to share the profit and to set the price at the joint profit-maximizing level.

15.3 Monopsony

In Chapter 11, we saw that a *monopoly*, a single *seller*, picks a point—a price and a quantity combination—on the market *demand curve* that maximizes its profit. A *monopsony*, a single *buyer* in a market, chooses a price-quantity combination from the industry *supply curve* that maximizes its profit. A monopsony is the mirror image of a monopoly, and it exercises its market power by buying at a price *below* the price that competitive buyers would pay.

A U.S. manufacturer of state-of-the-art weapon systems can legally sell only to the federal government. U.S. professional football teams, which act collectively, are the only U.S. firms that hire professional football players.[9] Many fisheries have only one buyer of fish (or at most a small number of buyers, an *oligopsony*).

Monopsony Profit Maximization

Suppose that a firm is the sole employer in town—a monopsony in the local labor market. The firm uses only one factor, labor (L), to produce a final good. The value the firm places on the last worker it hires is the marginal revenue product of that worker—the value of the extra output the worker produces—which is the height of the firm's labor demand curve for the number of workers the firm employs.

The firm has a downward-sloping demand curve in panel a of Figure 15.9. The firm faces an upward-sloping supply curve of labor: The higher its daily wage, w, the more people want to work for the firm. The firm's *marginal expenditure*—the additional cost of hiring one more worker—depends on the shape of the supply curve.

The supply curve shows the average expenditure, or wage, the monopsony pays to hire a worker for one hour. For example, the monopsony's average expenditure or wage is $w_m = \$20$ if it hires $L = 20$ workers per hour. If the monopsony wants to hire one more worker, it must raise its wage because the supply curve is upward

[9] Football players belong to a union that acts collectively, like a monopoly, in an attempt to offset the monopsony market power of the football teams.

sloping. Because it pays all workers the same wage, the monopsony must also pay more to each worker it was already employing. Thus, the monopsony's marginal expenditure on the last worker is greater than that worker's wage.[10] The marginal expenditure curve in the figure is twice as steep as the linear supply curve.[11]

In contrast, if the firm were a competitive price taker in the labor market, it would face a supply curve that was horizontal at the market wage. Consequently, such a competitive firm's marginal expenditure to hire one more worker would be the market wage.

Any buyer—including a monopsony (Appendix 15B) or a competitive firm—hires workers up to the point at which the marginal value of the last worker's labor services equals the firm's marginal expenditure. If the labor services of the last worker is worth more to the buyer than its marginal expenditure, the buyer hires another worker. Similarly, if the services of the last worker is less valuable than its marginal expenditure, the buyer hires one fewer worker.

The monopsony hires 20 workers. The intersection of its marginal expenditure curve and the demand curve determines the monopsony optimum, e_m. The monopsony values the labor services of the last worker at \$40 (the height of its demand curve), and its marginal expenditure on that worker (the height of its marginal expenditure curve) is \$40. It pays only \$20 (the height of the supply curve). In other words, the monopsony values the last worker at \$20 more than it actually has to pay.

If the market in Figure 15.9 were competitive, the intersection of the market demand curve and the market supply curve would determine the competitive equilibrium at e_c, where buyers purchase 30 units at $w_c = \$30$ per unit. Thus, the monopsony hires fewer workers, 20 versus 30, than a competitive market would hire and pays a lower wage, \$20 versus \$30.

Monopsony power is the ability of a single buyer to pay less than the competitive price profitably. The size of the gap between the value the monopsony places on the last worker (the height of its demand curve) and the wage it pays (the height of the supply curve) depends on the elasticity of supply at the monopsony optimum. The markup of the marginal expenditure (which equals the value to the monopsony) over the wage is inversely proportional to the elasticity of supply at the optimum (Appendix 15B):

$$\frac{ME - w}{w} = \frac{1}{\eta}.$$

By comparing panels a and b in Figure 15.9, we see that the less elastic the supply curve is at the optimum, the greater the gap between marginal expenditure and the wage. At the monopsony optimum, the supply curve in panel b is more elastic than the supply curve in panel a.[12] The gap between the marginal expenditure and the

---

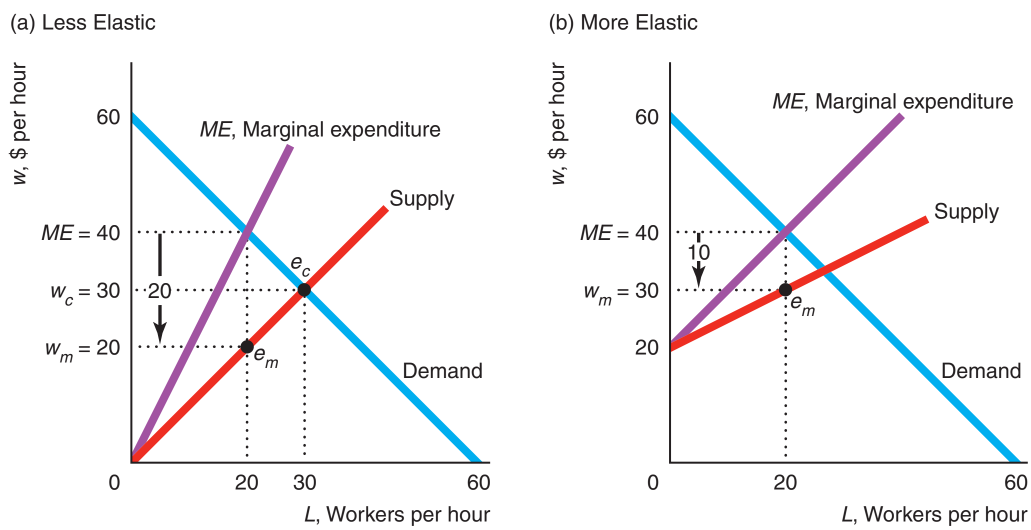
[10]The monopsony's total expenditure is $E = w(L)L$, where $w(L)$ is the wage given by the supply curve. Its marginal expenditure is $ME = \mathrm{d}E/\mathrm{d}L = w(L) + L[\mathrm{d}w(L)/\mathrm{d}L]$, where $w(L)$ is the wage paid the last worker and $L[\mathrm{d}w(L)/\mathrm{d}L]$ is the extra amount the monopsony pays the workers it was already employing. Because the supply curve is upward sloping, $\mathrm{d}w(L)/\mathrm{d}L > 0$, the marginal expenditure, ME, is greater than the average expenditure, $w(L)$.

[11]Appendix 15B shows that the ME curve is twice as steep as the labor supply curve for any linear labor supply curve.

[12]The supply curve in panel a is $w = L$, while that in panel b is $w = 20 + \frac{1}{2}L$. The elasticity of supply, $\eta = (\mathrm{d}L/\mathrm{d}w)(w/L)$, at the optimum is $w/L = 20/20 = 1$ in panel a and $2w/L = 2 \times 30/20 = 3$ in panel b. Consequently, the supply curve at the optimum is three times as elastic in panel b as in panel a.

Figure 15.9 Monopsony

(a) The marginal expenditure curve—the monopsony's marginal cost of buying one more unit—lies above the upward-sloping market supply curve. The monopsony equilibrium, e_m, occurs where the marginal expenditure curve intersects the monopsony's demand curve. The monopsony hires fewer workers at a lower wage, $w_m = \$20$, than would firms in a competitive market, which pay $w_c = \$30$. (b) The supply curve is more elastic at the optimum than in (a), so the value that the monopsony places on the last unit (which equals the marginal expenditure of \$40) exceeds the price the monopsony pays, $w_m = \$30$, by less than in (a).



wage is greater in panel a, $ME - w = \$20$, than in panel b, $ME - w = \$10$. Similarly, the markup in panel a, $(ME - w)/w = 20/20 = 1$, is much greater than that in panel b, $(ME - w)/w = 10/30 = \frac{1}{3}$.

Application

Monopsony and the Gender Wage Gap

If workers cannot easily leave their current jobs to look for work elsewhere, their employers can exercise monopsony power, lowering their wages. Many married women face different job market opportunities than do men. If they need to stay near home because of childcare and other household responsibilities, their alternative job opportunities are limited, which may partially explain why women earn less than comparably skilled men (Madden, 1973).

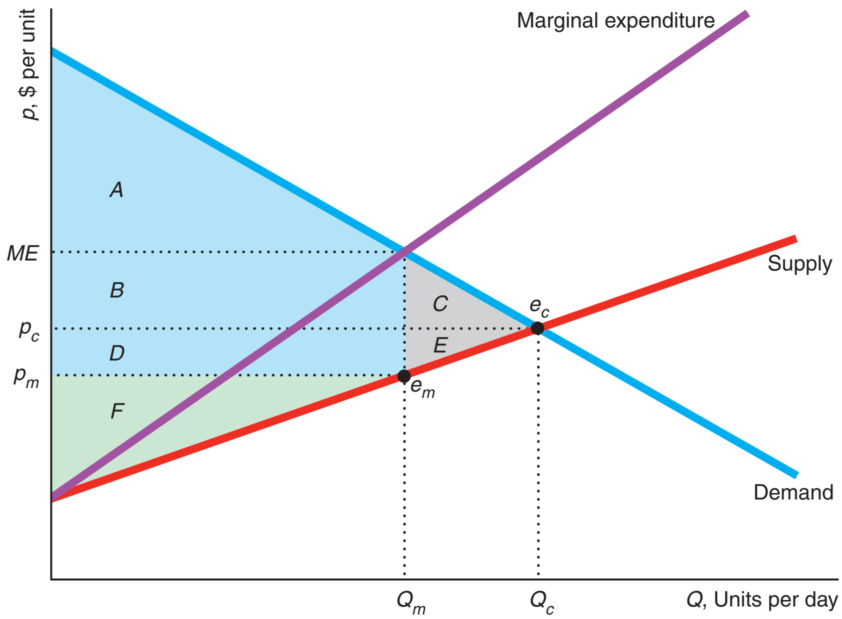
Webber (2016) estimated that being married or having children present reduces a firm's labor-supply elasticity for women (but does not affect men), which explains roughly 60% of the gender elasticity differential. The total difference in elasticities explains 3.3% lower earnings for women, or about 14% of the gender earnings gap.

Welfare Effects of Monopsony

By creating a wedge between the value to the monopsony and the value to the suppliers, the monopsony causes a welfare loss in comparison to a competitive market. If the market in Figure 15.10 is competitive, in the competitive equilibrium, e_c, the quantity is Q_c and the price is p_c. However, at a monopsony optimum, e_m, the quantity is Q_m and the price is p_m.

Figure 15.10 Welfare Effects of Monopsony

By setting a price, p_m, below the competitive level, p_c, a monopsony causes too little to be sold by the supplying market, reducing welfare.



	Competition	Monopsony	Change
Consumer Surplus, *CS*	$A + B + C$	$A + B + D$	$D - C = \Delta CS$
Producer Surplus, *PS*	$D + E + F$	F	$-D - E = \Delta PS$
Welfare, $W = CS + PS$	$A + B + C + D + E + F$	$A + B + D + F$	$-C - E = \Delta W = DWL$

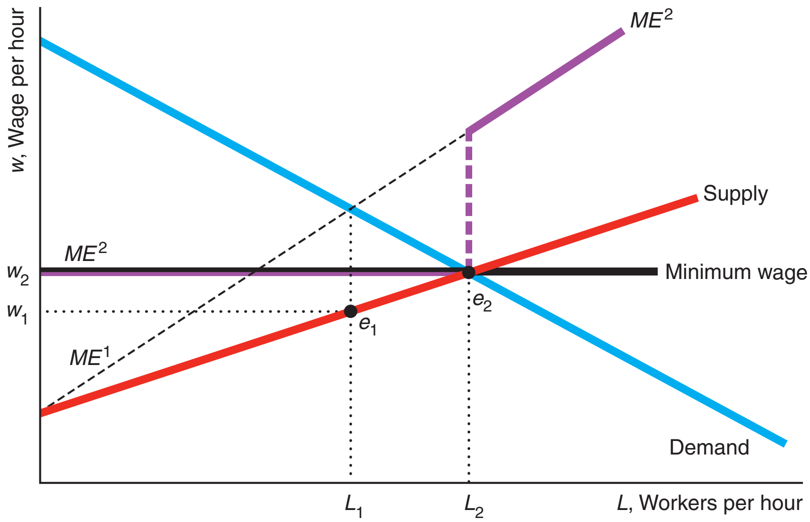
Compared to the competitive equilibrium, sellers lose producer surplus, $D + E$, because the monopsony price, p_m, for the good is below the competitive price, p_c. Area D is a transfer from the sellers to the monopsony and represents the savings of $p_c - p_m$ on the Q_m units the monopsony buys. The monopsony loses C because suppliers sell it less output, Q_m instead of Q_c, at the low price. Thus, the deadweight loss of monopsony is area $C + E$. This loss is due to the wedge between the value the monopsony places on the Q_m units, the monopoly expenditure ME in the figure, and the price it pays, p_m. The greater the difference between Q_c and Q_m and the larger the gap between ME and p_m, the greater the deadweight loss.

Solved Problem 15.4

How does the equilibrium in a labor market with a monopsony employer change if a minimum wage is set at the competitive level?

Answer

1. *Determine the original monopsony equilibrium.* Given the supply curve in the graph, the marginal expenditure curve is ME^1. The intersection of ME^1 and the demand curve determines the monopsony equilibrium, e_1. The monopsony hires L_1 workers at a wage of w_1.



2. *Determine the effect of the minimum wage on the marginal expenditure curve.* The minimum wage makes the supply curve, as viewed by the monopsony, flat in the range where the minimum wage is above the original supply curve (fewer than L_2 workers). The new marginal expenditure curve, ME^2, is flat where the supply curve is flat. Where the supply curve is upward sloping, ME^2 is the same as ME^1.

3. *Determine the post-minimum-wage equilibrium.* The monopsony operates where its new marginal expenditure curve, ME^2, intersects the demand curve. With the minimum wage, the demand curve crosses the ME^2 curve at the end of the flat section. Thus, at the new equilibrium, e_2, the monopsony pays the minimum wage, w_2, and employs L_2 workers.

4. *Compare the equilibria.* The post-minimum-wage equilibrium is the same as the competitive equilibrium determined by the intersection of the supply and demand curves. Workers receive a higher wage, and more are employed than in the monopsony equilibrium. The minimum wage helps workers and hurts the monopsony.

Challenge Solution

Athletes' Salaries and Ticket Prices

In the Challenge at the beginning of the chapter, we asked whether, when a baseball team re-signs its star player at a much higher salary, it would raise the ticket price as a result. We've seen that an increase in the wage does lead to a higher competitive market price. Is the sports example different than the competitive market, and if so, why?

When a sports team re-signs a star player at a higher salary, it does not raise its ticket price. The sports case differs in two major ways. First, a sports team can set prices—it is a local monopoly (or at least a member of an oligopoly). Second, the team pays a player a fixed salary for the year, which does not vary with the number of fans in the stadium.

A monopolistic baseball club sets its ticket price such that its marginal revenue curve intersects its marginal cost curve. When a team raises a star player's salary, it increases its fixed cost but not its marginal cost. The player's salary doesn't affect the cost of bringing one more fan to the stadium. Indeed, if a stadium has unfilled seats, the marginal cost of the last fan is essentially zero.

Thus, to maximize its profit, the firm should set its price to maximize its revenue. That is, if the team starts paying a higher salary to a current player it should not raise its ticket price, because the profit-maximizing price is unchanged by a shift in a fixed cost.[13]

If salaries determined ticket prices, then we would expect changes in salaries to be highly correlated with ticket prices, but the correlation is not strong. The share of team revenue that went to salaries rose from 25% in 1976 to 54% in 2006, but fell below 40% by 2012.

Between 1990 and 2005, the average player salary increased 100%, while the average baseball ticket price rose 120%. In 2010, of the 14 teams that reduced their payroll, 4 lowered their average ticket price, 7 kept their price unchanged, and 3 raised their price. Of the 16 teams that had a higher payroll, 2 lowered the average ticket price (for example, the Tigers' payroll increased by 6.8%, but its average ticket price fell by 14.2%), 6 kept their price unchanged, and 8 raised the price. Of the six teams that raised their prices, one had its payroll fall—the Rangers raised the average ticket price 6.4% even though the team's payroll fell 19%—and another's payroll was essentially unchanged.

The main reason for a correlation between player payrolls and ticket prices is that higher ticket prices "cause" higher salaries rather than the other way around. Star players capture some of the team's profit. Indeed, the Dodgers raised their ticket prices before they signed Brown to a high salary. Teams in cities where they can earn the highest revenues tend to pay the highest salaries. The New York Yankees' YES cable television network pays them about $350 million in revenue per year (plus the team owns a third of the company), which more than covers the Yankees' payroll ($229 million in 2013) before a single fan entered Yankee Stadium. The five other highest payrolls are paid by teams with new, large stadiums that double as virtual cash registers and provide funds to hire players.

Summary

1. **Competitive Factor Market.** Any firm maximizes its profit by choosing the quantity of a factor such that the marginal revenue product (*MRP*) of that factor—the marginal revenue times the marginal product of the factor—equals the factor price. The *MRP* is the firm's factor demand. A competitive firm's marginal revenue is the market price, so its *MRP* is the market price times the marginal product. The firm's long-run factor demand is usually flatter than its short-run demand because it can adjust more factors, thus giving it more flexibility. The market demand for a factor reflects how changes in factor prices affect output prices and hence output levels in product markets.

2. **Effects of Monopolies on Factor Markets.** If firms exercise market power to raise price above marginal cost in an output market or factor market, the quantity demanded by consumers falls. Because the quantity of output and the quantity of inputs are closely related, a reduction in the quantity of an input reduces output, and a reduction in output reduces the demand for inputs.

3. **Monopsony.** A profit-maximizing monopsony—a single buyer—sets its price so that the marginal value to the monopsony equals its marginal expenditure. Because the monopsony pays a price below the competitive level, fewer units are sold than in a competitive market, producers of factors are worse off, the monopsony earns higher profits than it would if it were a price taker, and society suffers a deadweight loss. A monopsony may also price discriminate.

[13]In contrast, if by hiring a new star player (rather than re-signing one) the team attracts more fans, the demand curve it faces may shift so that it pays for the team to raise ticket prices. However, the team raises its price due to the shift in the demand curve rather than for cost reasons.

Questions

*Select questions are available on MyLab Economics; * = answer appears at the back of this book;* **A** *= algebra problem;* **C** *= calculus problem.*

1. Competitive Factor Market

1.1 Explain how a competitive firm selects the amount of labor it employs. Assume that the labor market is also competitive.

*1.2 A firm's production function is Cobb-Douglas: $q = AL^aK^b$. What is the firm's marginal revenue product of labor? (*Hint*: Use Appendix 6C and see Solved Problem 15.1.) **A**

1.3 The Cobb-Douglas production function for a U.S. tobacco products firm is $q = L^{0.2}K^{0.3}$ (see the Application "Returns to Scale in Various Industries" in Chapter 6). Derive the marginal revenue product of labor for this firm. (*Hint*: Use Appendix 6C and see Solved Problem 15.1.) **A**

*1.4 A competitive firm's production function is $q = 2LK$. What is its marginal revenue product of labor? (*Hint*: $MP_L = 2K$ and see Solved Problem 15.1.) **A**

*1.5 If the marginal product of labor is $MP_L = 9 - L$, the market price is $p = 10$ and the wage rate is $w = 10$, how many units of labor will a competitive firm hire? How does your answer change if the government imposes a specific tax of 5? (*Hint*: See Solved Problem 15.2.)

1.6 How does a rise in the rental price of capital affect a firm's demand for labor in the long run?

1.7 Suppose that competitive firms use a fixed-proportions production process to assemble bicycles using one frame and two tires. Each frame costs p_F and each tyre costs p_T. What is the market demand for bicycle frames?

1.8 Suppose that a modern plague (AIDS, SARS, Ebola virus, avian flu) wipes out or incapacitates a major share of a small country's work force. If this country's labor market is competitive, what effect will this disaster have on wages in this country?

1.9 In July 2016, *ABC Rural* reported that a more environmentally friendly and less costly water-jet technology is being developed to compete with hydraulic fracturing in which water, sand, and chemicals are pumped under high pressure into geological structures to fracture the rock and increase the volumes of oil and gas that can be recovered. In terms of a marginal revenue product and marginal cost diagram, how might improvements in drilling technology result in more oil and gas wells drilled?

1.10 Suppose that a firm's production function is $q = L + K$. Can it be a competitive firm? Explain. **A**

2. Effects of Monopolies on Factor Markets

2.1 How does a monopoly's demand for labor change if a second firm enters its output market and the result is a Stackelberg duopoly equilibrium, where the former monopoly becomes the Stackelberg leader? Assume the inverse market demand curve is $p = 50 - 2Q$, and the firms have constant and identical marginal costs, $MC_1 = MC_2 = 2$. (*Hint*: See Equation 13.7.)

2.2 The inverse market demand curve for a final good is $P = 50 - Q$ and the marginal cost of supplying labor is $MC_L = 20$. Each unit of output requires half a unit of labor, L, and no other factor, $Q = 2L$. If factor and output markets are competitive, what are the equilibria in both markets? How does your answer change if the factor market is a monopoly but the output market is competitive?

2.3 If a monopoly has a Cobb-Douglas production function, $Q = L^aK^b$, and faces an inverse demand function of $p = Q^{-d}$, what is its marginal revenue product of labor? (*Hint*: Use Appendix 6C, and note that the monopoly's marginal revenue function is $MR = [1 - d]Q^{-d} = [1 - d]p$.) **A**

2.4 The inverse market demand curve for a final good is $P = 50 - Q$ and the wage rate is $w = 20$. Each unit of output requires half a unit of labor, L, and no other factor, $Q = 2L$. If factor and output markets are competitive, what are the equilibria in both factor and output markets? How does your answer change if the factor market is competitive but the output market is a monopoly?

2.5 The sale of commercial rights for the matches in the qualifying competitions of the 2016 European Football Championship and the 2018 FIFA World Cup was centralized under the Union of European Football Associations (UEFA). Instead of members securing their own deals individually, UEFA contracted on behalf of all of its 53 member states. Media rights were awarded to various broadcasters in different countries. In the United Kingdom, for example, ITV secured the exclusive rights to televise all of England's qualifying games in a £100 million deal. What would be the expected effect of centralizing sales with UEFA on advertising rates and the number of commercials, and why?

2.6 The inverse market demand curve for a final good is $p = 50 - Q$ and the wage rate is $w = 20$. Each unit of output requires half a unit of labor, L, and no other factor, $Q = 2L$. If factor and output markets are competitive, what are the equilibria in both factor and output markets? How does your answer change if both the output and factor markets are a monopoly?

2.7 Can a merger of an upstream and a downstream monopoly help consumers? Explain. (*Hint*: See Solved Problem 15.3.)

2.8 For the first five years after Apple introduced the iPhone, Apple sold it in the United States with the requirement that it be used only on the AT&T cell phone network. Indeed, Apple took a series of steps to prevent customers from "unlocking" the phone so that it could be used on other networks. The Orange network in France began selling the first iPhone for €399 ($588) with a two-year subscription. Unlike in the United States, one could get an unlocked iPhone in France from the vendor. Orange would unlock an iPhone for an additional €100 ($144) if the customer chose an iPhone service plan, €150 if the customer stayed with the carrier and had a non-iPhone plan (which didn't allow one to use the iPhone's special features), and €250 if the customer didn't have a plan with Orange (Stan Beer, "Orange iPhone Unlock Starts Demise of Exclusive Carrier Model," *ITWire*, November 28, 2007). Give plausible explanations why Apple chose to have an exclusive deal with AT&T, why AT&T wanted Apple to enforce exclusivity, and why Orange was being more flexible. Was Apple or the phone service "extending monopoly power?"

2.9 Apple sold its iPhone to AT&T, which in turn sold it to the final consumers. Suppose that the consumers' constant elasticity demand function for the iPhone was $Q = Ap^{-}$, Apple's marginal cost of production was m, and AT&T's marginal cost of reselling the phone was c. If both Apple and AT&T were monopolies and set prices independently, what price would they set? If they were to have merged, what price would they have set? (*Hint*: See Solved Problem 15.3.) **C**

2.10 In August 2011, the United Kingdom's big supermarkets and certain dairy processors were found to have colluded to increase the prices consumers paid for milk, butter, and cheese in 2002 and 2003. Suppose that the retail price supermarkets charge for these products is 25% higher than the price they pay the dairy processors. If the dairy processors were to form a cartel, compare the market price where they sell directly to consumers to the market price where consumers buy through supermarkets. For this purpose, assume the marginal cost of producing a dairy product is constant.

3. Monopsony

3.1 Can a monopsony exercise monopsony power—profitably setting its price below the competitive level—if the supply curve it faces is horizontal?

3.2 A monopsony faces a supply curve of $p = 10 + Q$. What is its marginal expenditure curve? **A**

3.3 Suppose that the original labor supply curve, S^1, for a monopsony shifts to the right to S^2 if the firm spends $1,000 in advertising. Under what condition should the monopsony engage in this advertising? (*Hint*: See the monopoly advertising analysis in Chapter 12.)

3.4 Use a diagram of a competitive labor market to compare the equilibrium effects and incidence of a per-unit tax on earnings when the buyer of the labor services is competitive compared to when it is a monopsonist.

3.5 Suppose that a modern plague (AIDS, SARS, Ebola virus, avian flu) wipes out or incapacitates a major share of a small country's work force. If this country's labor market is monopsonistic, what effect will this disaster have on wages in this country? Compare your answer to that in Question 1.8.

3.6 A firm is a monopoly in the output market and a monopsony in the input market. Its only input is the finished good, which it buys from a competitive market with an upward-sloping supply curve. The firm sells the same good to competitive buyers in the output market. Determine its profit-maximizing output. What price does it charge in the output market? What price does it pay to its suppliers?

3.7 Compare the equilibrium in a market in which a firm is both a monopoly and a monopsony (as in the previous question) to the competitive equilibrium.

3.8 Compare the equilibrium quantity and price in two markets: one in which a firm is both a monopsony and a monopoly (as in Question 3.6) and one in which the firm buys inputs competitively but has a monopoly in the output market.

3.9 If the monopsony faces a supply curve of $p = 10 + Q$ and has a demand curve of $p = 50 - Q$, what are the equilibrium quantity and price? How does this equilibrium differ from the competitive equilibrium? **A**

3.10 Compare welfare in a market where a firm is both a monopsony and a monopoly (as in Question 3.6) to welfare in markets in which the firm has a monopsony in the input market but acts as a price taker in the output market.

3.11 What happens to the monopsony equilibrium if the minimum wage is set slightly above or below the competitive wage? (*Hint*: See Solved Problem 15.4.)

3.12 What effect does a price support have on a monopsony? In particular, describe the equilibrium if the price support is set at the price where the supply curve intersects the demand curve. (*Hint*: See Solved Problem 15.4.)

3.13 Some employment contracts contain a non-compete clause that prohibits an employee from engaging in certain activities for its company's competitors for a certain amount of time after the employment agreement has come to an end. Sometimes one competitor sues another when the latter hires the former's ex-employee in violation of a covenant not to compete. At other times, competitors enter into agreements not to hire each other's employees. However, mutual no-hire agreements run afoul of antitrust laws. What would be the purpose of a no-hire agreement among competitors? What type of market structure would these firms be trying to achieve?

3.14 Humphreys and Pyun (2016) estimated that $(MRP - \text{salary})/MRP = 0.89$ for rookie major league baseball players, while the comparable figure for players who are eligible to be free agents is 0.21. Explain the reason for this difference.

4. Challenge

4.1 The Challenge points out that if a ball club raises a player's salary, it increases its fixed cost but not its variable cost. Use a figure to show what effect such an increase has if (a) the firm is competitive or (b) the firm is a monopoly.

Interest Rates, Investments, and Capital Markets

16

I'd gladly pay you Tuesday for a hamburger today. —Wimpy

Challenge

Does Going to College Pay?

For most of your childhood, your parents, teachers, or other adults urged you to go to college. According to a 2015 Gallup poll, virtually all (96%) of U.S. adults believe that it is somewhat or very important to have a degree or professional certificate beyond a high school diploma.

However, is college worth the cost? U.S. adults who considered college to be a good investment plummeted from 81% in 2008 to 50% in 2015. A 2015 survey found that 60% of people between the ages of 25 and 44 who had finished high school but lacked a college degree believed that "a college education is worth the cost." Another 2015 survey found that three-quarters of college students believed college was a good investment.

In the fall of 2015, over two-thirds (69%) of 2015 high school graduates were enrolled in colleges or universities. Enrollment in U.S. colleges and universities rose by almost one third between fall 2000 and fall 2015. The share of adults over 25 who have completed a bachelor's or higher degree increased from 23% in 1990 to 36% in 2015.

Going to college is expensive. In the 2014–2015 school year, the average total cost of tuition, fees, room, and board was \$21,728 across all institutions of higher education and \$41,970 at private nonprofit schools. In 2015, *U.S. News & World Report* claimed that college grads earn \$1 million more than people without college degrees over their lifetimes. How should we weigh the costs and benefits to determine if an investment in a college education pays financially?

This chapter continues the treatment of factor markets in Chapter 15 by focusing on capital markets and examining which investments pay. People invest in capital and other durable goods: products that are usable for years. Firms use durable goods—such as manufacturing plants, machines, and trucks—to produce and distribute goods and services. (By comparison, nondurable goods and services, such as ice-cream cones or seeing a movie in a theater, are consumed or used once at the time of payment or soon thereafter.)

Consumers spend one in every eight of their dollars on durable goods such as houses, cars, and refrigerators. Just as a firm considers whether or not to invest in physical capital, individuals decide whether to invest in their own *human capital*. Where a firm chooses the durability of a piece of equipment, some people invest in lengthening their expected life spans by exercising or purchasing medical care. Where a firm buys machinery and other capital to produce more output and increase its future profits, individuals invest in education to raise their productivity and their future earnings.

If a firm rents a durable good by the week, it faces a decision similar to buying a nondurable good or service. A firm demands workers' services (or other nondurable input) up to the point at which its *current* marginal cost (the wage) equals its *current* marginal benefit (the marginal revenue product of the workers' services). A firm that rents a durable good, such as a truck, by the month can use the same rule to decide how many trucks to employ per month. The firm rents trucks up to the point at which the *current* marginal rental cost equals its *current* marginal benefit—the marginal revenue product of the trucks.

If the capital good must be bought or built rather than rented, the firm cannot apply this rule on the basis of current costs and benefits alone. (Many types of capital, such as factories or specialized pieces of equipment, *cannot* be rented.) In deciding whether to build a long-lived factory, a firm must compare the *current* cost of the capital to the *future* higher profits it will make from using the plant.

stock
a quantity or value that is measured independently of time

Often such comparisons involve *stocks* and *flows*. A **stock** is a quantity or value that is measured independently of time. Because a durable good lasts for many periods, its stock is discussed without reference to its use within a particular period. We say that a firm owns "an apartment building *this* year" (not "an apartment building *per* year"). If a firm buys the apartment building for $500,000, we say that it has a capital stock worth $500,000 today.

flow
a quantity or value that is measured per unit of time

A **flow** is a quantity or value that is measured per unit of time. The consumption of nondurable goods, such as the number of ice-cream cones you eat per week, is a flow. Similarly, the stock of a durable good provides a flow of services. A firm's apartment building—its capital stock—provides a flow of housing services (apartments rented per month or year) to tenants. In exchange for these housing services, the firm receives a flow of rental payments from the tenants.

Does it pay for the firm to buy the apartment building? To answer this question, we need to extend our analysis in two ways. First, we must develop a method of comparing a flow of dollars in the future to a dollar today, as we do in this chapter. Second, we need to consider the role of uncertainty about the future (can the firm rent all the apartments each month?), a subject that we discuss in Chapter 17.

In this chapter, we examine four main topics

1. **Comparing Money Today to Money in the Future.** Interest rates tell us how much more money is worth today than in the future.
2. **Choices over Time.** Investing money in a project pays if the return from that investment is greater than that on the best alternative when both returns are expressed on a comparable basis.
3. **Exhaustible Resources.** Scarcity, rising costs of extraction, and positive interest rates may cause the price of exhaustible resources like coal and gold to rise exponentially over time.
4. **Capital Markets, Interest Rates, and Investments.** Supply and demand in capital markets determine the market interest rate, which affects how much people invest.

16.1 Comparing Money Today to Money in the Future

Even without inflation—so a bundle of goods would sell for the same price today, next year, and 100 years from now—most people would still value receiving a dollar today more than a dollar to be received tomorrow. Wouldn't you rather eat a dollar's worth of chocolate today than wait ten years to eat that same amount of chocolate?

Interest Rates

interest rate
the percentage more that must be repaid to borrow money for a fixed period of time

Because virtually everyone values having a dollar today more than having a dollar in the future, getting someone to loan you a dollar today requires agreeing to pay back more than a dollar in the future. You may have borrowed money to pay for your college education in exchange for a credible promise to repay a greater amount after you graduate. How much more you must pay in the future is specified by an **interest rate**: the percentage more that must be repaid to borrow money for a fixed period of time.[1]

If you put money in a savings account, you are lending the bank your money, which it may in turn loan to someone who wants to buy a car or a house. For the use of your deposited funds for one year, the bank agrees to pay you an interest rate, i, of, say, 4%. That is, the bank promises to return to you \$1.04 ($= 1 + i$) one year from now for every dollar you loan it. If you put \$100 in your savings account, you will have your \$100 plus interest of $\$100 \times 0.04 = \$4$ for a total of \$104 at the end of the year. (See MyLab Economics, Chapter 16, "Usury," for a discussion of ancient people's opposition to paying interest, and current restrictions on Islamic banks.)

discount rate
a rate reflecting the relative value an individual places on future consumption compared to current consumption

Discount Rate You may value future consumption more or less than do other members of society. If you knew you had two years to live, you would place less value on payments three or more years in the future than most other people do. We call an individual's personal "interest" rate that person's **discount rate**: a rate reflecting the relative value an individual places on future consumption compared to current consumption.

A person's willingness to borrow or lend depends on whether his or her discount rate is greater or less than the market interest rate. If your discount rate is nearly zero—you view current and future consumption as equally desirable—you would gladly loan money in exchange for a positive interest rate. Similarly, if your discount rate is high—current consumption is much more valuable to you than future consumption—you would be willing to borrow at a lower interest rate. In the following discussion, we assume for simplicity that an individual's discount rate is the same as the market interest rate unless we explicitly state otherwise.

Compounding If you place \$100 in a bank account that pays 4%, at the end of a year, you can take out the interest payment of \$4 and leave your \$100 in the bank to earn more interest in the future. If you leave your \$100 in the bank indefinitely and the interest rate remains constant over time, you will receive a payment of \$4 each year. In this way, you can convert your \$100 stock into a flow of \$4-a-year payments forever.

In contrast, if you leave both your \$100 and your \$4 interest payment in the bank, the bank must pay you interest on \$104 at end of the second year. The bank owes you interest of \$4 on your original deposit of \$100 and interest of $\$4 \times 0.04 = \$0.16$ on your interest from the first year, for a total of \$4.16.

Thus, at the end of Year 1, your account contains

$$\$104.00 = \$100 \times 1.04 = \$100 \times 1.04^1.$$

[1]For simplicity, we refer to *the* interest rate throughout this chapter, but most economies have many interest rates. For example, a bank charges a higher interest rate to loan you money than it pays you to borrow your money.

By the end of Year 2, you have

$$\$108.16 = \$104 \times 1.04 = \$100 \times 1.04^2.$$

At the end of Year 3, your account has

$$\$112.49 \approx \$108.16 \times 1.04 = \$100 \times 1.04^3.$$

If we extend this reasoning, by the end of Year t, you have

$$\$100 \times 1.04^t.$$

In general, if you let your interest accumulate in your account, for every dollar you loan the bank, it owes you $1 + i$ dollars after one year, $(1 + i) \times (1 + i) = (1 + i)^2$ dollars after two years, $(1 + i) \times (1 + i) \times (1 + i) = (1 + i)^3$ after three years, and $(1 + i)^t$ dollars at the end of t years. This accumulation of interest on interest is called *compounding*.

Frequency of Compounding To get the highest return on your savings account, you need to check both the interest rate and the frequency of compounding. We have assumed that interest is paid only at the end of the year. However, many banks pay interest more frequently than once a year. If you leave your interest in the bank for the entire year, you receive compounded interest—interest on the interest.

If a bank's annual interest rate is $i = 4\%$, but it pays interest two times a year, the bank pays you half a year's interest, $i/2 = 2\%$, after six months. For every dollar in your account, the bank pays you $(1 + i/2) = 1.02$ dollars after six months. If you leave the interest in the bank, at the end of the year, the bank must pay you interest on your original dollar and on the interest you received at the end of the first six months. At the end of the year, the bank owes you $(1 + i/2) \times (1 + i/2) = (1 + i/2)^2 = (1.02)^2 = \$1.0404$, which is your original \$1 plus 4.04¢ in interest.

If the bank were to compound your money more frequently, you would earn even more interest. Some banks offer continuous compounding, paying interest at every instant. Such compounding is only slightly better for you than daily compounding. Table 16.1 shows that the amount you would earn after one year of investing \$10,000 at either a 4% or an 18% annual interest rate depends on the frequency of compounding.

Because most people cannot easily perform such calculations, the 1968 U.S. Truth-in-Lending Act requires lenders to tell borrowers the equivalent noncompounded annual percentage rate (APR) of interest. As the table shows, twice-a-year compounding at 4% has an APR of 4.04%. That is, over a year, an account with a

Table 16.1 Interest and the Frequency of Compounding

	Interest Payments on a \$10,000 Investment at the End of 1 Year, \$	
Frequency of Compounding	**4%**	**18%**
Once a year	400.00	1,800.00
Twice a year	404.00	1,881.00
Four times a year	406.04	1,925.19
Daily	408.08	1,971.64
Continuous	408.11	1,972.17

noncompounded interest rate of 4.04% pays you the same interest as a 4% account that was compounded twice during the year.

Thus, when considering various loans or interest rates, you should compare the APRs; comparing rates that are compounded at different frequencies can be misleading. If you use credit cards to borrow money, it's particularly important that you compare APRs across accounts because credit card interest rates are usually high. If the interest rate on your card is 18%, a continuously compounded rate has an APR of over 19.7%. If you borrow $10,000 for a year, you'll owe $1,972.17 with continuous compounding, which is 9.6% more than the $1,800 you'd owe with annual compounding. From now on, we assume that compounding takes place annually.

Using Interest Rates to Connect the Present and Future

Interest rates connect the value of the money you put in the bank today, the *present value* (PV), with the *future value* (FV) that you are later repaid, which is the present value plus interest. Understanding this relationship allows us to evaluate the attractiveness of investments involving payments today for profits in the future and of purchases made today but paid for later. Knowing the interest rate and the present value allows us to calculate the future value. Similarly, we can determine the present value if we know the future value and the interest rate.

Future Value If you deposit PV dollars in the bank today and allow the interest to compound for t years, how much money will you have at the end? The future value, FV, is the present value times a term that reflects the compounding of the interest payments:

$$FV = PV \times (1 + i)^t. \tag{16.1}$$

Table 16.2 shows how much $1 put in the bank today will be worth in the future at various annually compounded interest rates. For example, $1 left in the bank for 50 years will be worth only $1.64 at a 1% interest rate. However, that same investment is worth $7.11 at a 4% interest rate, $117.39 at a 10% rate, and $9,100.44 at a 20% rate.

Table 16.2 Future Value, FV, to Which $1 Grows by the End of Year t at Various Interest Rates, i, Compounded Annually, $

t, Years	1%	4%	5%	10%	20%
1	1.01	1.04	1.05	1.10	1.20
5	1.05	1.22	1.28	1.61	2.49
10	1.10	1.48	1.63	2.59	6.19
25	1.28	2.67	3.39	10.83	95.40
50	1.64	7.11	11.47	117.39	9,100.44

Note: $FV = (1 + i)^t$, where FV is the future value of $1 invested for t years at an annual interest rate of i.

Application

Power of Compounding

One thousand dollars left to earn interest at 8% a year will grow to $43 quadrillion in 400 years, but the first 100 years are the hardest —Sidney Homer, Salomon Brothers analyst

No doubt you've read that the Dutch got a good deal buying Manhattan from the original inhabitants in 1626 for about \$24 worth of beads and trinkets. However, that conclusion may be wrong. If these Native Americans had had the opportunity to sell the beads and invest in tax-free bonds with an APR of 7%, the bond would be worth \$6.9 trillion in 2016, which is much more than the assessed value of Manhattan Island of \$1.1 trillion. On the other hand, if the United States had taken the \$7.2 million it paid for the purchase of Alaska from Russia in 1867 and invested in the same type of bonds, that money would be worth only \$172 billion in 2016, which is much less than Alaska's current value.

Present Value Instead of asking how much a dollar today is worth in the future, we can ask how much a dollar in the future is worth today, given the market interest rate. For example, we may want to know how much money, PV, we have to put in the bank today at an interest rate of i to get a specific amount of money, FV, in the future. If we want to have $FV = \$100$ at the end of a year and the interest rate is $i = 4\%$, then from Equation 16.1 we know that $PV \times 1.04 = \$100$. Dividing both sides of this expression by 1.04, we learn that we need to put $PV = \$100/1.04 = \$96.15$ in the bank today to have \$100 next year.

A more general formula relating money t periods in the future to money today is obtained by dividing both sides of Equation 16.1 by $(1 + i)^t$ to obtain

$$PV = \frac{FV}{(1 + i)^t}. \tag{16.2}$$

This equation tells us what FV dollars in year t are worth today at an interest rate of i compounded annually. Table 16.3 and Figure 16.1 show what \$1 in the future is worth today at various interest rates. At high interest rates, money in the future is virtually worthless today: A dollar paid to you in 25 years is worth only 1¢ today at a 20% interest rate.

Stream of Payments

Sometimes we need to deal with payments per period, which are flow measures, rather than a present value or future value, which are stock measures. Often a firm pays for a new factory or an individual pays for a house by making monthly mortgage

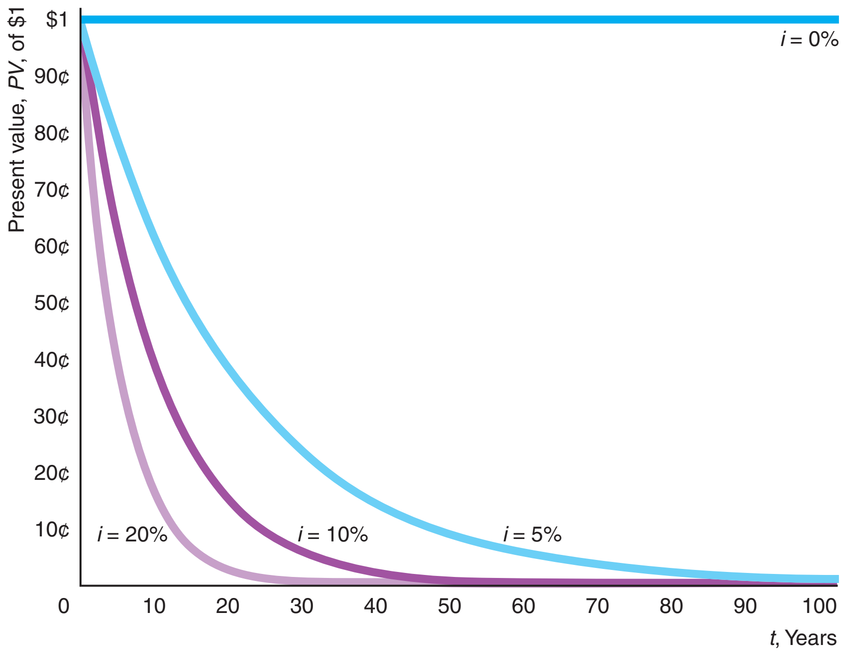
Table 16.3 Present Value, *PV*, of a Payment of \$1 at the End of Year *t* at Various Interest Rates, *i*, Compounded Annually, \$

t, Years	1%	4%	5%	10%	20%
1	0.99	0.96	0.95	0.91	0.83
5	0.95	0.82	0.78	0.62	0.40
10	0.91	0.68	0.61	0.39	0.16
25	0.78	0.38	0.30	0.09	0.01
50	0.61	0.14	0.09	0.009	0.00011

Note: $PV = 1/(1 + i)^t$, where PV is the present value of \$1 at the end of year t at an annual interest rate of i.

Figure 16.1 Present Value of a Dollar in the Future

The present value of a dollar is lower the farther in the future it is paid. At a given time in the future, the present value is lower when the interest rate is higher.



payments. In deciding whether to purchase the factory or house, the decision maker compares the value of the stock (factory or home) to a flow of payments over time.

Present Value of Payments over Time One way to make such an evaluation is to use our knowledge of the relationship between present and future value to determine the present value of the stream of payments. To do so, we calculate the present value of each future payment and sum them.

Payments for a Finite Number of Years We start with a specific example. Suppose that you agree to pay $10 at the end of each year for three years to repay a debt. If the interest rate is 10%, the present value of this series of payments is

$$PV = \frac{\$10}{1.1} + \frac{\$10}{1.1^2} + \frac{\$10}{1.1^3} \approx \$24.87.$$

More generally, if you make a *future payment* of f per year for t years at an interest rate of i, the present value (stock) of this flow of payments is

$$PV = f\left[\frac{1}{(1+i)^1} + \frac{1}{(1+i)^2} + \cdots + \frac{1}{(1+i)^t}\right] = \frac{f}{i}\left[1 - \frac{1}{(1+i)^t}\right], \quad (16.3)$$

where the last equality is derived in Appendix 16A. Table 16.4 shows that the present value of a payment of $f = \$10$ a year for five years is $43 at 5%, $38 at 10%, and $30 at 20% annual interest.

Payments Forever If these payments must be made at the end of each year forever, the present value formula is easier to calculate than Equation 16.3. If you put PV dollars into a bank account earning an interest rate of i, you can get an interest or future

Table 16.4 Present Value, *PV*, of a Flow of $10 a Year for *t* Years at Various Interest Rates, *i*, Compounded Annually, $

t, Years	5%	10%	20%
5	43	38	30
10	77	61	42
50	183	99	50*
100	198	100*	50*
∞	200	100	50

*The actual numbers are a fraction of a cent below the rounded numbers in the table. For example, the *PV* at 10% for 100 years is $99.9927.

Note: The payments are made at the end of the year.

payment of $f = i \times PV$ at the end of the year. Dividing both sides of this expression by i, we find that to get a payment of f each year forever, you'd need to put

$$PV = \frac{f}{i} \tag{16.4}$$

in the bank (see Appendix 16A). Thus, you'd have to deposit $10/$i$ in the bank to ensure a future payment of $f =$ $10 forever. Using this formula, we determine that the present value of $10 per year forever is $200 at 5%, $100 at 10%, and $50 at 20%.[2]

Solved Problem 16.1

MyLab Economics Solved Problem

Melody Toyota advertises that it will sell you a Corolla for $14,000 or lease it to you. To lease it, you must make a down payment of $1,650 and agree to pay $1,800 at the end of each of the next two years. After the last lease payment, you may buy the car for $12,000. If you plan to keep the car until it falls apart (at least a decade) and the interest rate is 10%, which approach has a lower present value of costs (assuming no inflation)?

Answer

1. *Calculate the present value of leasing.* To calculate the present value of leasing, we use Equation 16.2, $PV = FV/(1 + i)^t$. The present value of leasing the car and then buying it is the sum of the down payment of $1,650, the present value of paying $f =$ $1,800 at the end of the first year ($1,800/1.1 ≈ $1,636.36), the present value of paying $f =$ $1,800 at the end of the first year ($1,800/1.1^2 ≈ $1,487.60), and the present value of purchasing the car for

[2]This payment-in-perpetuity formula, Equation 16.4, provides a good approximation of a payment for a large but finite number of years. As Table 16.4 shows, at a 5% interest rate, the present value of a payment of $10 a year for 100 years, $198, is close to the present value of a permanent stream of payments, $200. At higher interest rates, this approximation is nearly perfect. At 10%, the present value of payments for 100 years is $99.9927 compared to $100 for perpetual payments. The reason this approximation works better at high rates is that a dollar paid more than 50 or 100 years from now is essentially worthless today, as Table 16.3 shows.

$FV = \$12{,}000$ at the end of two years ($\$12{,}000/1.1^2 \approx \$9{,}917.36$). Thus, the present value of leasing the car and then buying it is approximately

$$\$1{,}650 + \$1{,}636.36 + \$1{,}487.60 + \$9{,}917.36 = \$14{,}691.32.$$

2. *Compare leasing to buying the car*. The present value of buying the car is \$14,000, which is about \$691 less than the present value of leasing it.

Future Value of Payments over Time We just calculated the present value of a stream of payments. This type of computation can help you decide whether to buy something today that you'll pay for over time. Sometimes, however, we want to know about the future value of a stream of payments.

For example, suppose that you want to know how much you'll have in your savings account, FV, at some future time if you save f each year. The first year, you place f dollars in your account. The second year, you add another f and you have the first year's payment plus its accumulated interest, $f(1 + i)^1$. Thus, at the end of the second year, your account has $f[1 + (1 + i)^1]$. In the third year, you have the third year's payment, f, plus the current value of the second year's payment, $f(1 + i)$, plus the current value of the first year's payment, $f(1 + i)^2$, for a total of $f[1 + (1 + i) + (1 + i)^2]$. Continuing in this way, we see that, at the end of t years, the account has[3]

$$FV = f[1 + (1 + i)^1 + (1 + i)^2 + \cdots + (1 + i)^{t-1}] = \frac{f}{i}[(1 + i)^t - 1]. \quad (16.5)$$

where the last equality is derived in Appendix 16B.

Application

Saving for Retirement

If all goes well, you'll live long enough to retire. Will you live like royalty off your savings, or will you depend on Social Security to provide enough income that you can avoid having to eat dog food to stay alive? (When I retire, I'm going to be a Velcro farmer.)

You almost certainly don't want to hear this, but it isn't too early to think about saving for retirement. Thanks to the power of compounding, if you start saving when you're young, you don't have to save as much per year as you would if you start saving when you're middle-aged.

Suppose that you plan to work full time from age 22 until you retire at 70 and that you can earn 7% on your retirement savings account. Let's consider two approaches to savings:

- **Early bird.** You save \$3,000 a year for the first 15 years of your working life and then let your savings accumulate interest until you retire.
- **Late bloomer.** After not saving for the first 15 years, you save \$3,000 a year for the next 33 years until retirement.

Which scenario leads to a bigger retirement nest egg? To answer this question, we calculate the future value at retirement of each of these streams of investments.

[3]From Equation 16.3, we know that the present value of a stream of payments of f for t years at an interest rate of i is $PV = (f/i)[1 - 1/(1 + i)^t]$. We convert that expression to a future value at the end of t years by multiplying by $(1 + i)^t$, which gives us the last equality in Equation 16.5.

The early bird adds \$3,000 each year for 15 years into a retirement account. Using Equation 16.5, we calculate that the account has

$$\$3{,}000(1 + 1.07^1 + 1.07^2 + \cdots + 1.07^{14}) = \frac{\$3{,}000}{0.07}(1.07^{15} - 1) \approx \$75{,}387$$

at the end of 15 years. This amount then grows as the interest compounds for the next 33 years. Using Equation 16.1, we determine that the fund grows about 9.3 times to

$$\$75{,}387.07 \times 1.07^{33} = \$703{,}010$$

by retirement.

The late bloomer makes no investments for 15 years and then invests \$3,000 a year until retirement. Again using Equation 16.5, we calculate that the funds at retirement are

$$\$3{,}000(1 + 1.07^1 + 1.07^2 + \cdots + 1.07^{32}) = \frac{\$3{,}000}{0.07}(1.07^{33} - 1) \approx \$356{,}800.$$

Thus, even though the late bloomer contributes to the account for more than twice as long as the early bird, the late bloomer has saved only about half as much at retirement. Indeed, to have roughly the same amount at retirement as the early bird, the late bloomer would have to save nearly \$6,000 a year for the next 33 years. (By the way, someone who saved \$3,000 each year for all 48 years would have \$703,010 + \$356,800 = \$1,059,810 salted away by retirement.)

Inflation and Discounting

So far, we've ignored inflation (implicitly assumed an inflation rate of zero). Now we suppose that general inflation occurs so that *nominal prices*—actual prices that are not adjusted for inflation—rise at a constant rate over time. By adjusting for this rate of inflation (Chapter 5), we can convert nominal prices to *real prices*, which are prices that are adjusted for inflation. To calculate the real present value of future payments, we adjust for inflation and use interest rates to discount future real payments.

To illustrate this process, we calculate the real present value of a payment made next year. First, we adjust for inflation so as to convert next year's nominal payment to a real amount. Then we determine the real interest rate. Finally, we use the real interest rate to convert the real future payment to a real present value.

Adjusting for Inflation Suppose that the rate of inflation is g and the nominal amount you pay next year is $\tilde{f}$ This future debt in today's dollars—the real amount you owe—is $f = \tilde{f}/(1 + g)$. If the rate of inflation is $g = 10\%$, a nominal payment of $\tilde{f}$ next year is $\tilde{f}/1.1 \approx 0.909\tilde{f}$ in today's dollars.

Nominal and Real Rates of Interest To calculate the present value of this future real payment, we discount using an interest rate. Just as we converted the future payments into real values by adjusting for inflation, we convert a nominal interest rate into a real interest rate by adjusting for inflation.

Without inflation, a dollar today is worth $1 + i$ next year, where i is the real interest rate. With an inflation rate of g, a dollar today is worth $(1 + i)(1 + g)$ nominal dollars tomorrow. If $i = 5\%$ and $g = 10\%$, a dollar today is worth $1.05 \times 1.1 = 1.155$ nominal dollars next year.

Banks pay a nominal interest rate, $\tilde{i}$, rather than a real one. If they're going to get people whose real discount rate is i to save, banks' nominal interest rate must be such that a dollar pays $(1 + i)(1 + g)$ dollars next year. Because $1 + \tilde{i} = (1 + i)(1 + g) = 1 + i + ig + g$, the nominal rate is

$$\tilde{i} = i + ig + g.$$

By rearranging this equation, we see that the real interest rate depends on the nominal interest rate and the rate of inflation:

$$i = \frac{\tilde{i} - g}{1 + g}. \tag{16.6}$$

Equation 16.6 shows that the real interest rate is less than the nominal rate in the presence of inflation.

If the inflation rate is small, the denominator of Equation 16.6, $1 + g$, is close to 1. As a result, many people approximate the real interest rate as the nominal interest rate minus the rate of inflation:

$$\tilde{i} - g.$$

If the nominal interest rate is 15.5% and the rate of inflation is 10%, the real interest rate is $(15.5\% - 10\%)/1.1 = 5\%$. The approximation to the real rate, $15.5\% - 10\% = 5.5\%$, is above the true rate by half a percentage point. The lower the rate of inflation, the closer the approximation is to the real interest rate. If the inflation rate falls to $g = 2\%$ while the nominal rate remains 15.5%, the approximation to the real rate, 13.5%, is above the real rate, 13.24%, by only slightly more than a quarter of a percentage point.

Calculating the Present Value with Inflation To obtain the present value of a payment one year from now, we discount the future real payment of $f = \tilde{f}/(1 + g)$ using the real interest rate:

$$PV = \frac{f}{1 + i} = \frac{\tilde{f}}{(1 + g)(1 + i)} = \frac{\tilde{f}}{1 + \tilde{i}}.$$

Thus, we obtain the real present value by adjusting the payment for inflation, f, and by discounting using the real interest rate, i. Equivalently, we can discount the nominal payment, $\tilde{f}$, using the nominal interest rate, $\tilde{i}$.

Suppose that you sign a contract with a store to pay \$69.30 next year for a Blu-ray player you get today. The rate of inflation is $g = 10\%$, and the real interest rate is $i = 5\%$. We calculate the real present value by converting the future payment into real dollars and by using the real interest rate to discount. Next year's nominal payment of \$69.30 is only \$69.30/1.1 = \$63 in real dollars. Discounting by the real interest rate, we find that the real present value of that payment is \$63/1.05 = \$60.

If everyone anticipates a low inflation rate, g, then, using Equation 16.6, the real interest rate is roughly $i = \tilde{i} - g$. Suppose, however, that the inflation rate turns out to be higher than the anticipated rate of g. Such unanticipated inflation helps debtors because it lowers the real cost of future payments that are set in nominal rather than real terms.

When you buy the Blu-ray player, if no one expects inflation ($g = 0$), both you and the store's owner believe that the present value of your future payment is

\$69.30/1.05 = \$66. Immediately after you make the deal, the inflation rate increases to $g = 10\%$, so the actual present value is only \$60. Thus, because of the unexpected inflation, the present value of what you owe is less than you and the store's owner anticipated.

Application

Winning the Lottery

In 2016, George Collins won a "\$12 million" Florida Lotto jackpot. He chose a lump-sum cash payout of \$9.2 million, before taxes, instead of \$12 million in 30 annual payments. By offering these options, the lottery was implicitly acknowledging that money in the future is worth less than money today.[4]

Many states boast that their lottery pays a winner \$1 million. This claim is misleading (translation: They lie through their teeth). Typically, a lottery winner gets \$50,000 a year for 20 years, which means that the winner receives 20 × \$50,000 = 1 million nominal dollars over time. However, after adjustment for inflation and discounting, the real present value of these prize payments over time is much less than \$1 million.

What is the present value of a payment of \$50,000 each year for 20 years? If the first payment is made today, its real present value is \$50,000, regardless of the inflation and interest rates. The later payments need to be adjusted for inflation and discounted to the present to be comparable to this year's payment.

If the inflation rate is 5% and the real interest rate is 4%, a \$50,000 payment next year is worth only \$45,788 ≈ \$50,000/(1.05 × 1.04) this year. Generalizing, we determine that the real present value of a dollar t years from now is

$$\frac{1}{(1.05)^t(1.04)^t}.$$

The $(1.05)^t$ term in the denominator adjusts for inflation between now and the year t: It expresses the payment in the future in terms of today's dollars. The $(1.04)^t$ term in the denominator converts the payment in year t to a present value.

At these rates, the real present value of the 20 payments is less than half of the alleged value: \$491,396. Without inflation ($g = 0$), the real present value would be \$706,697. With 5% inflation and a real interest rate of 10%, the present value of the prize is only \$351,708.

16.2 Choices over Time

Earlier chapters discuss how consumers and firms make choices that do not involve time. Often, however, such decisions involve comparisons over time. Individuals and firms must choose between two or more options—such as investments and contracts—that have different present and future values. A land speculator decides whether to sell a plot of land today for \$100,000 or next year for \$200,000. Margi decides among putting \$1,000 into a bank account, buying \$1,000 worth of stocks, paying \$1,000 for a course in computer programming, and consuming the \$1,000 now. MGM, a conglomerate, decides whether to produce a movie that stars a muscle-bound hero who solves the pollution problem by beating up an evil capitalist, to

[4]Sheila Botelho was asked why she chose the single payment option after winning Rhode Island's Multi-State Powerball lottery. Mrs. Botelho and her husband responded, "At our age, we don't even buy green bananas."

build a new hotel in Reno, to buy a television studio, or to put money in a long-term savings account.

One way to make a choice involving time is to *pick the option with the highest present value*. By borrowing or lending at the market interest rate, we can shift wealth from one period to another. Thus, if we choose the option that has the highest present value, we can shift our wealth between periods so that we have more money in every period than we'd have if we made a less attractive choice.

Investing

Investment decisions may be made by comparing present values. *A firm makes an investment if the expected return from the investment is greater than the opportunity cost* (Chapter 7). The opportunity cost is the best alternative use of its money, which is what it would earn in the next best use of the money.

Thus, to decide whether to make an investment, the firm needs to compare the potential outlay of money to the firm's best alternative. One possibility is that its best alternative is to put the money that it would otherwise spend on this investment in an interest-bearing bank account. We consider two methods for making this comparison: the *net present value* approach and the *internal rate of return* approach.

Net Present Value Approach A firm has to decide whether to buy a truck for \$20,000. Because the opportunity cost is \$20,000, the firm should make the investment only if the present value of expected future returns from the truck is greater than \$20,000.

More generally, *a firm should make an investment only if the present value of the expected return exceeds the present value of the costs*. If R is the present value of the expected returns to an investment and C is the present value of the costs of the investment, the firm should make the investment if $R > C$.[5]

This rule is often restated in terms of the net present value, $NPV = R - C$, which is the difference between the present value of the returns, R, and the present value of the costs, C. *A firm should make an investment only if the net present value is positive*:

$$NPV = R - C > 0.$$

Assume that the initial year is $t = 0$, the firm's revenue in year t is R_t, and its cost in year t is C_t. If the last year in which either revenue or cost is nonzero is T, the net present value rule holds that the firm should invest if

$$\begin{aligned} NPV &= R - C \\ &= \left[R_0 + \frac{R_1}{(1 + i)^1} + \frac{R_2}{(1 + i)^2} + \cdots + \frac{R_T}{(1 + i)^T}\right] \\ &\quad - \left[C_0 + \frac{C_1}{(1 + i)^1} + \frac{C_2}{(1 + i)^2} + \cdots + \frac{C_T}{(1 + i)^T}\right] > 0. \end{aligned}$$

Instead of comparing the present values of the returns and costs, we can examine whether the present value of the *cash flow* in each year (loosely, the annual *profit*), $\pi_t = R_t - C_t$, is positive. By rearranging the terms in the previous expression, we can rewrite the net present value rule as

[5] This rule holds when future costs and returns are known with certainty and investments can be reversed but cannot be delayed (Dixit and Pindyck, 1994).

$$NPV = (R_0 - C_0) + \frac{R_1 - C_1}{(1+i)^1} + \frac{R_2 - C_2}{(1+i)^2} + \cdots + \frac{R_T - C_T}{(1+i)^T}$$

$$= \pi_0 + \frac{\pi_1}{(1+i)^1} + \frac{\pi_2}{(1+i)^2} + \cdots + \frac{\pi_T}{(1+i)^T} > 0. \qquad (16.7)$$

This rule does not restrict the firm to making investments only where its cash flow is positive each year. For example, a firm buys a piece of equipment for \$100 and spends the first year learning how to use it, so it makes no revenues from the machine and has a negative cash flow that year: $\pi_0 = -100$. The next year, its revenue is \$350 and the machine's maintenance cost is \$50, so its second year's cash flow is $\pi_1 = \$300$. At the end of that year, the machine wears out, so the annual cash flow from this investment is zero thereafter. Setting the interest rate at 5% in Equation 16.7, we learn that the firm's net present value is

$$NPV = -100 + 300/1.05 \approx \$185.71.$$

Because this net present value is positive, the firm makes the investment.

Solved Problem 16.2

Peter Guber and Joe Lacob bought the Golden State Warriors basketball team for \$450 million in 2010. *Forbes* magazine estimated that the team's net income for 2009 was \$11.9 million. If the new owners believed that they would continue to earn this annual profit (after adjusting for inflation), $f = \$11.9$ million, forever, was this investment more lucrative than putting the \$450 million in a savings account that pays a real interest rate of $i = 2\%$ or $i = 3\%$?

Answer

1. *Determine the net present value of the team at a real interest rate of 2%.* Using Equation 16.4, the present value of the stream of income was \$11.9 million/0.02 = \$595 million. Thus, the net present value of buying the Warriors was \$595 million minus the present value of the cost, which was the purchase price of \$450 million:

$$NPV = \$595 \text{ million} - \$450 \text{ million} = \$145 \text{ million} > 0.$$

Consequently, it paid for the investors to buy the Warriors if their best alternative investment paid 2%.

2. *Determine the net present value at a real interest rate of 3%.* At this higher interest rate, the present value of the income stream was only \$11.9 million/0.03 ≈ \$397 million, so the investment would not pay: $NPV = \$397 - \$450 = -53$ million < 0.

internal rate of return the discount rate that results in a net present value of an investment of zero

Internal Rate of Return Approach Whether the net present value of an investment is positive depends on the interest rate, as Solved Problem 16.2 shows. At what discount rate (rate of return) is a firm indifferent between making an investment and not? The **internal rate of return** (*irr*) is the discount rate such that the net present value of an investment is zero. Replacing the interest rate, *i*, in Equation 16.7 with *irr* and setting the *NPV* equal to zero, we implicitly determine the internal rate of return by solving

$$NPV = \pi_0 + \frac{\pi_1}{1 + irr} + \frac{\pi_2}{(1 + irr)^2} + \cdots + \frac{\pi_T}{(1 + irr)^T} = 0$$

for *irr*.

It is easier to calculate *irr* when the investment pays a steady stream of profit, f, forever and the cost of the investment is PV. The investment's rate of return is found by rearranging Equation 16.4 and replacing i with *irr*:

$$irr = \frac{f}{PV}. \tag{16.8}$$

Instead of using the net present value rule, we can decide whether to invest by comparing the internal rate of return to the interest rate. If the firm is borrowing money to make the investment, *it pays for the firm to borrow to make the investment if the internal rate of return on that investment exceeds that of the next best alternative* (which we assume is the interest rate):[6]

$$irr > i.$$

Solved Problem 16.3

Peter Guber and Joe Lacob can buy the Golden State Warriors basketball team for \$450 million, and they expect an annual real flow of payments (profits) of f = \$11.9 million forever. Using the internal rate of return approach, should they buy the team if the real interest rate is 2%?

Answer

Determine the internal rate of return to this investment and compare it to the interest rate. Using Equation 16.8, we calculate that the internal rate of return from buying the Warriors is

$$irr = \frac{f}{PV} = \frac{\$11.9 \text{ million}}{\$450 \text{ million}} \approx 2.6\%.$$

Because this internal rate of return, 2.6%, is greater than the real interest rate, 2%, they buy the team.

Rate of Return on Bonds

Instead of investing in capital or putting their money in a bank, firms or individuals may invest in a *bond*, a piece of paper issued by a government or a corporation that promises to repay the borrower with a payment stream. The amount borrowed is called the *face value* of the bond. Some bonds have a number of *coupons*. Each year, the holder of the bond clips one coupon, returns it to the issuer, and receives a payment of a fixed amount of money. At the *maturity date* shown on the bond—when no coupons remain—the borrower redeems the bond by returning the face value, the amount borrowed.

Some bonds, *perpetuities*, have no maturity date and the face value is never returned. Instead, the bondholder receives annual payments forever.

For example, last year Jerome paid PV = \$2,000 to buy a government-issued bond that guarantees the holder a payment of f = \$100 a year forever. According to Equation 16.8, the rate of return on Jerome's bond was 5% = \$100/\$2,000. At the

[6] The net present value approach always works. The internal rate of return method is inapplicable if *irr* is not unique. In Solved Problem 16.3, *irr* is unique, and using this approach gives the same answer as the net present value approach.

time, banks were paying 5% on comparable accounts and were expected to do so in the future. As a result, Jerome was indifferent between buying a bond and keeping his money in a bank account.

This year, however, because of *unanticipated* inflation, the nominal interest rate that banks paid *unexpectedly* rose to 10%, and everyone expects this new interest rate to persist. If the bonds were to continue to sell for \$2,000, the rate of return would remain 5%, so everyone would prefer to keep their money in the bank. Thus, if Jerome wants to sell his bond, he must lower the price until the rate of return on the bond reaches 10%. As a result, the present value of Jerome's bond falls to \$1,000 = \$100/0.1 this year, according to Equation 16.4. In general, a bond's selling price falls from the face value of the bond if the nominal interest rate rises over time (and the price rises if the interest rate falls).

Similarly, the real return to a bond that pays a nominal rate of return varies with the inflation rate. During the high-inflation 1970s and early 1980s, holders of U.S. bonds lost much of their wealth for this reason. Following Canada, Britain, and other countries, the United States in 1997 started offering bonds that adjust for the inflation rate. These bonds are supposed to provide a constant, real rate of return.

★ Behavioral Economics: Time-Varying Discounting

Tomorrow: One of the greatest labor saving devices of today.

People want immediate gratification.[7] We want rewards now and costs delayed until later: "Rain, rain, go away; Come again some other day; We want to go outside and play; Come again some other day."

Time Consistency So far in this chapter, we have explained such impatience by assuming that people discount future costs or benefits by using *exponential discounting*, as in Equation 16.2: The present value is the future value divided by $(1 + i)^t$, where t is the exponent and the discount rate, i, is constant over time. If people use this approach, their preferences are *time consistent*: They will discount an event that occurs a decade from the time they're asked by the same amount today as they will one year from now.

However, many of us indulge in immediate gratification in a manner that is inconsistent with our long-term preferences: Our "long-run self" disapproves of the lack of discipline of our "short-run self." Even though we plan today not to overeat tomorrow, tomorrow we may overindulge. We have *present-biased preferences*: When considering the trade-off between two future moments, we put more weight on the earlier moment as it gets closer. For example, if you are offered \$100 in 10 years or \$200 in 10 years and a day, you will almost certainly choose the larger amount one day later. After all, what's the cost of waiting one extra day a decade from now? However, if you are offered \$100 today or \$200 tomorrow, you may choose the smaller amount today because an extra day is an appreciable delay when your planning horizon is short.

One explanation that behavioral economists (see Chapter 4) use for procrastination and other time-inconsistent behavior is that people's personal discount rates are smaller in the far future than in the near future. For example, suppose that you know that you can mow your lawn today in two hours, but if you wait until next week, it will take you two-and-a-quarter hours because the grass will be longer. Your displeasure (negative utility) from spending 2 hours mowing is -20 and from spending 2.25 hours mowing is -22.5. The present value of mowing next week is $-22.5/(1 + i)$,

[7]This section draws heavily on Rabin (1998), O'Donoghue and Rabin (1999), and Karp (2005).

where i is your personal discount rate for a week. If today your discount rate is $i = 0.25$, then your present value of mowing in a week is $-22.5/1.25 = -18$, which is not as bad as -20, so you delay mowing. However, if you were asked six months in advance, your discount rate might be much smaller, say $i = 0.1$. At that interest rate, the present value is $-22.5/1.1 \approx -20.45$, which is worse than -20, so you would plan to mow on the first of the two dates. Thus, falling discount rates may explain this type of time-inconsistent behavior.

Falling Discount Rates and the Environment A social discount rate that declines over time may be useful in planning for global warming or other future environmental disasters (Karp, 2005). Suppose that the harmful effects of greenhouse gases will not be felt for a century and that society used traditional, exponential discounting. We would be willing to invest at most 37¢ today to avoid a dollar's worth of damages in a century if society's constant discount rate is 1%, and only 2¢ if the discount rate is 4%. Thus, even a modest discount rate makes us callous toward our distant descendants: We are unwilling to incur even moderate costs today to avoid large damages far in the future.

One alternative is for society to use a declining discount rate, although doing so will make our decisions time inconsistent. Parents today may care more about their existing children than their (not yet seen) grandchildren, and therefore may be willing to significantly discount the welfare of their grandchildren relative to that of their children. They probably have a smaller difference in their relative emotional attachment to the tenth future generation relative to the eleventh generation. If society agrees with such reasoning, our future social discount rate should be lower than our current rate. By reducing the discount rate over time, we are saying that the weights we place on the welfare of any two successive generations in the distant future are more similar than the weights on two successive generations in the near future.

Application

Falling Discount Rates and Self-Control

If people's discount rates fall over time, they have a *present bias* or a *self-control problem*, which means that they prefer immediate gratification to delayed gratification.[8] Several recent studies argue that governments should help people with this bias by providing self-control policies.

Shapiro (2004) finds that food stamp recipients' caloric intake declines by 10% to 15% over the food stamp month, implying that they prefer immediate consumption. With a constant discount rate, they would be more likely to spread their consumption evenly over the month. Governments can help people with a present bias by delivering food stamps at two-week intervals instead of once a month, as several states do with welfare payments.

Cigarette smokers often have inconsistent preferences with respect to smoking. Individuals with declining discount rates lack self-control and perpetually postpone quitting smoking. A 2013 Gallup poll found that 74% of U.S. smokers would like to give up smoking. According to a 2015 survey, 56% of Beijing smokers said they want to kick their habit. Consequently, a smoker who wants to quit may support the government's impositions of control devices. Based on a survey in Taiwan, Kan (2007) finds that a smoker who intends to quit is more likely to support a smoking

[8] In the famous marshmallow test, small children are offered one marshmallow now or a second one if they wait. See an excellent reenactment at **https://www.youtube.com/watch?v=QX_oy9614HQ**. Children who could delay gratification did better later in life: **http://www.newyorker.com/magazine/2009/05/18/dont-2**.

ban and a cigarette tax increase. In 2012, most (59%) New Zealand smokers supported more government action on tobacco, and nearly half (46%) supported banning sales of cigarettes in 10 years, provided effective nicotine substitutes were available. In 2014, 39% of smokers favored higher taxes (up from 29% in 2002).

In 2009, President Obama—a smoker who wanted to quit—signed a law bringing tobacco products under federal law for the first time. He said that this law, aimed at stopping children from starting to smoke, would have prevented him from taking up smoking. Perhaps the most striking evidence of smokers' mixed feelings is that Gruber and Mullainathan (2005) found that cigarette taxes make people with a propensity to smoke happier in both the United States and Canada.

16.3 Exhaustible Resources

The meek shall inherit the earth, but not the mineral rights. —J. Paul Getty

exhaustible resources nonrenewable natural assets that cannot be increased, only depleted

Discounting plays an important role in decision making about how fast to consume oil, gold, copper, uranium, and other **exhaustible resources**: nonrenewable natural assets that cannot be increased, only depleted. An owner of an exhaustible resource decides when to extract and sell it so as to maximize the present value of the resource. Scarcity of the resource, mining costs, and market structure affect whether the price of such a resource rises or falls over time.

When to Sell an Exhaustible Resource

Suppose that you own a coal mine. In what year do you mine the coal, and in what year do you sell it to maximize the present value of your coal? To illustrate how to answer these questions, we assume that you can sell the coal only this year or next in a competitive market, that the interest rate is i, and that the cost of mining each pound of coal, m, stays constant over time.

Given the last two of these assumptions, the present value of mining a pound of coal is m if you mine it this year and $m/(1 + i)$ if you mine it next year. As a result, if you're going to sell the coal next year, you're better off mining it next year because you postpone incurring the cost of mining. You mine the coal this year only if you plan to sell it this year.

Now that you have a rule that tells you when to mine the coal—at the last possible moment—your remaining problem is when to sell it. That decision depends on how the price of a pound of coal changes from one year to the next. Suppose that you know that the price of coal will increase from p_1 this year to p_2 next year.

To decide in which year to sell, you compare the present value of selling today to that of selling next year. The present value of your profit per pound of coal is $p_1 - m$ if you sell your coal this year and $(p_2 - m)/(1 + i)$ if you sell it next year. Thus, to maximize the present value from selling your coal:

- *You sell all the coal this year* if the present value of selling this year is greater than the present value of selling next year: $p_1 - m > (p_2 - m)/(1 + i)$.
- *You sell all the coal next year* if $p_1 - m < (p_2 - m)/(1 + i)$.
- *You sell the coal in either year* if $p_1 - m = (p_2 - m)/(1 + i)$.

The intuition behind these rules is that storing coal in the ground is like keeping money in the bank. You can sell a pound of coal today, netting $p_1 - m$, invest the money in the bank, and have $(p_1 - m)(1 + i)$ next year. Alternatively, you can keep the coal in the ground for a year and then sell it. If the amount you'll get next year,

$p_2 - m$, is less than what you can earn from selling now and keeping the money in a bank account, you sell the coal now. In contrast, if the price of coal is rising so rapidly that the coal will be worth more in the future than wealth left in a bank, you leave your wealth in the mine.

Price of a Scarce Exhaustible Resource

This two-period analysis generalizes to many periods (Hotelling, 1931). We use a multiperiod analysis to show how the price of an exhaustible resource changes over time.

The resource is sold both this year, year t, and next year, $t + 1$, only if the present value of a pound sold now is the same as the present value of a pound sold next year: $p_t - m = (p_{t+1} - m)/(1 + i)$, where the price is p_t in year t and p_{t+1} in the following year. Using algebra to rearrange this equation, we obtain an expression that tells us how price changes from one year to the next:

$$p_{t+1} = p_t + i(p_t - m). \tag{16.9}$$

If you're willing to sell the coal in both years, the price next year must exceed the price this year by $i(p_t - m)$, which is the interest payment you'd receive if you sold a pound of coal this year and put the profit in a bank that pays an interest rate of i.

The gap between the price and the constant marginal cost of mining grows over time, as Figure 16.2 shows. To see why, we subtract p_t from both sides of Equation 16.9 to obtain an expression for the change in the price from one year to the next:

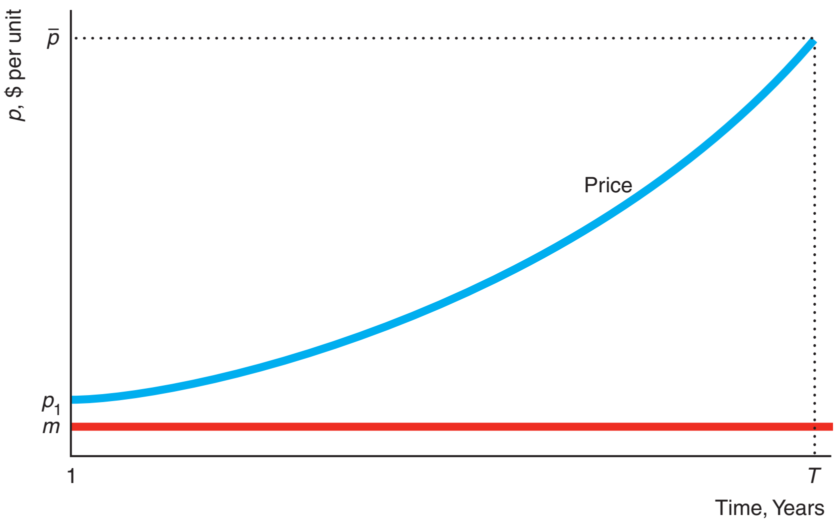
$$\Delta p = p_{t+1} - p_t = i(p_t - m).$$

This equation shows that the gap between this year's price and next year's price widens as your cash flow this year, $p_t - m$, increases. Thus, the price rises over time, and the gap between the price line and the flat marginal cost of mining line grows, as the figure illustrates.

Although we now understand how price changes over time, we need more information to determine the price in the first year and hence in each subsequent year.

Figure 16.2 Price of an Exhaustible Resource

The price of an exhaustible resource in year $t + 1$ is higher than the price in year t by the interest rate times the difference between the price in year t and the marginal cost of mining, $i(p_t - m)$. Thus, the gap between the price line and the marginal cost line, $p_t - m$, grows exponentially with the interest rate.



Suppose that mine owners know that the government will ban the use of coal in year T (or that a superior substitute will become available that year). They want to price the coal so that all of it is sold by the year T, because any resource that is unsold by then is worthless. The restriction that all the coal is used up by T and Equation 16.9 determine the price in the first year and the increase in the price thereafter.

Price in a Two-Period Example To illustrate how the price is determined in each year, we assume that the market has many identical competitive mines, no more coal will be sold after the second year because of a government ban, and the marginal cost of mining is zero in each period. Setting $m = 0$ in Equation 16.9, we learn that the price in the second year equals the price in the first year plus the interest rate times the first-year price:

$$p_2 = p_1 + (i \times p_1) = p_1(1 + i). \tag{16.10}$$

Thus, the price increases with the interest rate from the first year to the second year.

The mine owners face a resource constraint: They can't sell more coal than they have in their mines. The coal they sell in the first year, Q_1, plus the coal they sell in the second year, Q_2, equals the total amount of coal in the mines, Q. The mine owners want to sell all their coal within these two years because any coal they don't sell does them no good.

Suppose that the demand curve for coal is $Q_t = 200 - p_t$ in each year t. If the amount of coal in the ground is less than would be demanded at a zero price, the sum of the amount demanded in both years equals the total amount of coal in the ground:

$$Q_1 + Q_2 = (200 - p_1) + (200 - p_2) = Q.$$

Substituting the expression for p_2 from Equation 16.10 into this resource constraint to obtain $(200 - p_1) + [200 - p_1(1 + i)] = Q$ and rearranging terms, we find that

$$p_1 = (400 - Q)/(2 + i). \tag{16.11}$$

Thus, the first-year price depends on the amount of coal in the ground and the interest rate.

If the mines initially contain $Q = 169$ pounds of coal, p_1 is \$110 at a 10% interest rate and only \$105 at a 20% interest rate, as Table 16.5 shows. At the lower interest rate, the difference between the first- and second-year price is smaller (\$11 versus \$21), so relatively more of the original stock of coal is sold in the second year (47% versus 44%).

Table 16.5 Price and Quantity of Coal Reflecting the Amount of Coal and the Interest Rate

	$Q = 169$		$Q = 400$
	$i = 10\%$	$i = 20\%$	Any i
$P_1 = (400 - Q)/(2 + i)$	\$110	\$105	\$0
$P_2 = p_1(1 + i)$	\$121	\$126	\$0
$\Delta p \equiv p_2 - p_1 = i \times p_1$	11	21	0
$Q_1 = 200 - p_1$	90	95	200
$Q_2 = 200 - p_2$	79	74	200
Share sold in Year 2	47%	44%	50%

Rents If coal is a scarce good, its competitive price is above the marginal cost of mining the coal ($m = 0$ in our example). How can we reconcile this result with our earlier finding that price equals marginal cost in a competitive market? The answer is that when coal is scarce, it earns a *rent*: a payment to the owner of an input beyond the minimum necessary for the factor to be supplied (Chapter 9).

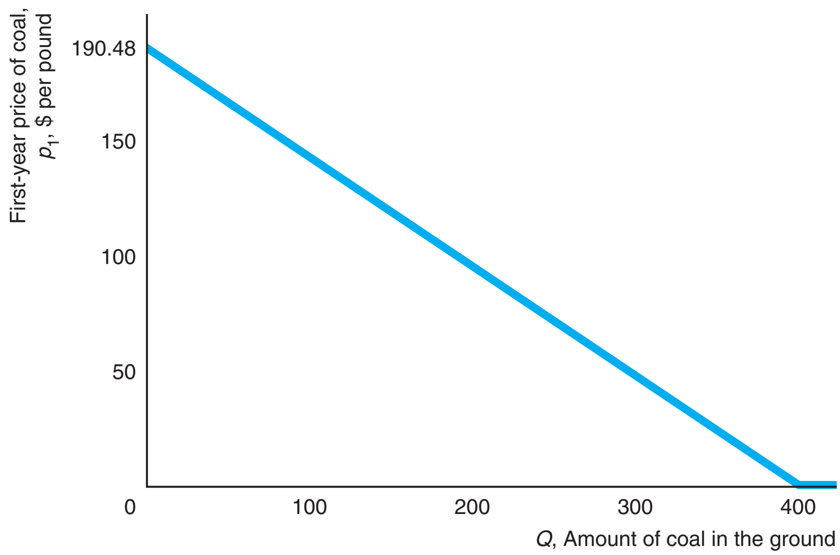
The owner of the coal need not be the same person who mines the coal. A miner could pay the owner for the right to take the coal out of the mine. After incurring the marginal cost of mining the coal, m, the miner earns $p_1 - m$. The owner of the mine, however, charges that amount in rent for the right to mine this scarce resource rather than giving any of this profit to the miner. Even if the owner of the coal and the miner are the same person, the amount beyond the marginal mining cost is a rent to scarcity.

If the coal were not scarce, no rent would be paid, and the price would equal the marginal cost of mining. Given the demand curve in the example, the most coal anyone would buy in a year is 200 pounds, which is the amount demanded at a price of zero. If the initial stock of coal in the group is 400 pounds—enough to provide 200 pounds in each year—coal is not scarce, so the price of coal in both years is zero (the marginal mining cost), as Table 16.5 shows.[9] As Figure 16.3 illustrates, the less coal in the ground initially, Q, the higher the initial price of coal.

Rising Prices Thus, according to our theory, the price of an exhaustible resource rises if the resource (1) is scarce, (2) can be mined at a marginal cost that remains constant over time, and (3) is sold in a competitive market. The price of old-growth redwood trees rose as predicted by this theory.

Figure 16.3 First-Year Price in a Two-Period Model

In a two-period model, the price of coal in the first year, p_1, falls as the amount of coal in the ground initially, Q, increases. This figure is based on an interest rate of 10%.



[9]Equation 16.11 holds only where coal is scarce: $Q \le 400$. According to this equation, $p_1 = 0$ when $Q = 400$. If the quantity of coal in the ground is even greater, $Q > 400$, coal is not scarce—people don't want all the coal even if the price is zero—so the price in the first year equals the marginal mining cost of zero. That is, the price is not negative, as Equation 16.11 would imply if it held for quantities greater than 400.

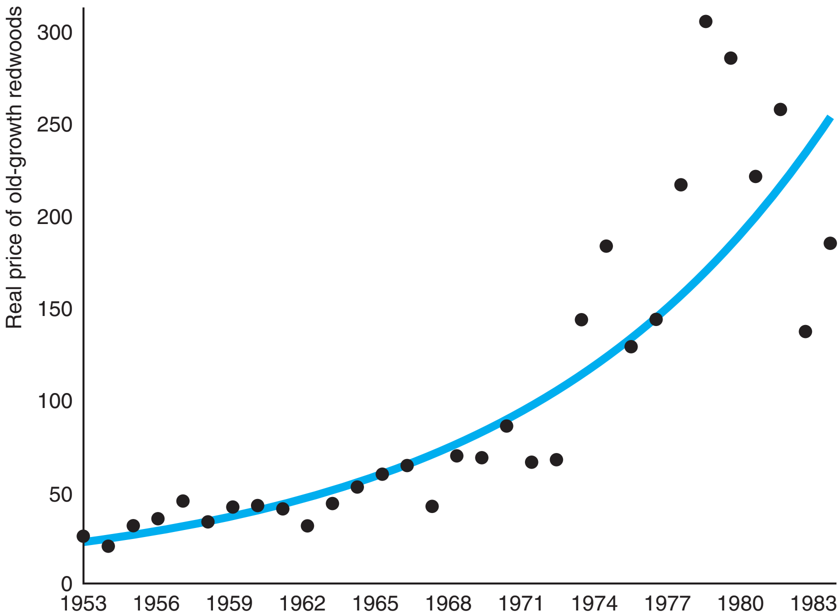
Application

Redwood Trees

Many of the majestic old-growth redwood trees in America's western forests are several hundred to several thousand years old. If a mature redwood is cut, young redwoods will not grow to a comparable size within our lifetime. Thus, an old-growth redwood forest, like fossil fuels, is effectively a nonrenewable resource, even though new redwoods are being created and grow very slowly. In contrast, many other types of trees, such as those grown as Christmas trees, are quickly replenished and therefore are renewable resources like fish.

The exponential trend line on the graph shows that the real price of redwoods rose from 1953 to 1983 at an average rate of 8% a year. By the end of this period, virtually no redwood trees were available for sale. The trees either had been harvested or were growing in protected forests. The last remaining privately owned stand was purchased by the U.S. government and the state of California from the Maxxam Corporation in 1996.

The unusually high prices observed in the late 1960s through the 1970s are in large part due to actions of the federal government, which used its power of eminent domain to buy a considerable fraction of all remaining old-growth redwoods for the Redwood National Park at the market price. The government bought 1.7 million million-board feet (MBF) in 1968 and 1.4 million MBF in 1978. The latter purchase represented about two and a quarter years of cutting at previous rates. These two government purchases combined equaled 43% of private holdings in 1978 of about 7.3 million MBF. Thus, the government purchases were so large that they moved up the time of exhaustion of privately held redwoods by several years, causing the price to jump to the level it would have reached several years later.



Why Price May Be Constant or Fall

If any one of the three conditions we've been assuming—*scarcity*, *constant marginal mining costs*, and *competition*—is not met, the price of an exhaustible resource may remain steady or fall.[10] Most exhaustible resources, such as aluminum, coal, lead, natural gas, silver, and zinc, have had decades-long periods of falling or constant real prices. Indeed, the real price of each major mineral, each metal, and oil was lower in 1998 than in 1980.

Abundance As we've already seen, the initial price is set at essentially the marginal cost of mining if the exhaustible resource is not scarce. The gap between the price and the marginal cost grows with the interest rate. If the good is so abundant that the initial gap is zero, the gap does not grow and the price stays constant at the marginal cost. Further, if the gap is initially very small, it has to grow for a long time before the increase becomes noticeable.

Because of abundance, the real prices for many exhaustible resources have remained relatively constant for decades. Moreover, the price falls when the discovery of a large deposit of the resource is announced.

The amount of a resource that can be profitably recovered using current technology is called a *reserve*. Known reserves of some resources are enormous; others are more limited.[11] The world has enough silicon (from sand) and magnesium to last virtually forever at current rates of extraction. Known reserves of zinc will last 17 years; lead, 16 years; gold, 19 years; and silver, 20 years. Known reserves of aluminum (bauxite) will last over a century, and additional reserves are constantly being discovered. Because of this abundance, the real price of aluminum has remained virtually constant for the past 50 years.

Technical Progress Improved technology increased potential natural gas reserves substantially. Over long periods of time, steady technical progress has reduced the marginal cost of mining many natural resources and has thereby lowered the price of those exhaustible resources. A large enough drop in the marginal mining cost may more than offset the increase in the price due to the interest rate, so the price falls from one year to the next.[12]

The era spanning the end of the nineteenth century and the beginning of the twentieth century witnessed many advances in mining. As a result of technical progress in mining and discoveries of new supplies, the real prices of many exhaustible resources fell. For example, the real price of aluminum in 1945 was only 12% of the price 50 years earlier. Eventually, as mines play out, prospectors have to dig ever deeper to find resources, causing marginal costs to increase and prices to rise faster than they would with constant marginal costs. New means of recovering natural gas from shale increased estimates of reserves by 38% in the United States and by 48% in the rest of the world.[13]

Changing Market Power Changes in market structure can result in either a rise or a fall in the price of an exhaustible resource. The real price of oil remained virtually constant from 1880 through 1972. But when the Organization of Petroleum Exporting Countries (OPEC) started to act as a cartel in 1973, the price of oil climbed

---

[10]The following discussion of why prices of exhaustible resources may not rise and the accompanying examples are based on Berck and Roberts (1996) and additional data supplied by these authors. Their paper also shows that pollution and other environmental controls can keep resource prices from rising. Additional data are from Brown and Wolk (2000).

[11]Data are from **http://minerals.usgs.gov/minerals/pubs/mcs/2015/mcs2015.pdf**.

[12]When the marginal cost of mining is constant at m, Equation 16.9 shows that $p_{t+1} = p_t + i(p_t - m)$, so p_{t+1} must be more than p_t. If we allow mining costs to vary from year to year, then $p_{t+1} = p_t + i(p_t - m_t) + (m_{t+1} - m_t)$. Thus, if the drop in the mining costs, $m_{t+1} - m_t$, is greater than $i(p_t - m)$, the price p_{t+1} is less than p_t.

[13]**http://www.eia.gov/analysis/studies/worldshalegas**.

rapidly. At its peak in 1981, the real price of oil was nearly five times higher than its nearly constant level during the period 1880–1972. When Iran and Iraq went to war in 1980, the OPEC cartel began to fall apart and the real price of oil sank to traditional levels, where it remained through the 1990s. In the first decade of the new millennium, the price increased substantially, in large part due to worldwide increases in demand. Then, up to 2014, the price fluctuated due to political uncertainty. Since then, the price fell due to production increases.

16.4 Capital Markets, Interest Rates, and Investments

We've seen that an individual's decision about whether to make an investment depends on the market interest rate. The interest rate is determined in the capital market, where the interest rate is the price, the quantity supplied is the amount of funds loaned, and the quantity demanded is the amount of funds borrowed.

Did you realize that government spending affects your investment opportunities (such as paying to go to college)?

Common Confusion: Government spending does not affect me personally.

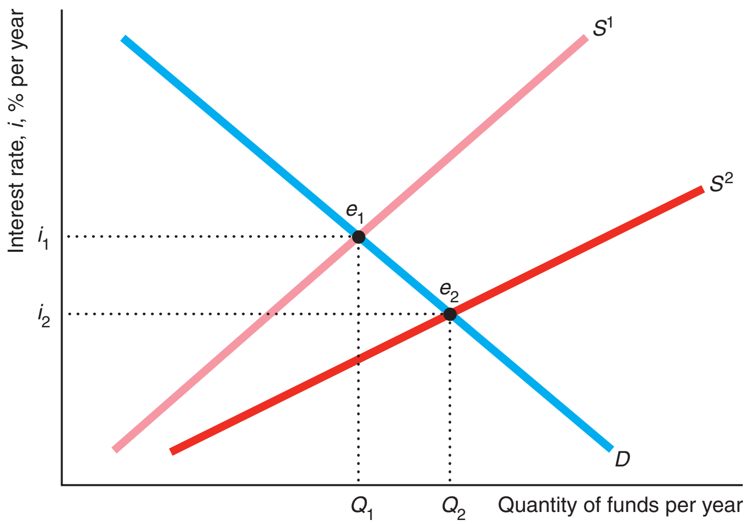
If the government borrows to pay for its spending, it affects the market interest rate, which affects your investment opportunities.

Because the capital market is competitive, the interest rate and the quantity of funds loaned and borrowed is determined by the intersection of the supply curve for funds and the demand curve for funds, as in Figure 16.4. Funds are demanded by individuals buying homes or paying for a college education, governments borrowing money to build roads or wage wars, and firms investing in new plants or equipment. The demand curve, D, is downward sloping because more is borrowed as the interest rate falls.

The supply curve reflects loans made by individuals and firms. Many people, when their earnings are relatively high, save money in bank accounts and buy bonds (which

Figure 16.4 Capital Market Equilibrium

The initial equilibrium, e_1, is determined by the intersection of the demand curve for loans, D, and the initial supply curve, S^1. Changes in laws that induce more people to save shift the supply curve to S^2. The interest rate, i_2, at the new equilibrium, e_2, is lower than the original interest rate, i_1. More funds are loaned than originally: $Q_2 > Q_1$.



they convert back to money for consumption when they retire or during lean times). Firms that have no alternative investments with higher returns may also loan money to banks or others. Higher interest rates induce greater savings by both groups, so the initial supply curve, S^1, is upward sloping.

The initial equilibrium is e_1, with an equilibrium interest rate of i_1 and an equilibrium quantity of funds loaned and borrowed of Q_1. As usual, this equilibrium changes if any of the variables—such as tastes and government regulations—that affect supply and demand shifts.

The supply curve of funds may shift to the right for many reasons. The government may remove a restriction on investment by foreigners. Or the government may make Individual Retirement Accounts (IRAs) tax exempt until retirement, a policy that induces additional savings at any given interest rate.

Such a change causes the supply curve to shift to the right to S^2 in Figure 16.4. The new equilibrium is e_2, with a lower interest rate, i_2. At the lower interest rate, firms and others undertake investment projects with lower rates of return than before the shift. They borrow more funds, so the new equilibrium is at $Q_2 > Q_1$.

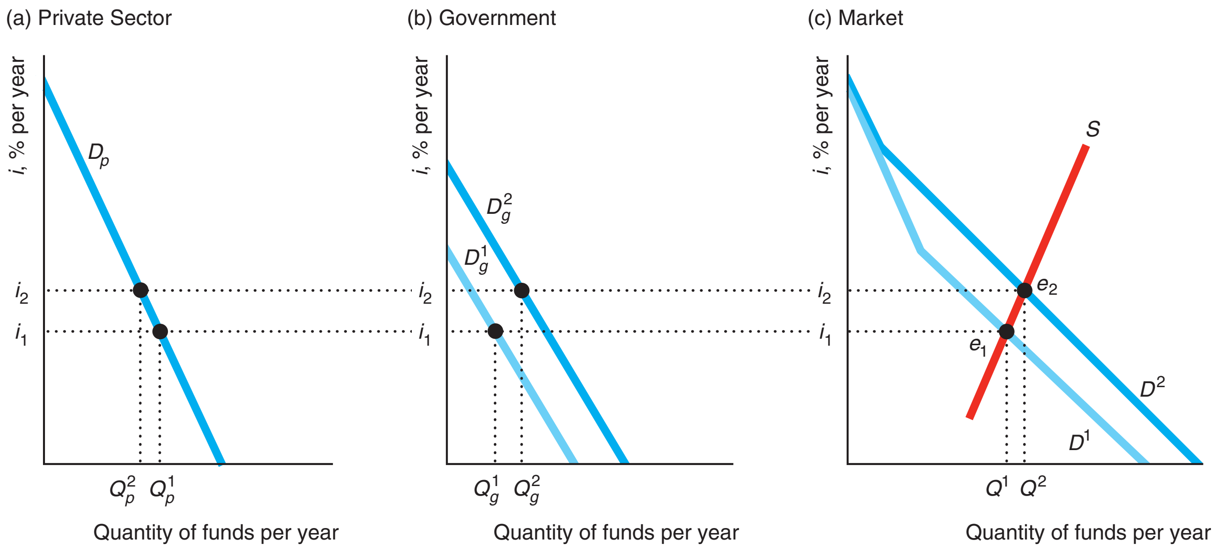
Solved Problem 16.4

MyLab Economics Solved Problem

Suppose the government needs to borrow money to pay for fighting a war in a foreign land. Show that increased borrowing by the government—an increase in the government's demand for money at any given interest rate—raises the equilibrium interest rate, which discourages or *crowds out* private investment.

Answer

Using three side-by-side graphs, show how an outward shift of the government's demand curve affects the equilibrium interest rate and thereby reduces private investment. In the figure, panel a shows the private sector demand curve for funds, D_p, which are funds that private firms and individuals borrow to make investments. Panel b shows that the government sector demand curve shifts to the right from D_g^1 to D_g^2. As a result, in panel c, the total demand curve—the horizontal sum of the private and government demand curves—shifts from D^1 to D^2. Panel c also shows the supply curve of money, S.



The initial equilibrium, e_1 in panel c, is determined by the intersection of the initial total demand for funds, D^1, and the supply curve, S, where the interest rate is i_1 and the quantity of funds borrowed is Q_1. After the government demand curve shifts out, the new equilibrium is e_2, where the interest rate is higher, $i_2 > i_1$, and more funds are borrowed, $Q_2 > Q_1$. The higher market interest rate causes private investment to fall from Q_p^1 to Q_p^2 (panel a). That is, the government borrowing crowds out some private investment.

Challenge Solution

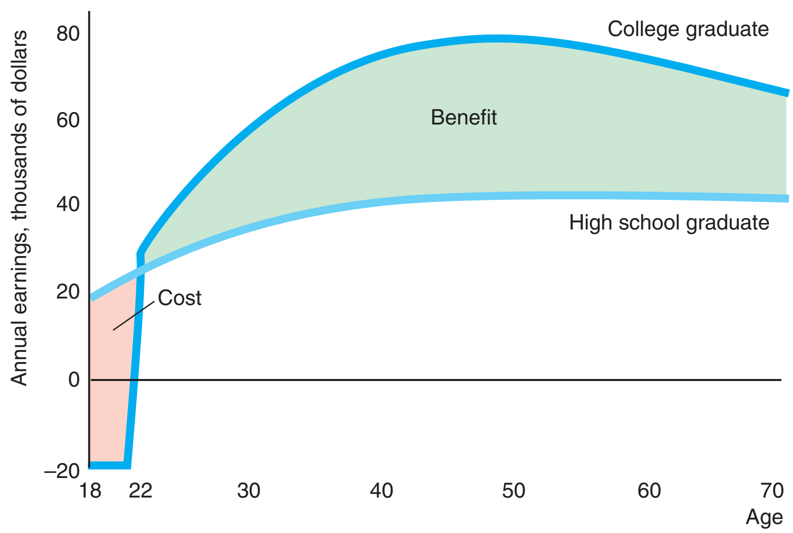
Does Going to College Pay?

I have often thought that if there had been a good rap group around in those days, I would have chosen a career in music instead of politics.
—Richard Nixon

Probably the most important human capital decision you've had to make was whether to attend college. If you opted to go to college solely for the purpose of increasing your lifetime earnings, have you made a good investment?

Let's look at the decision you faced during your last year of high school. You have to decide whether to invest in a college education or go directly into the job market. If you go straight into the job market, we assume that you work full time (35 hours or more a week) from age 18 until you retire at age 70.

If you want to attend college to increase your lifetime earnings, you should start college upon finishing high school so that you can earn a higher salary for as long as possible. To keep the analysis relatively simple, we'll assume that you graduate from college in four years, during which time you do not work and you spend $20,000 a year on tuition and other schooling expenses such as books and fees. When you graduate from college, you work full time from age 22 to 70. Thus, the opportunity cost of a college education includes the tuition payments plus the four years of forgone earnings for someone with a high school diploma. The expected benefit is the stream of higher earnings in the future.



The figure shows how much the typical person earns with a high school diploma and with a college degree at each age.[14] At age 22, a typical college grad earns \$29,335, and those with only a high school diploma earn \$25,009. The college grad's earnings peak at 51 years of age, at \$77,865. A high school grad's earnings also reach a maximum at 51 years, at \$42,707.

If one stream of earnings is higher than the other at every age, we would pick the higher stream. Because these streams of earnings cross at age 22, we cannot use that simple approach to answer the question. One way to decide whether investing in a college education pays is to compare the present values at age 18 of the two earnings streams. The present values depend on the interest rate used, as the table shows.

	Present Value, Thousands of 2009 Dollars	
Discount Rate, %	**High School**	**College**
0	2,007	**3,225**
2	1,196	**1,823**
4	779	**1,103**
6	547	**708**
8	410	**476**
10	323	**332**
10.42	**309**	**309**
12	**264**	238
14	**223**	174

If potential college students can borrow money at an interest rate of 0%, money in the future is worth as much as money today, so the present value equals the sum of earnings over time. According to the table, the sum of a college graduate's earnings (including the initial negative earnings) is \$3.23 million (first row of the table), which is 61% more than the lifetime earnings of a high school grad, \$2.01 million. Thus, it pays to go to college if the interest rate is 0%. The figure also illustrates that attending college pays at a 0% discount rate because the sum of the (negative) cost and (positive) benefit areas—the difference in earnings between going to college and going to work after high school—is positive.

The table demonstrates that the present value of earnings for a college grad equals that of a high school grad at an interest rate of 10.42%. That is, the average internal rate of return to the college education is 10.42%. Because the present value of earnings for a college grad exceeds that of a high school grad if the real interest rate at which they can borrow or invest is less than 10.42%, income-maximizing people should go to college if the real interest rate is less than that rate.[15]

---

[14]The statistical analysis controls for age, education, and demographic characteristics but not innate ability. See Sources for information about the data, which are for 2009. I thank Yann Panassie, a former student in my intermediate microeconomics course, for estimating this model. We assumed that wages increase at the same rate as inflation, so real earnings are constant over time. No adjustment is made for the greater incidence of unemployment among high school graduates, which was nearly twice that of college graduates in 2015.

[15]The government-subsidized nominal interest rate on federal Stafford loans was 3.4% in 2011–2013 and 4.29% for the 2015–2016 academic year. Some poor people who cannot borrow to pay for college at all—effectively, they face extremely high interest rates—do not go to college, unlike wealthier people with comparable abilities.

According to Payscale.com in 2015, the average internal rate of return of going to college is higher for students at some schools than others: 12.5% at Harvey Mudd and Georgia Institute of Technology; 11.4% at Colorado School of Mines and Texas A&M; 10.4% at Iowa State University (in-state); 10.0% at New Mexico State University (in-state); 9.9% at Rutgers and the University of California, Berkeley (in-state); 7.9% at Stanford; 7.1% at Harvard; 6.9% at Lehigh University; and 0.7% at the School of the Art Institute of Chicago.[16]

The decision whether to go to college is more complex for people for whom education has a consumption component. Somebody who loves school may want to go to college even if alternative investments pay more. Someone who hates going to school invests in a college education only if the financial rewards are much higher than those for alternative investments.

Summary

1. **Comparing Money Today to Money in the Future.** Inflation aside, most people value money in the future less than money today. An interest rate reflects how much more people value a dollar today than a dollar in the future. To compare a payment made in the future to one made today, we can express the future payment in terms of current dollars by adjusting it using the interest rate. Similarly, a flow of payments over time is related to the present or future value of these payments by the interest rate.
2. **Choices over Time.** An individual or a firm may choose between two options with different cash flows over time by picking the one with the higher present value. Similarly, a firm invests in a project if its net present value is positive or its internal rate of return is greater than the interest rate. If people have a decreasing discount rate over time, they are not consistent in their behavior over time: They lack self-control and procrastinate.
3. **Exhaustible Resources.** Nonrenewable resources such as coal, gold, and oil are used up over time and cannot be replenished. If these resources are scarce, the marginal cost of mining them is constant or increasing, and the market structure remains unchanged, their prices rise rapidly over time because of positive interest rates. However, if the resources are abundant, the marginal cost of mining falls over time, or the market becomes more competitive, nonrenewable resource prices may remain constant or fall over time.
4. **Capital Markets, Interest Rates, and Investments.** Supply and demand in capital markets determine the market interest rate. A shock that shifts the supply curve to the left or the demand curve to the right raises the interest rate. As the interest rate increases, firms want to make fewer investments.

Questions

*Select questions are available on MyLab Economics; * = answer appears at the back of this book;* **A** *= algebra problem.*

1. Comparing Money Today to Money in the Future

1.1 People and firms invest in human and physical capital. Both types of capital are durable goods, the benefits from which can last for many years. Entering into a credit agreement is often key in obtaining the funds necessary to enable such investments. How do credit arrangements benefit both borrowers and lenders?

1.2 Under Section 347 of the Canadian Criminal Code, an annual rate of interest that exceeds 60% under a credit agreement is a "criminal" rate of interest. The term "interest" is defined to include not only ordinary commercial interest but also any fees, fines, expenses, shares, and royalties payable to a lender. What effect does a usury law like this have on lenders who are still willing to loan money at illegally high rates with the hope that the authorities do not catch them?

1.3 The website **http://www.timetravelfund.com** discusses investing $1 at 5% interest, which it says will be worth $39,323,261,827.22 in 500 years. Is its calculation correct, and, if so, for what frequency of compounding? If you wish, you may also

[16]For more schools, see **http://www.payscale.com/college-roi**. The Payscale's calculations, though similar to the one used in this Challenge Solution, differ in not controlling for individual characteristics and in several other ways.

discuss how good an investment you think this site provides. (*Hint*: See the Application "Power of Compounding.") **A**

1.4 Many retirement funds charge an administrative fee equal to 0.25% on managed assets. Suppose that Alexx and Spenser each invest $5,000 in the same stock this year. Alexx invests directly and earns 5% a year. Spenser uses a retirement fund and earns 4.75%. After 30 years, how much more will Alexx have than Spenser? **A**

*1.5 How does an individual with a zero discount rate weight current and future consumption? How does your answer change if the discount rate is infinite?

1.6 Discussing the $350 price of a ticket for one of her concerts, Barbra Streisand said, "If you amortize the money over 28 years, it's $12.50 a year. So is it worth $12.50 a year to see me sing? To hear me sing live?"[17] Under what condition is it useful for an individual to apply Ms. Streisand's rule to decide whether to go to the concert? What do we know about the discount rate of a person who made such a purchase?

1.7 If you buy a car for $100 down and $100 a year for two more years, what is the present value of these payments at a 5% interest rate? **A**

1.8 What is the present value of $100 paid a year from now and another $100 paid two years from now if the interest rate is i? **A**

1.9 Dell Computer makes its suppliers wait 37 days on average to be paid for their goods; however, Dell is paid by its customers immediately. Thus, Dell earns interest on this *float*, the money that it is implicitly borrowing. If Dell can earn an annual interest rate of 4%, what is this float worth to Dell per dollar spent on inputs? **A**

1.10 What is the present value of a stream of payments of f per year for t years that starts T years from now if the interest rate is i? **A**

1.11 How much money do you have to put into a bank account that pays 10% interest compounded annually to receive annual payments of $200 forever? **A**

1.12 Horizon Ford advertises that it will sell you a Taurus for $24,000 or lease it to you. To lease it, you must make a down payment of $3,000 and agree to pay $3,000 at the end of each of the next two years. After the last lease payment, you may buy the car for $20,000. If you plan to keep the car until it falls apart (at least a decade) and the interest rate is 10%, which approach has a lower present value of costs? (*Hint*: See Solved Problem 16.1.) **A**

*1.13 How much money do you have to put into a bank account that pays 10% interest compounded annually to receive perpetual annual payments of $200 in today's dollars if the rate of inflation is 5%? **A**

*1.14 You buy a new 4K TV today for €1,000, paying half today and the remaining half in one year. If the inflation rate is 2% and the nominal interest rate is 3%, what is the present value of your TV? **A**

1.15 If you spend $4 a day on a latte (in real dollars) for the rest of your life (essentially forever), what is your present discounted value at a 3% interest rate? **A**

1.16 You win a lottery. Your prize is either two annual payments of $50,000 at the end of each year or a lump-sum payment of $87,000 today. You expect the rate of inflation to be 5% and the real interest rate to be 4%. Which prize has a higher present value? (*Hint*: See the Application "Winning the Lottery.") **A**

2. Choices over Time

*2.1 A parts supplier is considering entering into a two-year contract to make its products available to one of two retail companies. Both retailers are offering an upfront payment of 20,000. Retailer 1 would pay the 80,000 balance of its total contract amount at the end of the first year. Retailer 2 would pay a total of 102,000 under its contract: 40,000 at the end of the first year and 42,000 at the end of the second year. If the discount rate is 3%, which contract should the parts supplier enter into if its only consideration is maximizing revenue? Would its decision change if the discount rate was 7%? **A**

2.2 At a 10% interest rate, do you prefer to buy a phone for $100 or to rent the same phone for $10 a year? Does your answer depend on how long you think the phone will last?

2.3 You are in the market for a used car in Egypt, and you find one that you like for E£210,000 at a local dealership. The dealer offers you a choice in paying for it: a discount of E£15,000 if you pay the full amount immediately or a E£42,000 down payment plus zero percent financing of the balance, paid annually over four years, if you pay by credit (that is, equal payments of E£42,000 at the end of each year for the next four years). If your discount rate is 4%, which payment option should you choose? **A**

2.4 You plan to buy a used refrigerator this year for $200 and to sell it when you graduate in two years.

[17]"In Other Words. . ." *San Francisco Chronicle*, January 1, 1995, Sunday Section, p. 3. She divided the $350 ticket price by 28 years to get $12.50 as the payment per year.

Assuming no inflation, you sell the refrigerator for \$100 at graduation, and the interest rate is 5%, what is the true cost (your current outlay minus the resale value in current terms) of the refrigerator to you? **A**

2.5 You want to buy a room air conditioner. The price of one machine is \$200. It costs \$20 a year to operate. The price of the other air conditioner is \$300, but it costs only \$10 a year to operate. Assuming that both machines last 10 years, which is a better deal? (Do you need to do extensive calculations to answer this question?) **A**

2.6 A Brazilian employer offers its employees a choice between taking a lump-sum distribution of pension or a pension annuity for 25 years when they retire. The choice for one employee is between accepting an immediate lump-sum payment of R\$750,000 or a R\$60,000 annuity. What is the break-even personal discount rate at which the employee would be indifferent between the two options? Would the employee take the lump-sum payment if the discount rate was 5%? (*Hint:* See Equation 16.3 and use an Excel worksheet.) **A**

2.7 Firms are increasingly offering retiring employees a choice of whether to take a lump-sum payment now or receive monthly payments for the rest of their lives (**http://www.fidelity.com/viewpoints/personal-finance/lump-sum-or-monthly-pension**). Discuss the benefits and drawbacks of accepting the lump-sum payment taking into account discounting, inflation, and uncertainty.

*2.8 Your gas-guzzling car gets only ten miles to the gallon and has no resale value, but you are sure that it will last five years. You know that you can always buy a used car for \$8,000 that gets 20 miles to the gallon. A gallon of gas costs \$2.00 and you drive 6,000 miles a year. If the interest rate is 5% and you are interested only in saving money, should you buy a new car now rather than wait until your current car dies? Would you make the same decision if you faced a 10% interest rate? **A**

2.9 You are buying a new \$20,000 car and have the option to pay for the car with a 0% loan or to receive \$500 cash back at the time of the purchase. With the loan, you pay \$5,000 down when you purchase the car and then make three \$5,000 payments, one at the end of each year of the loan. You currently have \$50,000 in your savings account.

a. The interest rate on your savings account is 4% and will remain so for the next three years. Which payment method should you choose?

b. What interest rate, *i*, makes you indifferent between the two payment methods? **A**

2.10 You are debating whether to either purchase a car or continue to take public transportation and taxis at an annual cost of €6,000 to get where you want to go. The purchase price of the car you are interested in is €24,000, and you estimate the annual expense of operating it (maintenance, gas, and parking) is €2,600. If the nominal interest rate is 5% and the rate of inflation is 2% each year, how long would you have to own the car before your investment paid off? (*Hint:* See Equation 16.3 and use an Excel worksheet.) **A**

2.11 An economic consultant explaining the effect on labor demand of increasing health care costs, interviewed for the *Wall Street Journal*'s Capital column (David Wessel, "Health-Care Costs Blamed for Hiring Gap," March 11, 2004, A2), states, "Medical costs are rising more rapidly than anything else in the economy—more than prices, wages or profits. It isn't only current medical costs, but also the present value of the stream of endlessly high cost increases that retards hiring."

a. Why does the present value of the stream of health care costs, and not just the current health care costs, affect a firm's decision whether to create a new position?

b. Why should an employer discount the future health care costs in its decision whether to create a new position? **A**

2.12 A group of Turkish investors is considering purchasing a firm for TL140 million. The net income of the firm in 2017 was TL5.6 million and financial analysts expect this level of profitability (adjusted for inflation) to continue indefinitely. Calculate the internal rate of return. If the real interest rate was 5%, would it be worthwhile for the investors to buy the firm? (*Hint*: See Solved Problems 16.2 and 16.3.) **A**

2.13 As discussed in the previous question, a group of investors is considering purchasing a firm. The net income of the firm in 2017 was TL5.6 million and financial analysts expect this level of profitability (adjusted for inflation) to continue indefinitely. At what price would it be worthwhile for the investors to buy the firm if the real interest rate was 5%? (*Hint*: See Solved Problems 16.2 and 16.3.) **A**

2.14 On January 15, 2016, the United Kingdom's Competition and Market Authority approved the £12.5 billion takeover of telecommunications company EE Limited (a joint venture between the French company Orange and the German company Deutsche Telekom AG) by the United Kingdom telecommunications company BT Group plc. The regulatory authority found that the merger was not expected to result in a substantial lessening of competition within any markets in the United Kingdom. EE had reported earnings before interest, taxes, depreciation, and amortization of £1,781 million for 2015 (EE Limited, 2016). If EE had been expected to continue to earn this amount each year in the future, what was the internal rate of return for this investment by BT? (*Hint*: See Solved Problem 16.3.) **A**

2.15 As discussed in the previous question, the United Kingdom telecommunications company BT Group plc paid £12.5 billion to acquire the telecommunications company EE Limited in January 2016. EE had reported earnings before interest, taxes, depreciation, and amortization of £1,781 million for 2015. Assuming that EE can maintain this earnings flow indefinitely, would it have been worthwhile for BT to acquire it if the real interest rate was 7%? Why? **A**

2.16 A group that includes former Lakers star Magic Johnson bought the Los Angeles Dodgers for \$2.15 billion in 2012. The Dodgers reported earnings of \$11.3 million (John Gittelsohn and Nadja Brandt, "Dodgers Costing \$2.15 Billion Hinges on Property Return," **http://www.businessweek.com**, April 4, 2012). Financial experts, puzzled by the high price, speculated that the new owners expect to make additional earnings from the surrounding land they purchased. If the real interest rate is 2%, how much do they have to earn annually for this purchase to make sense? (*Hint*: See Solved Problems 16.2 and 16.3.) **A**

2.17 A firm's profit is π = revenue − labor costs − capital costs. Its capital cost can be stated as its internal rate of return on capital, *irr*, times the value of its capital, $p_K K$, where p_K is the price of a unit of capital and K is the number of units of capital. What is the firm's implicit rate of return on its capital? (*Hint*: Set profit equal to zero and solve for the *irr*.) **A**

*2.18 A firm is considering an investment where its cash flow is $\pi_1 = \$1$ (million), $\pi_2 = -\$12$, $\pi_3 = \$20$, and $\pi_t = 0$ for all other t. The interest rate is 7%. Use the net present value rule to determine whether the firm should make the investment. Can the firm use the internal rate of return rule to make this decision? **A**

2.19 A company is considering whether or not to purchase a piece of factory equipment for AU\$500,000. The equipment would last three years, is expected to generate AU\$200,000 of additional annual profit over that period, and can be sold for scrap at the end of the third year for AU\$20,000. What is the present value of this investment if the interest rate is 10%? Should the company purchase the equipment? Should it buy the equipment if the interest rate is 12%? **A**

3. Exhaustible Resources

3.1 You have a barrel of oil that you can sell today for p dollars. Assuming no inflation and no storage cost, how high would the price have to be next year for you to sell the oil next year rather than now? **A**

3.2 Trees, wine, and cattle become more valuable over time and then possibly decrease in value. Draw a figure with present value on the vertical axis and years (age) on the horizontal axis and show this relationship. Show in what year the owner should "harvest" such a good assuming that the cost to harvesting is zero. How would your answer change if the interest rate were zero? Show in a figure. (*Hint*: If the good's present value is P_0 and we take that money and invest it at interest rate i, which is a small number such as 2% or 4%, then its value in year t is $P_0(1 + i)^t$; or if we allow continuous compounding, $P_0 e^{it}$. Such a curve increases exponentially over time and looks like the curve labeled *Price* in Figure 16.2. Draw curves with different possible present values. Use those curves to choose the optimal harvest time.) **A**

3.3 If all the coal in the ground, Q, is to be consumed in two years and the demand for coal is $Q_t = A(p_t)^{-}$ in each year t where is a constant demand elasticity, what is the price of coal each year? **A**

4. Capital Markets, Interest Rates, and Investments

4.1 If the government bars foreign lenders from loaning money to its citizens, how does the capital market equilibrium change?

4.2 In the figure in Solved Problem 16.4, suppose that the government's demand curve remains constant at D_g^1 but the government starts to tax private earnings, collecting 1% of all interest earnings. How does the capital market equilibrium change? What is the effect on private borrowers?

5. Challenge

5.1 Consider the following scenarios, each of which span 34 years. If you go to university, you spend €25,000 per year for four years on school-related expenses (for example, tuition and books) and then earn €50,000 per year for 8 years, €80,000 per year for 15 years, and €70,000 per year for your remaining 7 years before retirement. If you don't go to university, you earn €30,000 per year for 12 years, €45,000 per year for 15 years, and €40,000 per year for your remaining 7 years before retirement. If the interest rate is zero, should you go to college?

If the internal rate of return is 9.7%, should you go to college if the interest rate is 10%? Comparing your two answers, state a simple rule for determining whether you should go to college. Assume that all payments are made at the end of each year. (*Hint*: See the Challenge Solution.) **A**

5.2 At current interest rates, it pays for Bob to go to college if he graduates in four years. If it takes an extra year to graduate from college, does going to college still pay? Show how the figure in the Challenge Solution changes. Illustrate how the present value calculation changes using a formula and variables. **A**

5.3 Which is worth more to you: (a) a \$10,000 payment today or (b) a \$1,000-per-year higher salary for as long as you work? At what interest rate would (a) be worth more to you than (b)? Does your answer depend on how many years you expect to work? **A**

5.4 In 2015, the minimum starting salary for a teacher with four years of training in schools in New South Wales, Australia, was AU\$62,282; for a teacher with five years of training, it was AU\$65,486 (The University of Sydney, n.d.). (For simplicity, assume that these salaries stay constant and do not increase with experience.) To maximize your life-time earnings, is it worthwhile to take one extra year of schooling at a cost of AU\$15,000? In your calculations, assume that in each case, you'll work for 30 years and then retire, and consider interest rates of 3% and 4%. (*Hint*: See Equation 16.3.) **A**

17

Uncertainty

Lottery: A tax on people who are bad at math.

Challenge

BP and Limited Liability

On April 20, 2010, a massive explosion occurred on the Transocean Deepwater Horizon oil rig, which was leased by the oil company BP. The explosion killed 11 workers and seriously injured 17 others. In addition, many of the 90,000 workers who participated in the cleanup suffered significant health problems from exposure to various toxins. Safeguards in the well to automatically cap the oil in case of an accident did not work as expected. Consequently, a massive spill of roughly 200 million gallons of oil polluted the Gulf of Mexico before the well was finally capped. This catastrophic oil spill inflicted gigantic costs for cleaning up Louisiana and other Gulf states and inflicted very large losses on the Gulf fishing and tourism industries.

It does not necessarily follow that because the outcome was bad that BP made bad decisions before the event. BP could have taken reasonable safety precautions and merely been unlucky. However, government agencies concluded that the explosion and the resulting massive oil leak were largely due to a failure to take appropriate safety and other precautions by BP and its subcontractors.

BP may have ignored or underestimated the chance of these expensive calamities, improperly reasoning that such major disasters had not happened to them before and would therefore never happen in the future (or at least that the chances were miniscule). However, a more likely explanation for BP's behavior is that it did not expect to bear the full cost if a catastrophe occurred. In 1990, Congress passed a law that limited liability beyond cleanup costs to $75 million for a rig spill, a tiny fraction of the harm in this case.

In the face of international condemnation for the massive Gulf spill, BP agreed to waive this cap. In 2012, BP pled guilty to 11 counts of seaman's manslaughter and received a record $4 billion fine. In addition, BP was liable for much more due to cleanup costs, civil lawsuits, and other fines and penalties. In 2015, BP struck a $20.8 billion agreement to settle damages with Gulf Coast states and the federal government. That brought its total costs to at least $56 billion—747 times larger than $75 million. These losses are substantial compared to BP shareholders' equity of $106 billion in 2015.

BP made a calculated decision that reflected its estimate of the risk of a catastrophic oil spill, presumably taking the $75 million cap on liability into account. How does a cap on liability affect a firm's willingness to make a risky investment or to invest less than the optimal amount in safety? How does a cap affect the amount of risk that the firm and others in society bear? How does a cap affect the amount of insurance against the costs of an oil spill that a firm buys?

Life's a series of gambles. Will you get a good summer job? Will you avoid disasters such as air crashes, disease, earthquakes, and fire? Will you receive Social Security when you retire? Will you win the lottery tomorrow? Will your stock increase in value?

In this chapter, we extend the model of decision making by individuals and firms to include uncertainty. We look at how uncertainty affects consumption decisions made by individuals and business decisions made by firms.

When making decisions about investments and other matters, consumers and firms consider the possible *outcomes* under various circumstances, or *states of nature*. For example, a pharmaceutical firm's drug may either be approved or rejected by a regulatory authority, so the two states of nature are *approve* or *reject*. Associated with each of these states of nature is an outcome: the value of the pharmaceutical firm's stock will be $100 per share if the drug is approved and only $75 if the drug is rejected.

Although we cannot know with certainty what the future outcome will be, we may know that some outcomes are more likely than others. Often an uncertain situation—one in which no single outcome is certain to occur—can be quantified in the sense that we can assign a probability to each possible outcome. For example, if we toss a coin, we can assign a probability of 50% to each of the two possible outcomes: heads or tails. When uncertainty can be quantified, it is sometimes called **risk**: The likelihood that each possible outcome is known or can be estimated, and no single possible outcome is certain to occur. However, because many people do not distinguish between the terms risk and *uncertainty*, we use these terms interchangeably. All the examples in this chapter concern quantifiable uncertainty.[1]

risk
the likelihood of each possible outcome is known or can be estimated and no single possible outcome is certain to occur

Consumers and firms modify their decisions about consumption and investment as the degree of risk varies. Indeed, most people are willing to spend money to reduce risk by buying insurance or taking preventive measures. Moreover, most people will choose a riskier investment over a less risky one only if they expect a higher return from the riskier investment.

In this chapter, we examine five main topics

1. **Assessing Risk.** Probability, expected value, and variance are important concepts that are used to assess the degree of risk and the likely profit from a risky undertaking.
2. **Attitudes Toward Risk.** Whether managers or consumers choose a risky option over a less risky alternative depends on their attitudes toward risk and on the expected payoffs of each option.
3. **Reducing Risk.** People try to reduce their overall risk by choosing safe rather than risky options, taking actions to reduce the likelihood of bad outcomes, obtaining insurance, pooling risks by combining offsetting risks, and in other ways.
4. **Investing Under Uncertainty.** Whether people make an investment depends on the riskiness of the payoff, the expected return, attitudes toward risk, the interest rate, and whether it is profitable to alter the likelihood of a good outcome.
5. **Behavioral Economics of Uncertainty.** Because some people do not choose among risky options the way that traditional economic theory predicts, some researchers have switched to new models that incorporate psychological factors.

[1]Uncertainty is unquantifiable when we do not know enough to assign meaningful probabilities to different outcomes or if we do not even know what the possible outcomes are. If asked "Who will be the U.S. president in 15 years?" most of us do not even know the likely contenders, let alone the probabilities.

17.1 Assessing Risk

In America, anyone can be president. That's one of the risks you take.
—Adlai Stevenson

Gregg is considering whether to schedule an outdoor concert on July 4th. Booking the concert is a gamble: He stands to make a tidy profit if the weather is good, but he'll lose a substantial amount if it rains.

To analyze this decision, Gregg needs a way to describe and quantify risk. A particular *event*—such as holding an outdoor concert—has a number of possible *outcomes*—here, either it rains or it does not rain. When deciding whether to schedule the concert, Gregg quantifies how risky each outcome is using a *probability* and then uses these probabilities to determine what he can expect to earn.

Probability

A *probability* is a number between 0 and 1 that indicates the likelihood that a particular outcome will occur. If an outcome cannot occur, it has a probability of 0. If the outcome is sure to happen, it has a probability of 1. If it rains one time in four on July 4th, the probability of rain is $\frac{1}{4}$ or 25%.

How can Gregg estimate the probability of rain on July 4th? Usually the best approach is to use the *frequency*, which tells us how often an uncertain event occurred in the past. Otherwise, one has to use a *subjective probability*, which is an estimate of the probability that may be based on other information, such as informal "best guesses" of experienced weather forecasters.

Frequency The probability is the actual chance that an outcome will occur. Gregg does not know the true probability so he has to estimate it. Because Gregg (or the weather department) knows how often it rained on July 4th over many years, he can use that information to estimate the probability that it will rain this year.

He calculates θ (theta), the frequency that it rained, by dividing n, the number of years that it rained on July 4th, by N, the total number of years for which he has data:

$$\theta = \frac{n}{N}.$$

For example, if it rained 20 times on July 4th in the last 40 years, $n = 20$, $N = 40$, and $\theta = 20/40 = 0.5$. Gregg then uses θ, the frequency, as his estimate of the true probability that it will rain this year.

Subjective Probability Unfortunately, we often lack a history of repeated events that allows us to calculate frequencies. For example, the disastrous magnitude 9 earthquake that struck Japan in 2011, with an accompanying tsunami and nuclear reactor crisis, was unprecedented in modern history.

Where events occur very infrequently, we cannot use a frequency calculation to predict a probability. We use whatever information we have to form a *subjective probability*, which is a best estimate of the likelihood that the outcome will occur—that is, our best, informed guess.

The subjective probability can combine frequencies and all other available information—even information that is not based on scientific observation. If Gregg is planning a concert months in advance, his best estimate of the probability of rain is based on the frequency of rain in the past. However, as the event approaches, a weather forecaster can give him a better estimate that takes into account atmospheric conditions and other information in addition to the historical frequency. Because the forecaster's probability estimate uses personal judgment in addition to an observed frequency, it is a subjective probability.

Probability Distribution A *probability distribution* relates the probability of occurrence to each possible outcome. Panel a of Figure 17.1 shows a probability distribution over five possible outcomes: zero to four days of rain per month in a relatively dry city. The probability that it rains on no days during a month is 10%, as is the probability of exactly four days of rain. The chance of two rainy days is 40%, and the chance of one or three rainy days is 20% each. The probability that it rains five or more days a month is 0%.

These weather outcomes are *mutually exclusive*—only one of these outcomes can occur at a given time—and *exhaustive*—no other outcomes than those listed are possible. Where outcomes are mutually exclusive and exhaustive, exactly one of these outcomes will occur with certainty, and the probabilities must add up to 100%. For simplicity, we concentrate on situations with only two possible outcomes.

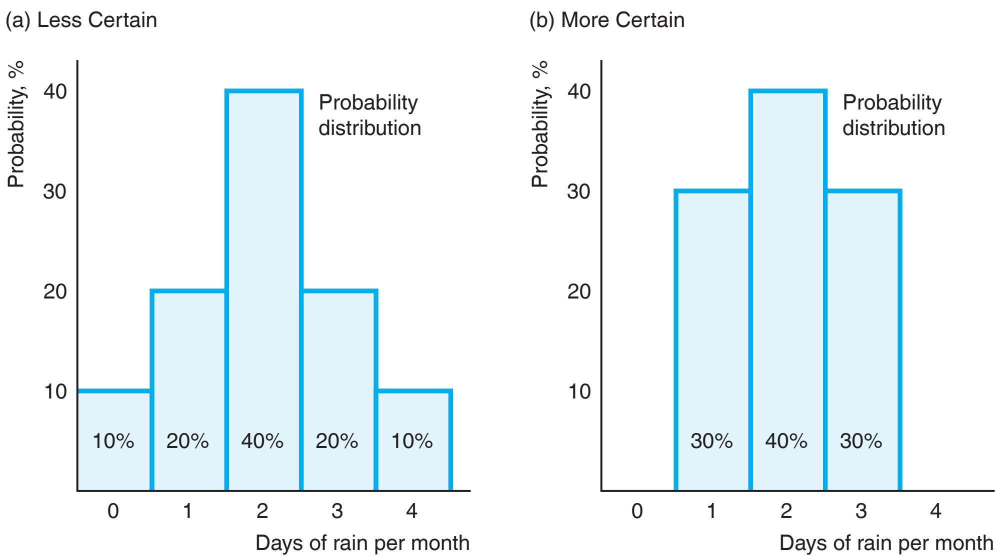
Expected Value

One of the common denominators I have found is that expectations rise above that which is expected. —George W. Bush

Gregg, a promoter, schedules an outdoor concert for tomorrow.[2] How much money he'll make depends on the weather. If it doesn't rain, his profit or value from the

Figure 17.1 Probability Distributions

The probability distribution shows the probability of occurrence for each of the mutually exclusive outcomes. Panel a shows five possible mutually exclusive outcomes. The probability that it rains exactly two days per month is 40%. The probability that it rains five or more days per month is 0%. The probability distributions in panels a and b have the same mean. The variance is smaller in panel b, where the probability distribution is more concentrated around the mean than the distribution in panel a.



concert is $V = 15$. (If it will make you happier—and it will certainly make Gregg happier—you can think of the profits in this example as \$15,000 instead of \$15.) If it rains, he'll have to cancel the concert and he'll lose $V = -5$, which he must pay the band. Although Gregg does not know what the weather will be with certainty, he knows that the weather department forecasts a 50% chance of rain.

Gregg may use the *mean* or the *average* of the values from both outcomes as a summary statistic of the likely payoff from booking this concert. The mean or average value is called his *expected value* (here, his *expected profit*). The expected value, EV, is the value of each possible outcome times the probability of that outcome:[3]

$$\begin{aligned} \mathrm{E}V &= [Pr(\text{no rain}) \times \text{Value(no rain)}] + [Pr(\text{rain}) \times \text{Value(rain)}] \\ &= \left[\tfrac{1}{2} \times 15\right] + \left[\tfrac{1}{2} \times (-5)\right] = 5, \end{aligned}$$

where Pr is the probability of an outcome, so Pr(rain) is the "probability that rain occurs."

The expected value is the amount Gregg would earn on average if the event were repeated many times. If he puts on such concerts many times over the years and the weather follows historical patterns, he will earn 15 at half of the concerts without rain, and he will get soaked for -5 at the other half of the concerts, at which it rains. Thus, he'll earn an average of 5 per concert over time.

Solved Problem 17.1

Suppose that Gregg could obtain perfect information so that he can accurately predict whether it will rain far enough before the concert so that he could book the band only if needed. How much would he expect to earn, knowing that he will eventually have this perfect information? How much does he gain by having this perfect information?

Answer

1. *Determine how much Gregg would earn if he had perfect information in each state of nature.* If Gregg knew with certainty that it would rain at the time of the concert, he would not book the band, so he would make no loss or profit: $V = 0$. If Gregg knew that it would not rain, he would hold the concert and make 15.

2. *Determine how much Gregg would expect to earn before he learns with certainty what the weather will be.* Gregg knows that he'll make 15 with a 50% probability ($= \frac{1}{2}$) and 0 with a 50% probability, so his expected value, given that he'll receive perfect information in time to act on it, is $\mathrm{E}V = \left(\frac{1}{2} \times 15\right) + \left(\frac{1}{2} \times 0\right) = 7.5$.

3. *Calculate his gain from perfect information as the difference between his expected earnings with perfect information and his expected earnings with imperfect information.* Gregg's gain from perfect information is the difference between the

---

[2]My brother Gregg, a successful concert promoter, wants me to inform you that the hero of the following story is some other Gregg who is a concert promoter.

[3]With n possible outcomes, the value of outcome i is V_i, and the probability of that outcome is Pr_i, the expected value is $\mathrm{E}V = Pr_1V_1 + Pr_2V_2 + \cdots + Pr_nV_n$.

expected earnings with perfect information, 7.5, and the expected earnings without perfect information, 5. Thus, Gregg expects to earn 2.5 (= 7.5 − 5) more with perfect information than with imperfect information.[4]

Comment: Having information has no value if it doesn't alter behavior. This information is valuable to Gregg because he avoids hiring the band if he knows that it will rain.

Variance and Standard Deviation

From the expected value, Gregg knows how much he is likely to earn on average if he books many similar concerts. However, the expected value does not summarize how risky a concert is.

If Gregg's earnings are the same whether it rains or not, he faces no risk and the actual return he receives is the expected value. If the possible outcomes differ from one another, he faces risk.

We can measure the risk Gregg faces in various ways. The most common approach is to use a measure based on how much the values of the possible outcomes differ from the expected value, EV. If it does not rain, the *difference* between Gregg's actual earnings, 15, and his expected earnings, 5, is 10. The difference if it does rain is $-5 - 5 = -10$. It is convenient to combine the two differences—one difference for each state of nature (possible outcome)—into a single measure of risk.

One such measure of risk is the *variance,* which measures the spread of the probability distribution. For example, the variance in panel a of Figure 17.1, where the probability distribution ranges from zero to four days of rain per month, is greater than the variance in panel b, where the probability distribution ranges from one to three days of rain per month.

Formally, the variance is the probability-weighted average of the squares of the differences between the observed outcome and the expected value.[5] The variance of the value Gregg obtains from the outdoor concert is

$$\begin{aligned}\text{Variance} &= [Pr(\text{no rain}) \times (\text{Value(no rain)} - \text{E}V)^2] + [Pr(\text{rain}) \\ &\quad \times (\text{Value(rain)} - \text{E}V)^2] \\ &= \left[\tfrac{1}{2} \times (15 - 5)^2\right] + \left[\tfrac{1}{2} \times (-5 - 5)^2\right] \\ &= \left[\tfrac{1}{2} \times (10)^2\right] + \left[\tfrac{1}{2} \times (-10)^2\right] = 100.\end{aligned}$$

Table 17.1 shows how to calculate the variance of the profit from this concert step by step. The first column lists the two outcomes: rain and no rain. The next column gives the probability. The third column shows the value or profit of each outcome. The next column calculates the difference between the values in the third column and the expected value, EV = \$5. The fifth column squares these differences, and the last column multiplies these squared differences by the probabilities in the second column. The sum of these probability-weighted differences, \$100, is the variance.

[4] This answer can be reached directly. The value of this information is his expected savings from not hiring the band when it rains: $\frac{1}{2} \times 5 = 2.5$.

[5] With n possible outcomes, an expected value of EV, a value of V_i for each outcome i, and Pr_i probability of each outcome i, the variance is $Pr_1(V_1 - \text{E}V)^2 + Pr_2(V_2 - \text{E}V)^2 + \cdots + Pr_n(V_n - \text{E}V)^2$. The variance puts more weight on large deviations from the expected value than on smaller ones.

Table 17.1 Variance and Standard Deviation: Measures of Risk

Outcome	Probability	Value	Deviation	Deviation^2	Deviation^2 × Probability
No rain	$\frac{1}{2}$	15	10	100	50
Rain	$\frac{1}{2}$	−5	−10	100	50
				Variance	100
				Standard Deviation	10

Note: Deviation = Value − EV = Value − 5.

Instead of describing risk using the variance, economists and businesspeople often report the *standard deviation*, which is the square root of the variance. The usual symbol for the standard deviation is σ (sigma), so the symbol for variance is σ^2. For the outdoor concert, the variance is $\sigma^2 = \$100$ and the standard deviation is $\sigma = \$10$.

17.2 Attitudes Toward Risk

Given the risks Gregg faces if he schedules a concert, will Gregg stage the concert? To answer this question, we need to know Gregg's attitude toward risk.

Expected Utility

If Gregg did not care about risk, then he would decide whether to promote the concert based on its expected value (profit) regardless of any difference in the risk. However, most people care about risk as well as expected value. Indeed, most people are *risk averse*—they dislike risk. They will choose a riskier option over a less risky option only if the expected value of the riskier option is sufficiently higher than that of the less risky one.

We need a formal way to judge the trade-off between expected value and risk—to determine if the expected value of a riskier option is sufficiently high to justify its greater risk. The most commonly used method extends the model of utility maximization. In Chapter 4, we noted that one can describe an individual's preferences over various bundles of goods by using a utility function. John von Neumann and Oskar Morgenstern (1944) extended the standard utility-maximizing model to include risk.[6] This approach can be used to show how people's taste for risk affects their choices among options that differ in both expected value and risk, such as career choices, the types of contracts to accept, where to build plants, whether to buy insurance, and which stocks to buy.

In this reformulation, we assume that the individual knows the value of each possible outcome and the probability that it will occur. A rational person maximizes *expected utility*. Expected utility is the probability-weighted average of the utility from each possible outcome. For example, Gregg's expected utility, EU, from promoting the concert is

$$\begin{aligned}\mathrm{E}U &= [Pr(\text{no rain}) \times U(\text{Value}(\text{no rain}))] + [Pr(\text{rain}) \times U(\text{Value}(\text{rain}))] \\ &= \left[\tfrac{1}{2} \times U(15)\right] + \left[\tfrac{1}{2} \times U(-5)\right],\end{aligned}$$

[6]This approach to handling choice under uncertainty is the most commonly used method. Schoemaker (1982) discusses the logic underlying this approach, the evidence for it, and several variants. Machina (1989) discusses a number of alternative methods. Here we treat utility as a cardinal measure rather than an ordinal measure.

where his utility function, U, is a function of his earnings. For example, $U(15)$ is the amount of utility Gregg gets from earnings or wealth of 15.[7]

In short, the expected utility calculation is similar to the expected value calculation. Both are weighted averages in which the weights are the probabilities that correspond to the various possible outcomes. The mathematical difference is that the expected value is the probability-weighted average of the monetary value, whereas the expected utility is the probability-weighted average of the utility from the monetary value. The key economic difference is that the expected utility captures the trade-off between risk and value, whereas the expected value considers only value.

If we know how an individual's utility increases with wealth, we can determine how that person reacts to risky situations. We refer to a risky situation as a *bet*. Thus, for example, if Gregg schedules his concert outdoors, he is betting that it will not rain. We can classify people in terms of their willingness to make a **fair bet**: a wager with an expected value of zero. An example of a fair bet is one in which you pay a dollar if a flipped coin comes up heads and receive a dollar if it comes up tails. Because you expect to win half the time and lose half the time, the expected value of this bet is zero:

$$\left[\tfrac{1}{2} \times (-1)\right] + \left[\tfrac{1}{2} \times 1\right] = 0.$$

fair bet
a wager with an expected value of zero

In contrast, a bet in which you pay 2 if you lose the coin flip and receive 4 if you win is an unfair bet that favors you, with an expected value of

$$\left[\tfrac{1}{2} \times (-2)\right] + \left[\tfrac{1}{2} \times 4\right] = 1.$$

risk averse
unwilling to make a fair bet

risk neutral
indifferent about making a fair bet

risk preferring
willing to make a fair bet

Someone who is unwilling to make a fair bet is **risk averse**. A person who is indifferent about making a fair bet is **risk neutral**. A person who is **risk preferring** will make a fair bet.[8]

Risk Aversion

Most people are risk averse. We can use our expected utility model to examine how Irma, who is risk averse, makes a choice under uncertainty. Figure 17.2 shows Irma's utility function. The utility function is concave to the wealth axis, indicating that Irma's utility rises with wealth but at a diminishing rate.[9] She has *diminishing marginal utility of wealth*: The extra pleasure from each extra dollar of wealth is smaller than the pleasure from the previous dollar. An individual whose utility function is concave to the wealth axis is risk averse, as we now illustrate.

Unwillingness to Take a Fair Bet Suppose that Irma has an initial wealth of 40 and can choose between two options. One option is to do nothing and keep the 40, so that her utility is $U(40) = 120$ (the height of point d in Figure 17.2) with certainty. Her other option is to buy a share (a unit of stock) in a start-up company. Her wealth will be 70 if the start-up is a big success and 10 otherwise.

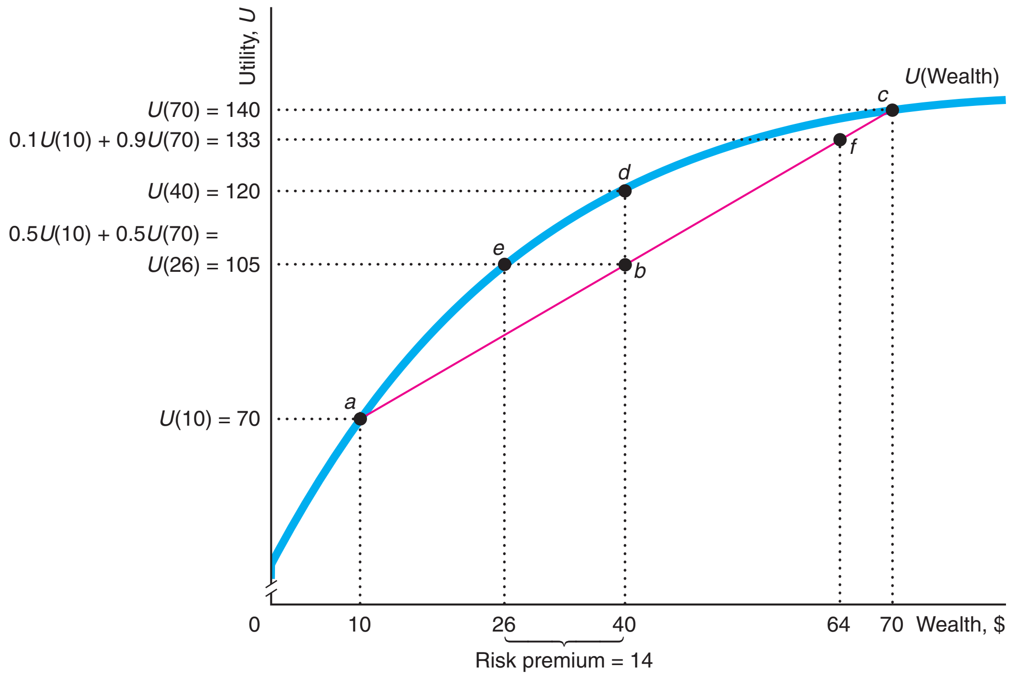
[7]People have preferences over the goods they consume. However, for simplicity, we'll say that a person receives utility from earnings or wealth, which can be spent on consumption goods.

[8]*Jargon alert*: The terms *risk loving* and *risk seeking* are common synonyms for risk preferring.

[9]Irma's utility from W wealth is $U(W)$. She has positive marginal utility from extra wealth, $dU(W)/dW > 0$; however, her utility increases with wealth at a diminishing rate, $d^2U(W)/dW^2 < 0$.

Figure 17.2 Risk Aversion

Initially, Irma's wealth is 40, so her utility is $U(40) = 120$, at point d. If she buys the stock and it's worth 70, her utility is $U(70) = 140$ at point c. If she buys the stock and it's worth only 10, she is at point a, where $U(10) = 70$. If her subjective probability that the stock will be worth 70 is 50%, the expected value of the stock is $40 = (0.5 \times 10) + (0.5 \times 70)$ and her expected utility from buying the stock is $0.5U(10) + 0.5U(70) = 105$, at point b, which is the midpoint of the line between the good outcome, point c, and the bad outcome, point a. Thus, her expected utility from buying the stock, 105, is less than her utility from having a certain wealth of 40, $U(40) = 120$, so she does not buy the stock. In contrast, if Irma's subjective probability that the stock will be worth 70 is 90%, her expected utility from buying the stock is $0.1U(10) + 0.9U(70) = 133$, point f, which is more than her utility with a certain wealth of 40, $U(40) = 120$, at d, so she buys the stock.



Irma's subjective probability is 50% that the firm will be a big success. Her expected value of wealth remains

$$40 = \left(\tfrac{1}{2} \times 10\right) + \left(\tfrac{1}{2} \times 70\right).$$

Thus, buying the stock is a fair bet because she has the same expected wealth whether she purchases the stock or not.

If Irma were risk neutral so that she only cared about her expected value and didn't care about risk, she would be indifferent between buying the stock or not. However, because Irma is risk averse, Irma prefers not buying the stock because both options have the same expected wealth and buying the stock carries more risk.

We can show that her expected utility is lower if she buys the stock than if she does not. If she buys the stock, her utility if the stock does well is $U(70) = 140$, at point c. If it doesn't do well, her utility is $U(10) = 70$, at point a. Thus, her expected utility from buying the stock is

$$\left[\tfrac{1}{2} \times U(10)\right] + \left[\tfrac{1}{2} \times U(70)\right] = \left[\tfrac{1}{2} \times 70\right] + \left[\tfrac{1}{2} \times 140\right] = 105.$$

Her expected utility is the height of point *b*, which is the midpoint of a line between points *a* and *c*. Because Irma's utility function is concave, her utility from certain wealth, 120 at point *d*, is greater than her expected utility from the risky activity, 105 at point *b*. As a result, she does not buy the stock. Buying the stock, which is a fair bet, increases the risk she faces without changing her expected wealth. Thus, Irma, because her utility function is concave, prefers not to take a fair bet and is risk averse. *A person whose utility function is concave picks the less risky choice if both choices have the same expected value.*

A risk-averse person chooses a riskier option only if it has a sufficiently higher expected value. Given her wealth of \$40, if Irma were much more confident that the stock would be valuable, her expected value would rise and she'd buy the stock, as Solved Problem 17.2 shows.

Solved Problem 17.2

MyLab Economics Solved Problem

Suppose that Irma's subjective probability is 90% that the stock will be valuable. What is her expected wealth if she buys the stock? What is her expected utility? Does she buy the stock?

Answer

1. *Calculate Irma's expected wealth*. Her expected value or wealth is 10% times her wealth if the stock bombs plus 90% times her wealth if the stock does well:

$$(0.1 \times 10) + (0.9 \times 70) = 64.$$

In Figure 17.2, 64 is the distance along the wealth axis corresponding to point *f*.

2. *Calculate Irma's expected utility*. Her expected utility is the probability-weighted average of her utility under the two outcomes:

$$[0.1 \times U(10)] + [0.9 \times U(70)] = [0.1 \times 70] + [0.9 \times 140] = 133.$$

Her expected utility is the height on the utility axis of point *f*. Point *f* is nine-tenths of the distance along the line connecting point *a* to point *c*.

3. *Compare Irma's expected utility to her certain utility if she does not buy the stock*. Irma's expected utility from buying the stock, 133 (at point *f*), is greater than her certain utility, 120 (at point *d*), if she does not. Thus, if Irma is this confident that the stock will do well, she buys it. Although the risk is greater from buying than from not buying, her expected wealth is sufficiently higher (64 instead of 40) that it's worth it to her to take the chance.

risk premium
the maximum amount that a risk-averse person would pay to avoid taking a risk

The Risk Premium The **risk premium** is the maximum amount that a decision maker would pay to avoid taking a risk. Equivalently, the risk premium is the minimum extra compensation (premium) that a decision maker would require to willingly incur a risk.

We can use Figure 17.2, where Irma owns the stock that has a 50% chance of being worth 70 and a 50% chance of being worth 10, to determine her risk premium. The risk premium is the difference between her expected wealth from the risky stock and the amount of wealth, called the *certainty equivalent*, that if held with certainty, would yield the same utility as this uncertain prospect.

Irma's expected wealth from holding the stock is 40, and the corresponding expected utility is 105. The certainty equivalent income is 26, because Irma's utility is 105 if she has 26 with certainty: $U(26) = 105$, which is the same as her expected utility from owning the stock. Thus, she would be indifferent between keeping the stock or selling it for a price of 26. Thus, her risk premium, the difference between the expected value of the uncertain prospect and the certainty equivalent, is $40 - 26 = 14$, as the figure shows.

Application

Stocks' Risk Premium

The value of most stocks is more variable over time than are bonds. Because stocks are riskier than bonds, for both to sell in the market to risk-averse investors, the rates of return on investing in stocks must exceed those on bonds over the period that the investor plans to hold these investments. This greater return is an investor's risk premium for stocks.

For example, a U.S. government bond is essentially free of any risk that the U.S. government will default. As Figure 17.2 illustrates, an investor will buy a stock only if it provides a risk premium over a risk-free U.S. government bond. That is, the investor buys the stock only if the expected return on the stock exceeds the rate of return on the bond.

In 2014, the stocks in the Standard and Poor's index of 500 leading stocks, the S&P 500, had a return of 13.5%, which exceeded the 10.8% return on 10-year U.S. treasury bonds. However, stocks do not always outperform safe government bonds. In 2015, the S&P 500 return, 1.4%, was virtually the same as for bonds, 1.3%. Moreover, stocks performed much worse than bonds in 2008 (−36.6% versus 20.1%) and in 2011 (2.1% versus 16.0%).

Nonetheless, stocks have had a higher rate of return over longer periods. For the 50-year period 1966–2015, the average annual return was 11.0% for S&P 500 stocks and 7.1% on long-term bonds.[10]

Risk Neutrality

Someone who is risk neutral is indifferent about taking a fair bet. Such a person has a constant marginal utility of wealth: Each extra dollar of wealth raises utility by the same amount as the previous dollar. With a constant marginal utility of wealth, the utility function is a straight line in a graph of utility against wealth. Consequently, a risk-neutral person's utility depends only on wealth and not on risk.

Suppose that Irma is risk neutral and has the straight-line utility function in panel a of Figure 17.3. She would be indifferent between buying the stock and receiving 40 with certainty if her subjective probability is 50% that it will do well. Her expected utility from buying the stock is the average of her utility at points *a* (10) and *c* (70):

$$\left[\tfrac{1}{2} \times U(10)\right] + \left[\tfrac{1}{2} \times U(70)\right] = \left[\tfrac{1}{2} \times 70\right] + \left[\tfrac{1}{2} \times 140\right] = 105.$$

Her expected utility exactly equals her utility with certain wealth of 40 (at point *b*) because the line connecting points *a* and *c* lies on the utility function and point *b* is the midpoint of that line.

Here Irma is indifferent between buying and receiving 40 with certainty, a fair bet, because she doesn't care how much risk she faces. Because the expected wealth from both options is 40, she is indifferent between them.

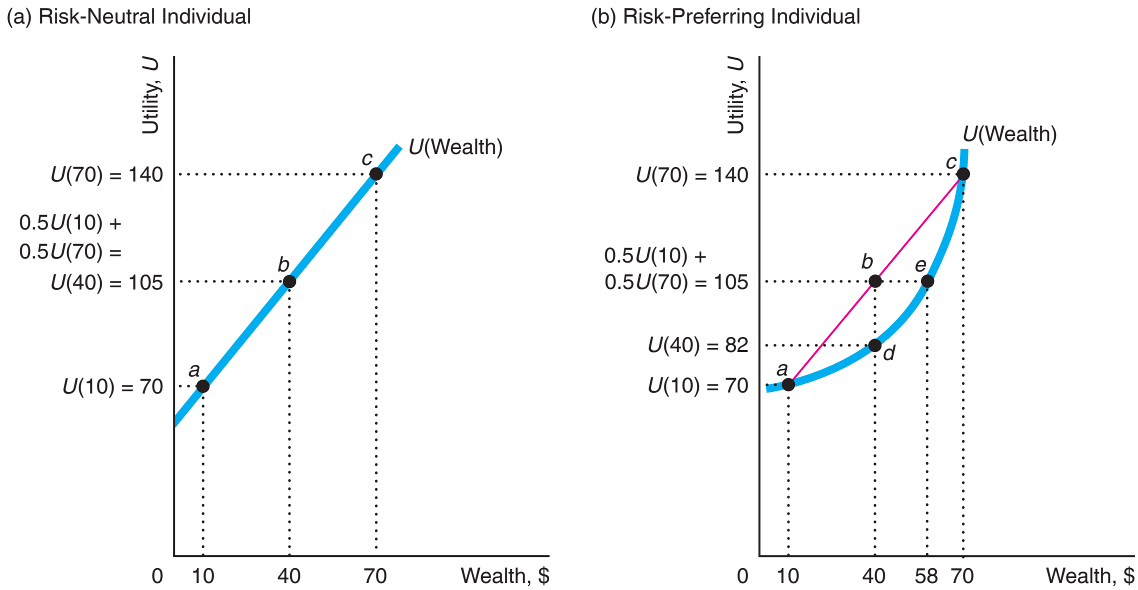
In general, *a risk-neutral person chooses the option with the highest expected value, because maximizing expected value maximizes utility*. A risk-neutral person

---

[10]Adjusting for inflation, which averaged 4.2% per year over this period, the average annual real rates of return were 6.9% on the S&P 500 and 3.0% on long-term U.S. government bonds.

Figure 17.3 Risk Neutrality and Risk Preference

(a) If Irma's utility function can be graphed as a straight line, she is risk neutral and is indifferent as to whether or not to make a fair bet. Her expected utility from buying the stock, 105 at *b*, is the same as from a certain wealth of 40 at *b*. (b) If the plot of Irma's utility function is convex to the horizontal axis, Irma has increasing marginal utility of wealth and is risk preferring. She buys the stock because her expected utility from buying the stock, 105 at *b*, is higher than her utility from a certain wealth of 40, 82 at *d*.



chooses the riskier option if it has even a slightly higher expected value than the less risky option. Equivalently, the risk premium for a risk-neutral person is zero.

Risk Preference

An individual with an increasing marginal utility of wealth is risk preferring and is happy to take a fair bet. If Irma has the utility function in panel b of Figure 17.3, she is risk preferring. Her expected utility from buying the stock, 105 at *b*, is higher than her certain utility if she does not buy the stock, 82 at *d*. Therefore, she buys the stock.

A risk-preferring person is willing to pay for the right to make a fair bet (a negative risk premium). As the figure shows, Irma's expected utility from buying the stock is the same as the utility from a certain wealth of 58. Given her initial wealth of 40, if you offer her the opportunity to buy the stock or offer to give her 18, she is indifferent. With any payment smaller than 18, she prefers to buy the stock.

Application

Gambling

Most people say that they don't like bearing risk.[11] Consistent with such statements, most consumers purchase car, homeowner's, medical, and other forms of insurance that reduce the risks they face. But many of these same people gamble.

[11]The following discussion of gambling is not intended to encourage you to gamble. For example, the important thing to remember about a Powerball lottery is that the probability of winning if you buy a ticket (about 1 in 185 million) is almost exactly the same as the probability of winning if you don't buy a ticket: zero.

According to one estimate, global gaming revenues were \$91.5 billion in 2015. Over half of the countries in the world have lotteries.

Not only do many people gamble, but they make unfair bets, in which the expected value of the gamble is negative. That is, if they play the game repeatedly, they are likely to lose money in the long run. For example, the British government keeps half of the total amount bet on its lottery. Americans lose 7% of all legal money bet.

According to a *Wall Street Journal* study, Internet gamblers win money 30% of the days they wager, but only 11% were in the black over two years (and most of those were ahead by less than \$150). Of the top 10% most frequent bettors, 95% lost money.

Why do people take unfair bets? Some people gamble because they are risk preferring or because they have a compulsion to gamble.[12] However, neither of these observations is likely to explain noncompulsive gambling by most people who exhibit risk-averse behavior in the other aspects of their lives, such as buying insurance. Risk-averse people may make unfair bets because they get pleasure from participating in the game or because they falsely believe that the gamble favors them.

The first explanation is that gambling provides entertainment as well as risk. Risk-averse people insure their property, such as their homes, because they do not want to bear the risk of theft, flooding, and fire. However, these same people may play poker or bet on horse races because they get enough pleasure from playing those games to put up with the financial risk and the expected loss.

Many people definitely like games of chance. One survey found that 65% of Americans say that they engage in games of chance even when the games involve no money or only trivial sums. That is, they play because they enjoy the games. The anticipation of possibly winning and the satisfaction and excitement arising from a win generate greater benefits than the negative feelings associated with a loss.

Instead, or in addition, people may gamble because they make mistakes.[13] Either people do not know the true probabilities or cannot properly calculate expected values, so they do not realize that they are participating in an unfair bet. And some gamblers are simply overconfident: they overestimate their likelihood of winning.

17.3 Reducing Risk

If most accidents occur at home, why not move away?

Risk-averse people want to eliminate or reduce the risks they face. Risk-neutral people avoid unfair bets that are stacked against them, and even risk-preferring people avoid very unfair bets. Individuals can avoid optional risky activities, but often they can't

[12] Friedman and Savage (1948) suggest that some gamblers are risk averse with respect to small gambles but risk preferring for large ones, such as a lottery.

[13] Economists, who know how to calculate expected values and derive most of their excitement from economic models, are apparently less likely to gamble than are other people. A number of years ago, an association of economists met in Reno, Nevada. Reno hotels charge low room rates on the assumption that they'll make plenty from guests' gambling losses. However, the economists gambled so little that they were asked pointedly not to return.

escape risk altogether. Property owners, for instance, always face the possibility that their property will be damaged or stolen. However, they may be able to reduce the harm they suffer if a bad event occurs by using information, diversifying, or buying insurance.

Just Say No

The simplest way to avoid risk is to abstain from optional risky activities. No one forces you to bet on the lottery, go into a high-risk occupation, or buy stock in a start-up biotech firm. If one brand of a product you use comes with a warranty and an otherwise comparable brand does not, you lower your risk by buying the guaranteed product.

Even when you can't avoid risk altogether, you can take precautions to reduce the probability of bad states of nature happening or the magnitude of any loss that might occur. By maintaining your car as the manufacturer recommends, you reduce the probability that it will break down. By locking your apartment door, you lower the chance of having your television stolen. Getting rid of the four-year-old collection of newspapers in your basement lessens the likelihood that your house will burn down. In 2015, DuPont Pioneer introduced a corn seed that it claims is more drought resistant than traditional seeds, which farmers can use to reduce their risk from drought.

Obtain and Use Information

Collecting accurate information before acting is one of the most important ways in which people can reduce risk and increase expected value and expected utility, as Solved Problem 17.1 illustrated. Armed with information, you may avoid a risky choice or you may be able to take actions that reduce the probability of a disaster or the size of the loss.

Before buying a car or refrigerator, many people read *Consumer Reports* to determine how frequently a particular brand is likely to need repairs. Similarly, before deciding where to locate a new plant, a prudent manager collects information about various locations concerning local crime rates, fire risks, and other potential hazards.

Diversify

Although it may sound paradoxical, individuals and firms often reduce their overall risk by making many risky investments instead of only one. This practice is called *risk pooling* or *diversification*. As your grandparents may have advised, "Don't put all your eggs in one basket."

Correlation and Diversification The extent to which diversification reduces risk depends on the degree to which the payoffs of various investments are correlated or move in the same direction.[14] If two investments are positively correlated, one

[14]A measure of the *correlation* between two random variables x and y is

$$\rho = \mathrm{E}\left(\frac{x - \bar{x}}{\sigma_x}\frac{y - \bar{y}}{\sigma_y}\right),$$

where the $\mathrm{E}(\cdot)$ means "take the expectation" of the term in parentheses, $\bar{x}$ and $\bar{y}$ are the means (expected values), and σ_x and σ_y are the standard deviations of x and y. This correlation can vary between -1 and 1. If $\rho = 1$, these random variables are perfectly positively correlated; if $\rho = -1$, they have a perfect negative correlation; and if $\rho = 0$, they are uncorrelated.

performs well when the other performs well. If two investments are negatively correlated, when one performs well, the other performs badly. If the performances of two investments move *independently*—do not move together in a predictable way—their payoffs are uncorrelated.

Diversification can eliminate risk if the returns to two investments are perfectly negatively correlated. Suppose that two firms are competing for a government contract and have an equal chance of winning it. Because only one firm can win, if one wins the other must lose. You can buy a share of stock in either firm for \$20. The stock of the firm that wins the contract will be worth \$40, whereas the stock of the loser will be worth \$10. Investments in these stocks have a perfect negative correlation. If one stock turns out to have the high value, 40, the other must have the low value, 10, and vice versa.

If you buy two shares of the same company, your shares are going to be worth either 80 or 20 after the contract is awarded. Thus, their expected value is

$$EV = \left(\tfrac{1}{2} \times 80\right) + \left(\tfrac{1}{2} \times 20\right) = 50$$

with a variance of

$$\sigma^2 = \left[\tfrac{1}{2} \times (80 - 50)^2\right] + \left[\tfrac{1}{2} \times (20 - 50)^2\right] = 900.$$

However, if you buy one share of each, your two shares will be worth \$50 no matter which firm wins, and the variance is zero—the risk has been completely eliminated by investing in these negatively correlated stocks.

For diversification to reduce risk, it is not necessary for the investments to have a perfect negative correlation. Indeed, it is not even necessary for the investments to have a negative correlation. Diversification reduces risk even if the two investments are uncorrelated or imperfectly positively correlated, although the risk reduction is not as dramatic as in the case of perfect negative correlation, where the risk can be completely eliminated.

Suppose, for example, that each of the two firms has a 50% chance of getting a government contract, but whether one firm gets a contract does *not* affect whether the other firm wins one. Thus, the stock values of the two firms are uncorrelated and each firm's stock price has an equal probability of being 40 or 10. The probability that both firms win contracts and have a stock price of 40 is $\frac{1}{4}$, the chance that one is worth 40 and the other is worth 10 is $\frac{1}{2}$, and the chance that each is worth 10 is $\frac{1}{4}$. If you buy one share of each firm, the expected value of these two shares is

$$EV = \left(\tfrac{1}{4} \times 80\right) + \left(\tfrac{1}{2} \times 50\right) + \left(\tfrac{1}{4} \times 20\right) = 50,$$

and the variance is

$$\sigma^2 = \left[\tfrac{1}{4} \times (80 - 50)^2\right] + \left[\tfrac{1}{2} \times (50 - 50)^2\right] + \left[\tfrac{1}{4} \times (20 - 50)^2\right] = 450.$$

This expected value is the same as from buying two shares of one firm, but the variance is only half as large. Thus, diversification lowers risk when the values are uncorrelated.

Diversification can reduce risk even if the investments are positively correlated provided that the correlation is not perfect. *Diversification does not reduce risk if two investments have a perfect positive correlation.* For example, if the government awards contracts only to both firms or to neither firm, the risks are perfectly positively correlated. The expected value of the stocks and the variance are the same whether you buy two shares of one firm or one share of each firm.

Diversification Through Mutual Funds Given that the value of the stock of most firms is not perfectly positively correlated with the value of other stocks, buying stock in several companies reduces risk compared to buying stock in only one company.

One way that investors can effectively own shares in a number of companies at once is by buying shares in a *mutual fund* of stocks. A mutual fund share is issued by a company that buys stocks in many other companies.

Several mutual funds are based on the *Standard & Poor's Composite Index of 500 Stocks* (S&P 500), which is a market value-weighted average of 500 large firms' stocks. The S&P 500 companies constitute only about 7% of all the publicly traded firms in the United States, but they represent approximately 80% of the total value of the U.S. stock market. A number of mutual funds cover even more stocks. The *Wilshire 5000 Index Portfolio* initially covered 5,000 stocks but now includes many more, as Wilshire seeks to include nearly all U.S. publicly traded stocks and adds new stocks as they are issued. Other mutual funds are based on bonds or on a mixture of stocks, bonds, and other types of investments.

Mutual funds allow investors to reduce the risk associated with uncorrelated price movements across stocks through diversification. However, a stock mutual fund has a *market-wide risk*—a risk that is common to the overall market—which arises because the prices of almost all stocks tend to rise when the economy is expanding and to fall when the economy is contracting. Buying a diversified stock mutual fund does not protect you against the systematic risks associated with shifts in the economy that have a similar effect on most stocks.

Application

Failure to Diversify

Foolishly, many corporate employees fail to diversify their portfolios. Much of their wealth is tied up in their employer's stock. Managers and other corporate employees may receive stock bonuses, which they do not sell. For others, their employer matches their investment in the company's 401(k) retirement plans with company stock.[15] Others invest voluntarily as a sign of loyalty.

If the firm fails, these employees not only lose their jobs but much of the value of their retirement portfolio, as happened to many of Radio Shack's employees when it declared bankruptcy in 2015. That's typical. Duan et al. (2015) analyzed 20 years of data for 729 troubled, large, publicly traded companies. They found that employees kept the amounts of money they held in company stock relatively stable during periods of trouble.

In 2007, at the beginning of the Great Recession, nearly two of every five employees participating in 401(k) retirement plans in large firms held 20% or more of their money in employer stock.[16] About one-sixth of participants invested 50% or more. On average, these funds held 16% in company stock.

It is very risky for workers to hold such a high percentage of their employer's stock. At the investment firm Bear Stearns, employees owned one-third of the company's stock as the firm faced bankruptcy in early 2008. The U.S. government bailed it out, but its 2008 stock value fell to 10% of its December 2007 value. If a Bear Stearns employee's 401(k) had invested $100,000 in a Standard & Poor's 500-stock index fund at the end of 2007, its value would have fallen to $90,760 by the end of the first quarter of 2008. However, if that employee shifted 16% into Bear Stearns stock, the investment would have fallen to slightly less than $77,838. Even worse,

---

[15]A 401(k) plan is a retirement program run by a firm for its employees. By investing in a 401(k), employees can defer paying taxes on their investment and returns until they start withdrawing funds after they reach 59.5 years.

[16]A 401(k) plan is a retirement program run by a firm for its employees under Section 401 (paragraph k) of the Internal Revenue Code. Employees defer paying taxes on investment returns provided they do not start withdrawing income until after age 59 $\frac{1}{2}$.

if all the funds had been in Bear Stearns stock, the 401(k) would have been worth only $10,000.

Consequently, many investment advisors recommend investing at most 5% in employer stock. The share of company stock in pension plans has decreased over time as participations become more educated and some firms restrict employees' ability to invest in company stock in their pension plans. In 2013, only 39% of firms offered their company stock as an investment option in their employees' retirement plans. According to *Pensions & Investments*' survey of the largest U.S. retirement plans, the average share of pension assets in company stock dropped from 26.1% in 2005 to 15.6% in 2014.

Buy Insurance

I detest life-insurance agents; they always argue that I shall some day die, which is not so. —Stephen Leacock

Individuals and organizations can also avoid or reduce risk by purchasing insurance. As we've already seen, a risk-averse person is willing to pay money—a risk premium—to avoid risk. The demand for risk reduction is met by insurance companies, which bear the risk for anyone who buys an insurance policy. Many risk-averse individuals and firms buy insurance, leading to an industry of enormous size: Global insurance revenues exceeded $4.613 trillion in 2012, over 6% of world GDP.[17]

Determine the Amount of Insurance to Buy Many individuals and firms buy insurance to shift some or all of the risk they face to an insurance company. A risk-averse person or firm pays a premium to the insurance company, and the insurance company transfers money to the policyholder if a bad outcome occurs, such as becoming ill, having an accident, or suffering a property loss due to theft or fire.

Because Scott is risk averse, he wants to insure his store, which is worth 500. The probability that his store will burn next year is 20%. If a fire occurs, the store will be worth nothing.

With no insurance, the expected value of his store is

$$EV = (0.2 \times 0) + (0.8 \times 500) = 400.$$

Scott faces a good deal of risk. The variance of the value of his store is

$$\sigma^2 = [0.2 \times (0 - 400)^2] + [0.8 \times (500 - 400)^2] = 40{,}000.$$

fair insurance
a contract between an insurer and a policyholder in which the value of the contract to the policyholder is zero

Suppose that an insurance company offers **fair insurance**: a contract between an insurer and a policyholder in which the expected value of the contract to the policyholder is zero. That is, the insurance is a fair bet. With fair insurance, for every 1 dollar that Scott pays the insurance company, the *insurance premium*, the company will pay Scott 5 dollars to cover the damage if the fire occurs, so that he has 1 dollar less if the fire does not occur, but 4 (= 5 − 1) dollars more if it does occur.[18]

Because Scott is risk averse and the insurance is fair, he wants to *fully insure* by buying enough insurance to eliminate his risk altogether. That is, he wants to buy the

[17]Based on **http://www.insurancejournal.com/news/international/2013/06/26/296846.htm** and **http://www.plunkettresearch.com/insurance-risk-management-market-research/industry-overview**.

[18]Following standard practice in the insurance industry, we use the term *insurance premium* (or just *premium*) in this section to refer to the amount *actually paid* for insurance. The *insurance premium* is different from a *risk premium*, which is a person's *willingness to pay* to avoid risk.

amount of fair insurance that will leave him equally well off in both states of nature. He pays a premium of x so that he has $500 - x$ if the fire does not occur, and has $4x$ if the fire occurs, such that $500 - x = 4x$, or $x = 100$.[19] If a fire does not occur, he pays a premium of 100 and has a store worth 500 for a net value of 400. If a fire does occur, Scott pays 100 but receives 500 from the insurance company for a net value of 400. Thus, Scott's wealth is 400 in either case.

Although Scott's expected value with full and fair insurance is the same as his expected value without insurance, the variance he faces drops from 40,000 without insurance to 0 with insurance. Scott is better off with full fair insurance because he has the same expected value and faces no risk. A risk-averse person always wants full insurance if the insurance is fair.

Sometimes insurance companies put limits on the amount of insurance offered. For example, the insurance company could offer Scott fair insurance but only up to a maximum gross payment of, for example, 400 rather than 500. Given this limit, Scott would buy the maximum amount of fair insurance that he could.

Solved Problem 17.3

MyLab Economics Solved Problem

The local government collects a property tax of 20 on Scott's home. If the tax is collected whether or not the home burns, how much fair insurance does Scott buy? If the tax is collected only if the home does not burn, how much fair insurance does Scott buy?

Answer

1. *Determine the after-tax expected value of the house with and without insurance.* If the tax is always collected, the house is worth $480 = 500 - 20$ if it does not burn and -20 if it does burn. Thus, the expected value of the house is

$$380 = [0.2 \times (-20)] + [0.8 \times 480].$$

If the tax is collected only if the fire does not occur, the expected value of the house is

$$384 = [0.2 \times 0] + [0.8 \times 480].$$

2. *Calculate the amount of fair insurance Scott buys if the tax is always collected.* Because Scott is risk averse, he wants to fully insure so that the after-tax value of his house is the same in both states of nature. If the tax is always collected, he pays a premium of x such that $500 - x - 20 = 4x - 20$, so $x = 100$. If no fire occurs, his net wealth is $500 - 100 - 20 = 380$. If a fire occurs, the insurance company pays 500, or a net payment of 400 above the cost of the insurance, and Scott pays 20 in taxes, leaving him with 380 once again. That is, he buys the same amount of insurance as he would without any taxes. The tax has no effect on his insurance decision because he owes the tax regardless of the state of nature.

3. *Calculate the amount of fair insurance Scott buys if the government collects the tax only if a fire occurs.* With this tax rule, Scott pays a premium of x such that $500 - x - 20 = 4x$, so $x = 96$. Scott pays the insurance company 96 and receives 480 if a fire occurs. Without a fire, Scott's wealth is

[19]The expected value of Scott's insurance contract is $[0.8 \times (-100)] + [0.2 \times 400] = 0$, which shows that the insurance is fair.

$500 - 96 - 20 = 384$. If a fire occurs, the insurance company pays 480, so Scott's wealth is $480 - 96 = 384$. Thus, he has the same after-tax wealth in both states of nature.

Comment: Because the tax system is partially insuring Scott by dropping the tax in the bad state of nature, he purchases less private insurance, 480, than the 500 he buys if the tax is collected in both states of nature.

Fairness and Insurance When fair insurance is offered, risk-averse people fully insure. If insurance companies charge more than the fair-insurance price, individuals buy less insurance.[20]

Because insurance companies do not offer fair insurance, most people do not fully insure. An insurance company could not stay in business if it offered fair insurance. With fair insurance, the insurance company's expected payments would equal the amount the insurance company collects. Because the insurance company has operating expenses—costs of maintaining offices, printing forms, hiring sales agents, and so forth—an insurance firm providing fair insurance would lose money. Insurance companies' rates must be high enough to cover their operating expenses, so the insurance is less than fair to policyholders.

How much can insurance companies charge for insurance? A monopoly insurance company could charge an amount up to the risk premium a person is willing to pay to avoid risk. For example, in Figure 17.2, Irma's risk premium is 14. She would be willing to pay up to \$14 for an insurance policy that would compensate her if her stock did not perform well. The more risk averse an individual is, the more a monopoly insurance company can charge. If many insurance companies compete for business, the price of an insurance policy is less than the maximum that risk-averse individuals are willing to pay—but still high enough that the firms can cover their operating expenses.

Application

Flight Insurance

That airline that doesn't kill me makes me stronger.

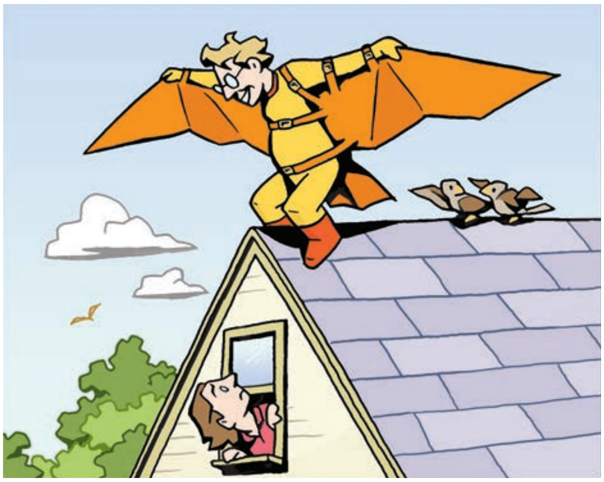
Many folks fear flying: "If flying is so safe, why do they call the airport the terminal?" Catering to these fears, many companies such as Travel Guard (TG) offer accidental death insurance of individual flights. If, just before I take my next regularly scheduled commercial flight, I pay TG \$25 and I die on that flight, TG will pay my family \$500,000. (Although I can get much larger amounts of air travel insurance, it seems a bad idea to make myself worth more to my family dead than alive.)

What are the chances of a given flight crashing? Given that probability, is TG's insurance fair?

If θ is my probability of dying on a flight, my family's expected value from this bet with TG is $[\theta \times 500{,}000] + [(1 - \theta) \times (-25)]$. For this insurance to be fair, this expected value must be zero, which is true if $\theta \approx 0.000115$. That is, one in every 8,696 passengers dies.

How great *is* the danger of being in a fatal commercial airline crash? According to the National Transportation Safety Board, no fatalities occurred on scheduled U.S.

[20]As Solved Problem 17.3 shows, tax laws may act to offset this problem, so that some insurance may be fair or more than fair after tax.



But dear! Flying is safer than driving.

commercial domestic airline flights in 2002, 2007, 2008, and from 2010 through at least June 2016.

In 2001, the probability was much higher than any other year because of the 525 on-board deaths caused by the terrorist hijacking and crashes on September 11 and the subsequent sharp reduction in the number of flights. However, even in 2001, the probability was 0.00000077, or 1 in 1.3 million fliers—much lower than the probability that makes TG's insurance a fair bet. From January 2005 through April 2016, the probability was 0.0000000103 or one fatality per 97 million passengers.

Given that probability, if I fly each day for 100 years, the probability of avoiding a fatal crash is 99.96%. The probability drops to 99.62% after flying every day for 1,000 years and to 96.31% after 10,000 years of flying every day. (For most people, the greatest risk of an airplane trip is the drive to and from the airport. More than twice as many people are killed in vehicle–deer collisions than in plane crashes.)

Given that my chance of dying in a fatal crash is $\theta = 0.0000000103$, the fair rate to pay for \$500,000 of flight insurance is about 0.52¢. Thus, TG is offering to charge me 4,854 times more than the fair rate for this insurance.

I would have to be incredibly risk averse to take TG up on their kind offer. Even if I were that risk averse, I would be much better off buying general life insurance, which is much less expensive than flight insurance and covers accidental death from all types of accidents and diseases.

Insurance Only for Diversifiable Risks Why is an insurance company willing to sell policies and take on risk? By pooling the risks of many people, the insurance company can lower its risk much below that of any individual. If the probability that one car is stolen is independent of whether other cars are stolen, the risk to an insurance company of insuring one person against theft is much greater than the average risk of insuring many people.

An insurance company sells policies only for risks that it can diversify. If the risks from disasters to its policyholders are highly positively correlated, an insurance company is not well diversified by holding many policies. A war affects all policyholders, so the outcomes that they face are perfectly correlated. Because wars are *nondiversifiable risks*, insurance companies do not offer policies insuring against wars.

Application

Limited Insurance for Natural Disasters

Year after year, the most costly insurance losses are due to natural disasters. In recent years, many insurance companies have started viewing some major natural disasters as nondiversifiable risks because such catastrophic events cause many insured people to suffer losses at the same time. As people build more homes in areas where damage from storms or earthquakes is likely, the size of the potential losses to insurers from nondiversifiable risks has grown.

Insurers paid out \$12.5 billion in claims to residential homeowners after the 1994 Los Angeles earthquake. Farmers Insurance Group reported that it paid out three times more for the Los Angeles earthquake than it had collected in earthquake premiums over the previous 30 years.

According to some estimates, Hurricane Katrina in 2005 caused $100 to $200 billion worth of damage and major loss of life. Private insurers paid out $41 billion, or between 20.5% and 41% of the total damages.

Japan's 2011 magnitude 9 earthquake and the associated tsunami was the most costly in history, with estimated damages of between $200 and $350 billion. However, the insurance industry paid only $35 billion, or no more than 17%, because of Japan's low levels of earthquake insurance protection. Two major Japanese earthquakes in 2016 caused at least $10 billion in damage, but insurance covered only about $2 billion.

Of the 353 disasters in 2015, 198 were natural catastrophes. Combined losses were $92 billion, of which $80 billion were due to natural catastrophes. Only 40% of these global economic losses were covered by private or government-sponsored insurance. Moreover, for the largest disaster that year, the 7.8 magnitude Nepal earthquake, insurance covered only 2% of the $10 billion (plus) economic loss.

Insurance companies now refuse to offer hurricane or earthquake insurance in many parts of the world for these relatively nondiversifiable risks. When Nationwide Insurance Company announced that it was sharply curtailing sales of new policies along the Gulf of Mexico and the eastern seaboard from Texas to Maine, a company official explained, "Prudence requires us to diligently manage our exposure to catastrophic losses."

The U.S. government has partially replaced private insurers. The National Flood Insurance Program insures Americans, primarily in Texas and Florida, against floods associated with hurricanes, tropical storms, heavy rains, and other conditions. However, Congress has not provided consistent funding for this program, which lapsed at least four times in 2010 alone. Thus, consumers may not be able to count on federal flood insurance always being available.[21]

In some high-risk areas, state-run insurance pools—such as the Florida Joint Underwriting Association and the California Earthquake Authority—provide households with insurance. However, not only do these policies provide less protection, but their rates are often three times more expensive than previously available commercial insurance, and they provide compensation only for damages beyond a specified level, called the *deductible*.

17.4 Investing Under Uncertainty

Attitudes toward risk affect people's willingness to invest. In some case, investors can pay to alter their probabilities of success.

In the following examples, the owner of a monopoly decides whether to open a new retail outlet. Because the firm is a monopoly, the return from the investment—the profit from the new store—does not depend on the actions of other firms. As a result, the owner of the monopoly faces no strategic considerations. The owner knows the cost of the investment but is unsure about how many people will patronize

[21]One argument against programs that subsidize insurance is that they provide incentives to engage in excessively risky behavior, such as building or buying homes in areas that have a high probability of being flooded.

the new store; hence, the store's future profit is uncertain. Because the investment has an uncertain payoff, the owner must take risk into account when deciding to invest in the new store.

We first consider the decision of Chris, a risk-neutral owner. Because she is risk neutral, she invests if the expected value of the firm rises due to the investment. Any action that increases her expected value must also increase her expected utility because she is indifferent to risk. In contrast, in the next example, Ken is risk averse, so he might not make an investment that increases his firm's expected value if the investment is very risky. That is, maximizing expected value does not necessarily maximize his expected utility.

Risk-Neutral Investing

Chris, the risk-neutral owner of the monopoly, uses a *decision tree* (panel a of Figure 17.4) to decide whether to invest. The rectangle, called a *decision node*, indicates that she must make a decision about whether to invest or not. The circle, a *chance node*, denotes that a random process determines the outcome (consistent with the given probabilities).

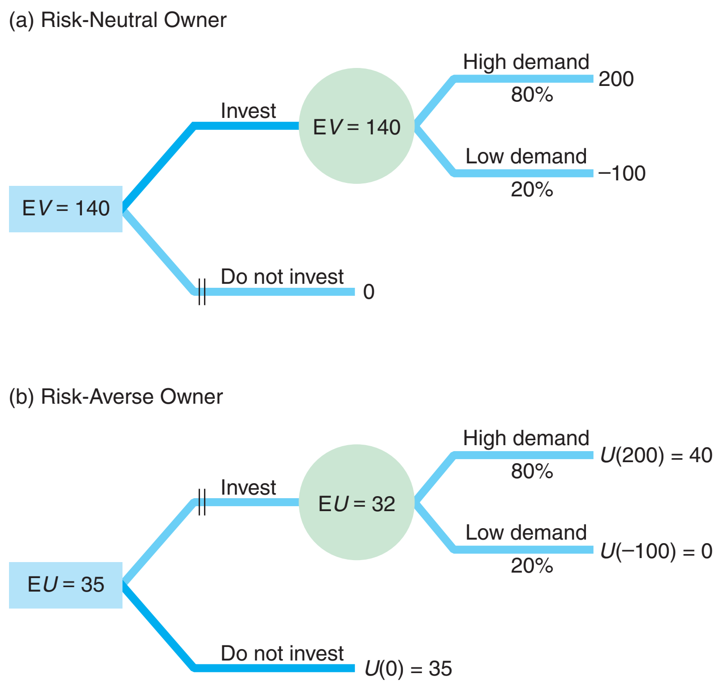
If Chris does not open the new store, she makes \$0. If she does open the new store, she expects to make \$200 (thousand) with 80% probability and to lose \$100 (thousand) with 20% probability. The expected value from a new store (see the circle in panel a) is

$$EV = [0.2 \times (-100)] + [0.8 \times 200] = 140.$$

Because Chris is risk neutral, she prefers an expected value of 140 to a certain one of 0, so she invests. Thus, her expected value in the rectangle is 140.

Figure 17.4 Investment Decision Trees MyLab Economics Video

Chris and Ken, each the owner of a monopoly, must decide whether to invest in a new store. (a) The expected value of the investment is 140, so it pays for Chris, who is risk neutral, to invest. (b) Ken is so risk averse that he does not invest even though the expected value of the investment is positive. His expected utility falls if he makes this risky investment.



Risk-Averse Investing

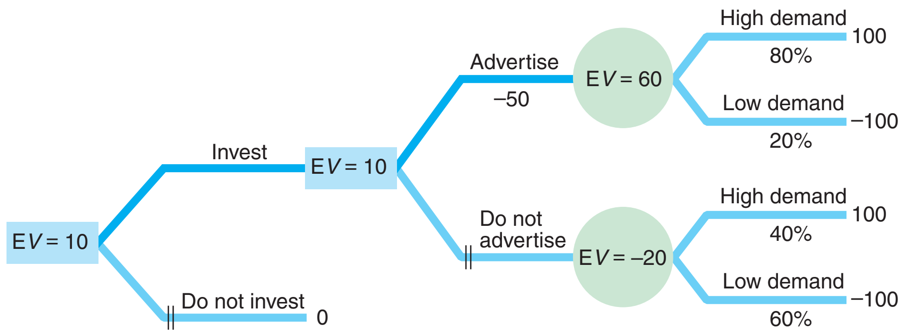
Let's compare Chris' decision-making process to that of Ken's, a risk-averse owner of a monopoly who faces the same investment decision. Ken invests in the new store if his expected utility from investing is greater than his certain utility from not investing. Panel b of Figure 17.4 shows his decision tree, which is based on a particular risk-averse utility function. The circle shows that Ken's expected utility from the investment is

$$\begin{aligned} \text{EU} &= [0.2 \times U(-100)] + [0.8 \times U(200)] \\ &= (0.2 \times 0) + (0.8 \times 40) = 32. \end{aligned}$$

Ken's certain utility from not investing is $U(0) = 35$, which is greater than 32. Thus, Ken does not invest. As a result, his expected utility in the rectangle is 35, his certain utility from not investing.

Solved Problem 17.4

We have been assuming that nature dictates the probabilities of various possible events. However, sometimes an investor can pay to alter the probabilities. Gautam, who is risk neutral, is considering whether to invest in a new store, as the figure shows. After investing, he can increase the probability that demand will be high at the new store by advertising at a cost of 50 ($50,000). If he makes the investment but does not advertise, he has a 40% probability of making 100 and a 60% probability of losing 100. Should he invest in the new store?



Answer

1. *Calculate the expected value of the investment if Gautam does not advertise*. If Gautam makes the investment but does not advertise, the expected value of his investment is

$$[0.4 \times 100] + [0.6 \times (-100)] = -20.$$

Thus, if he does not advertise, he expects to lose money if he makes this investment.

2. *Calculate the expected value of the investment if Gautam advertises and determine whether he should invest and whether he should advertise*. With advertising, Gautam's expected value before paying for the advertisements is

$$[0.8 \times 100] + [0.2 \times (-100)] = 60.$$

Thus, his expected value after paying for the advertisements is 10 (= 60 − 50). As a result, he is better off if he invests *and* advertises than if he does not invest or invests without advertising.

17.5 Behavioral Economics of Uncertainty

Many individuals make choices under uncertainty that are inconsistent with the predictions of expected utility theory. Economists and psychologists explain some of these departures from the predictions of the expected utility model using *behavioral economics*: the use of insights from psychology and research on human cognition and emotional biases to augment the rational economic model in an attempt to better predict economic decision making. (We discussed other applications of behavioral economics in Chapters 4, 11, and 14.)

Biased Assessment of Probabilities

People often have mistaken beliefs about the probability that an event will occur. These biases in estimating probabilities come from several sources, including false beliefs about causality and overconfidence.

Gambler's Fallacy Many—perhaps most—people subscribe to the *gambler's fallacy*:

Common Confusion: Past events affect current, independent outcomes.[22]

For example, suppose that you flip a fair coin and it comes up heads six times in a row. What are the odds that you'll get a tail on the next flip? Because past flips do not affect this one, the chance of a tail remains 50%, yet many people believe that a head is much more likely because they're on a "run." Others hold the opposite but equally false view that the chance of a tail is high because a tail is "due."

Suppose that you have an urn containing three black balls and two red ones. If you draw a ball without looking, your probability of getting a black ball is $\frac{3}{5} = 60\%$. If you replace the ball and draw again, the chance of picking a black ball remains the same. However, if you draw a black ball and do not replace it, the probability of drawing a black ball again falls to $\frac{2}{4} = 50\%$. Thus, the belief that a tail is due after several heads are tossed in a row is analogous to falsely believing that you are drawing without replacement when you are actually drawing with replacement.

Overconfidence Another common explanation for why some people make bets that the rest of us avoid is that these gamblers are overconfident. For example, Golec and Tamarkin (1995) found that football bettors tend to make low-probability bets because they greatly overestimate their probabilities of winning certain types of exotic football bets (an *exotic bet* depends on the outcome of more than one game). In one survey, gamblers estimated their chance of winning a particular bet at 45% when the objective probability was 20%.

Few groups exhibit more overconfidence than male high school athletes. Many U.S. high school basketball and football players believe they will get an athletic scholarship to attend college, but less than 5% receive one. Of this elite group, about 25% expect to become professional athletes, but only about 1.5% succeed.[23]

---

[22]The false belief that one event affects another independent event is captured by the joke about a man who brings a bomb on board a plane whenever he flies because he believes that "The chance of having one bomb on a plane is very small, so the chance of having two bombs on a plane is near zero!"

[23]See **http://www.ncaa.org/wps/wcm/connect/public/NCAA/Resources/Research/Probability+of+Going+Pro** and Rossi and Armstrong (1989).

Application

Biased Estimates

Factor by which Americans are more likely to be killed by a cow than by a shark: 27 — Harper's Index 2015

Do scare stories in newspapers, TV shows, and movies cause people to overestimate relatively rare events and underestimate relatively common ones? Newspapers are more likely to publish "man bites dog" stories than the more common "dog bites man" reports.[24]

If you have seen the movie *Jaws,* you can't help but think about sharks before wading into the ocean. In 2016, newspapers around the world reported shark attacks along the Florida, South Carolina, North Carolina, Texas, California, Australia, Madagascar, and Brazil coasts. Do you worry about shark attacks? You really shouldn't.

Only seven people were killed by sharks in U.S. waters in the decade from 2005 through 2015: an average of about 0.6 a year. That's slightly less likely than suffocating in a bean-bag chair. In 2015, 27 Americans were killed by lightning. The annual number of deaths from potable pool drownings is 23, car–deer collisions 211; motorcycles 2,500; firearm homicides 11,000; all homicides 16,000; car crashes 34,000; prostate cancer 40,000; breast cancer 46,000; cancer 500,000; tobacco-related 500,000; and heart disease 734,000.

Benjamin et al. (2001) reported that, when asked to estimate the frequency of deaths from various causes for the entire population, people overestimate the number of deaths from infrequent causes and underestimate those from more common causes. In contrast, if asked to estimate the number of deaths among their own age group from a variety of causes, their estimates are essentially unbiased. That is not to say that people know the true probabilities—only that their mistakes are not systematic. (However, you should know that, despite the widespread warnings issued every Christmas season, no one has died from eating poinsettia berries.)

Violations of Expected Utility Theory

Economists and psychologists have shown that some people's choices violate the basic assumptions of expected utility theory. One important class of violations arises because people change their choices in response to inessential changes in how choices are described or *framed*, even when the underlying probabilities and events do not change. Another class of violations arises because of a bias toward *certainty*.

Framing Many people (including many economists) believe that people are generally rational:

Common Confusion: People react the same way when given equivalent choices no matter how they are posed.

[24]For example, Indian papers reported on a man bites snake story, noting that Neeranjan Bhaskar has eaten more than 4,000 snakes (*Calcutta Telegraph*, August 1, 2005) and the even stranger "Cobra Dies after Biting Priest of Snake Temple!" (*Express India*, July 11, 2005).

However, experiments show that many people reverse their preferences when a problem is presented, or *framed,* in different but equivalent ways. Tversky and Kahneman (1981) posed a problem to college students: The United States government is considering two alternative programs to combat an unusual disease (e.g., avian flu) that otherwise is expected to kill 600 people. The "exact scientific estimates" of the consequences of these programs are:

- If Program A is adopted, 200 out of 600 people will be saved.
- If Program B is adopted, the probabilities are $\frac{1}{3}$ that 600 people will be saved and $\frac{2}{3}$ that no one will be saved.

When college students were asked to choose, 72% opted for the certain gains of Program A over the possibly larger but riskier gains of Program B.

A second group of students was asked to choose between an alternative pair of programs, and were told:

- If Program C is adopted, 400 out of 600 people will die.
- If Program D is adopted, the probabilities are $\frac{1}{3}$ that no one will die, and $\frac{2}{3}$ that 600 people will die.

When faced with this choice, 78% chose the potentially larger but uncertain losses of Program D over the certain losses of Program C. These results are surprising if people maximize their expected utility: Program A is identical to Program C and Program B is the same as Program D in the sense that these pairs have identical expected outcomes. Expected utility theory predicts consistent choices for the two pairs of programs, but many people make inconsistent choices, preferring Programs A and D. (Even after rereading the options and having the inconsistency problem explained, most of us still feel drawn toward Programs A and D.)

In many similar experiments, researchers have repeatedly observed this pattern, called the *reflection effect*: attitudes toward risk are reversed (reflected) for gains versus losses. People are often risk averse when making choices involving gains, but they are often risk preferring when making choices involving losses.

The Certainty Effect Many people put excessive weight on outcomes that they consider to be certain relative to risky outcomes. This *certainty effect* (or *Allais effect*, after the French economist who first noticed it) can be illustrated using another example from Kahneman and Tversky (1979). First, a group of subjects was asked to choose between two options:

- **Option A.** You receive \$4,000 with probability 80% and \$0 with probability 20%.
- **Option B.** You receive \$3,000 with certainty.

The vast majority, 80%, chose the certain outcome, B.

Then, the subjects were given another set of options:

- **Option C.** You receive \$4,000 with probability 20% and \$0 with probability 80%.
- **Option D.** You receive \$3,000 with probability 25% and \$0 with probability 75%.

Now, 65% prefer C.

Kahneman and Tversky found that over half the respondents violated expected utility theory by choosing B in the first experiment and C in the second one. If $U(0) = 0$, then choosing B over A implies that the expected utility from B is greater than the expected utility from A, so that $U(3{,}000) > 0.8U(4{,}000)$, or $U(3{,}000)/U(4{,}000) > 0.8$. Choosing C over D implies that $0.2U(4{,}000) > 0.25U(3{,}000)$, or $U(3{,}000)/U(4{,}000) < 0.8$ $(= 0.2/0.25)$. Thus, these choices are inconsistent with each other, and hence inconsistent with expected utility theory. The certainty of option B seems to give it extra attractiveness over and above what is implied by expected utility theory.

Expected utility theory is based on gambles with known probabilities, whereas most real-world situations involve unknown or subjective probabilities. Ellsberg (1961) pointed out that expected utility theory cannot account for an ambiguous situation in which many people are reluctant to put substantial decision weight on any outcome. He illustrated the problem in a "paradox." Each of two urns contains 100 balls that are either red or black. You know with certainty that the first urn has 50 red and 50 black balls. You do not know the ratio of red to black balls in the second urn. Most of us would agree that the known probability of drawing a red from the first urn, 50%, equals the subjective probability of drawing a red from the second urn. That is, not knowing how many red and black balls are in the second urn, we have no reason to believe that the probability of drawing a red is greater or less than 50%. Yet, most people would prefer to bet that a red ball will be drawn from the first urn rather than from the second urn.

Prospect Theory

Kahneman and Tversky's (1979) *prospect theory* is an alternative theory of decision making under uncertainty that can explain some of the choices people make that are inconsistent with expected utility theory. According to *prospect theory*, people are concerned about gains and losses—the changes in wealth—rather than the level of wealth, as in expected utility theory. People start with a reference point—a base level of wealth—and think about alternative outcomes as gains or losses relative to that reference level.

Comparing Expected Utility and Prospect Theories We can illustrate the differences in the two theories by comparing how people would act under the two theories when facing the same situation. Both Muzhe and Rui have initial wealth W. They may choose a gamble where they get A dollars with probability θ or B dollars with probability $1 - \theta$. For example, A might be negative, reflecting a loss, and B might be a positive, indicating a gain.

Muzhe wants to maximize his expected utility. If he does not gamble, his utility is $U(W)$. To calculate his expected utility if he gambles, Muzhe uses the probabilities θ and $1 - \theta$ to weight the utilities from the two possible outcomes:

$$\text{EU} = \theta U(W + A) + (1 - \theta)U(W + B),$$

where $U(W + A)$ is the utility he gets from his after-gambling wealth if A occurs and $U(W + B)$ is the utility if he receives B. He chooses to gamble if his expected utility from gambling exceeds his certain utility from his initial wealth: $EU > U(W)$.

In contrast, Rui's decisions are consistent with prospect theory. Rui compares the gamble to her current reference point, which is her initial situation where she has W with certainty. The value she places on her reference point is $V(0)$, where 0 indicates that she has neither a gain nor a loss with this certain outcome. The (negative) value that she places on losing is $V(A)$, and the value from winning is $V(B)$.

To determine the value from taking the gamble, Rui does not calculate the expectation using the probabilities θ and $1 - \theta$, as she would with expected utility theory. Rather, she uses *decision weights* $w(\theta)$ and $w(1 - \theta)$, where the w function assigns different weights from the original probabilities. If people assign disproportionately high weights to rare events (see the Application "Biased Estimates"), the weight $w(\theta)$ exceeds θ for low values of θ and is less for high values of θ.

Rui gambles if the value from not gambling, $V(0)$, is less than her evaluation of the gamble, which is the weighted average of her values in the two cases:

$$V(0) < [w(\theta) \times V(A)] + [w(1 - \theta) \times V(B)].$$

Thus, prospect theory differs from expected utility theory in both the valuation of outcomes and how they are weighted.

Properties of Prospect Theory To resolve various choice mysteries, the prospect theory value function, V, has an S-shape, as in Figure 17.5. This curve has three properties. First, the curve passes through the reference point at the origin, because gains and losses are determined relative to the initial situation.

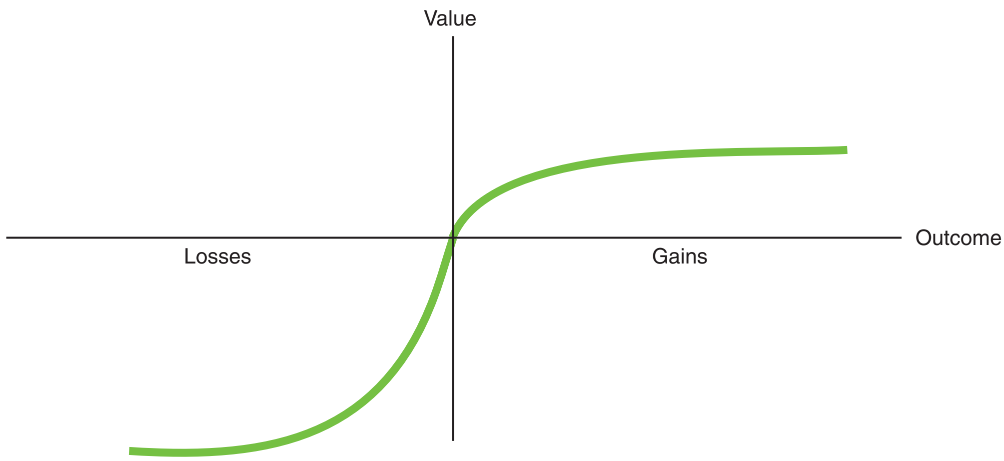
Second, both sections of the curve are concave to the horizontal, outcome axis. Because of this curvature, Rui is less sensitive to a given change in the outcome for large gains or losses than for small ones. For example, she cares more about whether she has a loss of \$1 rather than \$2 than she does about a loss of \$1,001 rather than \$1,002.

Third, the curve is asymmetric with respect to gains and losses. People treat gains and losses differently, in contrast to the predictions of expected utility theory. The S-curve in the figure shows a bigger impact to a loss than to a comparable size gain. That is, the value function reflects *loss aversion*: people hate making losses more than they like making gains.

Given the subjective weights, valuations based on gains and losses, and the shape of the value curve, prospect theory can resolve some of the behavioral mysteries of choice under uncertainty. Because the S-shaped curve shows that people treat gains and losses differently, it can explain the reflection effect in the disease experiment described earlier in this section, where people make different choices when identical outcomes are stated in terms of lives saved instead of lives lost. It also provides an

Figure 17.5 Prospect Theory Value Function MyLab Economics Video

The prospect theory value function has an S-shape. It passes through the reference point at the origin, because gains and losses are measured relative to the initial condition. Because both sections of the curve are concave to the outcome axis, decision makers are less sensitive to a given change in the outcome for large gains or losses than for small ones. Because the curve is asymmetric with respect to gains and losses, people treat gains and losses differently. This S-curve shows a bigger impact to a loss than to a comparable size gain, reflecting loss aversion.



explanation as to why some people engage in unfair lotteries: They put heavier weight on rare events than the true probability used in expected utility theory.

Similarly, we could use a weighting function to resolve the Ellsberg paradox. For example, with the urn containing an unknown ratio of black and red balls, an individual might put 40% on getting a black ball, 40% on getting a red ball, and leave 20% to capture an unwillingness to take a gamble when faced with substantial ambiguity. Doing so reduces the expected value of the gamble relative to that of the initial, certain situation where one does not gamble.

Application

Loss Aversion Contracts

How can a manager induce employees to work hard? If workers have loss aversion, then employees work harder if offered a loss contract—they lose a bonus if they do not work hard enough—than an equivalent gain contract—they receive a bonus if they work hard. Fryer et al. (2012), Hossain and List (2012), and Imas et al. (forthcoming) found that productivity is higher with a loss contract than with a gain contract.

However, the firm may not benefit from using a loss contract if they have to pay workers more to start work at the firm because they prefer a gain contract. Luckily for firms, Imas et al. found that workers prefer loss contracts.

Challenge Solution

BP and Limited Liability

We now address the three questions raised in the Challenge: How does a cap on liability affect a firm's willingness to make a risky investment or to invest less than the optimal amount in safety? How does a cap affect the amount of risk borne by the firm and by the rest of society? How does a cap affect the amount of insurance that a firm buys?

To illustrate the basic ideas, suppose that an oil rig firm expects to earn \$1 billion in the absence of a spill on its new rig and to lose \$39 billion if a spill occurs. The probability of a spill is θ. We start by considering whether the firm invests in a new rig (the analysis would be similar if it were deciding to invest in a given safety feature for a rig).

If the firm is risk neutral, then it invests in the new rig only if the expected return is positive, $[(1 - \theta) \times 1] + [\theta \times (-39)] > 0$, or if $\theta < 1/40 = 2.5\%$.[25] If the firm is risk averse, this *threshold probability*—the highest probability at which the firm is willing to invest—is less than 2.5%.

Now suppose that the firm's liability is capped at \$19 billion. If the firm is risk neutral, it invests in the new rig if $[(1 - \theta) \times 1] + [\theta \times (-19)] > 0$, or if $\theta < 1/20 = 5\%$. Similarly, if the firm is risk averse, the threshold probability is higher than it would be without the limit on liability.

A limit on liability increases society's total risk if it encourages the drilling company to drill when it would not otherwise. If the drilling company is risk neutral, the probability of a spill is $\theta = 3\%$, and the firm bears the full liability for the damages from a spill, then the company's expected earnings are $[0.97 \times 1] + [0.03 \times (-39)] = -0.2 < 0$, so it would not drill. However, if its liability is capped at \$19 billion, then its expected gain from drilling is $[0.97 \times 1] + [0.03 \times (-19)] = 0.4 > 0$, so it would drill. Because the firm is more likely to drill because of the liability cap, the cap causes the rest of society's

[25]The firm compares the expected return to that of the second-best investment opportunity, which we assume is zero for simplicity.

total risk to increase. Moreover, the rest of society bears the risk from the $20 billion ($39 billion − $19 billion) for which it is now responsible if a spill occurs.

If the firm is risk averse, it wants to buy fair insurance to cover its risk. To illustrate the effect of the cap on its decision as to how much insurance the firm buys, we now assume that the probability of a disaster is $\theta = 1\%$. Without either a liability cap or insurance, the firm's expected gain is $[0.99 \times 1] + [0.01 \times (-39)] = \$0.6$ billion. If an insurance company would provide fair insurance, the drilling firm could buy $100 of insurance for each $1 spent. Given that the drilling company is risk averse, it fully insures, so that if a spill occurs, the insurance company pays $39 billion. To buy this much insurance, the drilling company pays $0.39 billion, so that the expected value of the insurance contract is $-0.39 + [0.01 \times 39] = \$0$. With the insurance, the company earns $[0.99 \times 1] - 0.39 = \$0.6$ billion whether or not a spill occurs.

If the drilling company's liability is capped at $19 billion, it buys $19 billion worth of insurance for $0.19 billion, so that its expected gain in either state of nature is $1 - 0.19 = \$0.81$ billion. That is, the drilling company's expected profit increases by $0.21 billion due to the limit on its liability. This amount is a transfer from the rest of society to the firm, because society will be responsible for the extra $20 billion in damages if the spill occurs.

Summary

1. **Assessing Risk.** A probability measures the likelihood that a particular state of nature occurs. People may use historical frequencies, if available, to calculate probabilities. Lacking detailed information, people may form subjective estimates of a probability on the basis of available information. The expected value is the probability-weighted average of the values in each state of nature. One widely used measure of risk is the variance (or the standard deviation, which is the square root of the variance). The variance is the probability-weighted average of the squared difference between the value in each state of nature and the expected value.

2. **Attitudes Toward Risk.** Whether people choose a risky option over a nonrisky one depends on their attitudes toward risk and the expected payoffs of the various options. Most people are *risk averse* and will choose a riskier option only if its expected value is sufficiently higher than that of a less risky option. *Risk-neutral* people choose whichever option has the higher rate of return because they do not care about risk. *Risk-preferring* people may choose the riskier option even if it has a lower rate of return because they like risk and are willing to give up some expected return to take on more risk. An individual's utility function reflects that person's attitude toward risk. Expected utility is the probability-weighted average of the utility from the outcomes in the various states of nature. According to expected utility theory, decision makers choose the option that provides the highest expected utility.

3. **Reducing Risk.** People try to reduce the risk they face in several ways. They avoid some risks altogether and, when risks cannot be completely avoided, take actions that lower the probabilities of bad events or reduce the harm from bad events when they do occur. By collecting information before acting, investors can make better choices. People can reduce risk by diversifying over a range of investments unless the various investments are perfectly positively correlated. Insurance companies offer policies for risks that they can diversify by pooling risks across many individuals. Insurance is called fair if the expected return to the policyholder is zero: the expected payout equals the premium paid. Risk-averse people fully insure if they are offered fair insurance. Because insurance companies must earn enough income to cover their full operating costs, they offer insurance that is less than fair. Although risk-averse people may buy unfair insurance, they buy less than with full insurance.

4. **Investing Under Uncertainty.** Whether a person makes an investment depends on the uncertainty of the payoff, the expected return, the individual's

attitudes toward risk, the interest rate, and the cost of altering the probabilities of various outcomes. For a risk-neutral person, an investment pays if the expected value is positive. A risk-averse person invests only if that person's expected utility is higher after investing. Thus, risk-averse people make risky investments if the investments pay sufficiently higher rates of return than do safer investments. People pay to alter the probabilities of various outcomes from an investment if doing so raises their expected utility.

5. **Behavioral Economics of Uncertainty.** Economists and psychologists have identified behavior under uncertainty that is inconsistent with expected utility theory. These choices may be due to biased estimates of probabilities or different objectives than expected utility. For example, some people care more about losses than about gains. One alternative theory that is consistent with many of these puzzling choices is prospect theory, which allows people to treat gains and losses asymmetrically and to weight outcomes differently than with the probabilities used in expected utility theory.

Questions

Select questions are available on MyLab Economics*; * = answer appears at the back of this book;* **A** *= algebra problem;* **C** *= calculus problem.*

1. Assessing Risk

1.1 In a neighborhood with 1,000 houses, 5 catch fire, 7 are damaged by high winds, and the rest are unharmed during a one-year period. What do you estimate is the probability that a house is harmed by fire or high winds? **A**

*1.2 Asa buys a painting. The probability that the artist will become famous and the painting will be worth \$1,000 is 20%. The probability that the painting will be destroyed by fire or some other disaster is 10%. If the painting is not destroyed and the artist does not become famous, it will be worth \$500. What is the expected value of the painting? **A**

*1.3 On the advice of your financial advisor, you buy 50 shares of a company at the price of €17 per share. She has told you that there is an 80% probability that the stock price will increase to €25 over the course of the next six months and a 20% chance that the stock price will fall to €15 per share. What is the expected value and the variance of your purchase? **A**

1.4 Tiffany plans to sell pies that she will buy this evening at a street fair tomorrow. If the weather is nice, she will earn \$200; however, if it rains, she will lose \$60, the amount she would pay for the pies that she won't be able to sell. The weather forecast says that the chance of rain is 25%. What does she expect to earn? How much more would she expect to earn if she had perfect information about the probability of rain far enough before the street fair that she would only buy the pies she could sell? (*Hint*: See Solved Problem 17.1.) **A**

1.5 An Australian company accepts a contract to drill a water well for household use. Its profit will be AU\$3,500 if the well depth necessary to provide an adequate supply of water is 35 meters, but it will lose AU\$1,600 if the well depth is 80 meters. Given the firm's experience and knowledge of the area, the probability of having to drill to 80 meters is 30%. What is the most the firm would be willing to pay for a seismic-reflection survey that would determine exactly how deep the company would need to drill the water well? (*Hint*: See Solved Problem 17.1.)

*1.6 A city has a problem with people parking their cars illegally. To reduce the number of offenses and improve road and pedestrian safety, it can either hire more parking enforcement officers or increase the fine it charges owners of illegally parked vehicles. Explain why either method can be used to deal with the problem. Which approach is the city government likely to prefer, and why?

1.7 Suppose that the fine for illegally parking a car is ₹3,200 and that a fine amount of ₹1,760 is sufficient for discouraging most people from parking their cars illegally. How high must the probability of being caught and fined be to discourage most illegal parking? **A**

2. Attitudes Toward Risk

2.1 Ryan offers to bet Kristin that if a six-sided die comes up with one or two dots showing, he will pay her \$3, but if it comes up with any other number of dots, she'll owe him \$2. Is that a fair bet for Kristin? **A**

2.2 Suppose that Maoyong's utility function with respect to wealth is $U(W) = \ln W$ (where "$\ln W$" means the natural logarithm of W). Plot this utility function and illustrate in your figure why Maoyong is risk averse.

2.3 Jen's utility function with respect to wealth is $U(W) = \sqrt{W}$. Plot this utility function and illustrate in your figure why Jen is risk averse.

2.4 Suppose that Laura's utility function is $U(W) = W^{0.5}$, where W is wealth. Is she risk averse? Show mathematically. **C**

*2.5 Given the information in Solved Problem 17.2, Irma prefers to buy the stock. Show graphically how high her certain income would have to be for her to choose not to buy the stock.

2.6 In Question 1.3, you paid €850 to purchase stock in a company on the advice of your financial advisor. Now suppose you are risk-averse and have the concave utility function $U(W) = \sqrt{W}$, where W is wealth, and it is equally likely that the stock will not do as well. Its value will be $850 + x$ if it does well, and $850 - x$ if it does not do well. What is your risk premium if $x = 400$ and if $x = 160$? In comparing your two answers, what can you deduce about the relationship between the risk premium and the value of x (the variability of the prospect)?

2.7 Suppose that Laura has a utility function of $U(W) = W^{0.5}$ and an initial wealth of $W = \$100$. How much of a risk premium would she want to participate in a gamble that has a 50% probability of raising her wealth to \$120 and a 50% probability of lowering her wealth to \$80? (*Hint*: See Solved Problem 17.2.) **A**

2.8 What is the risk premium if, in Question 2.7, Laura's utility function were ln W? (*Hint*: See Solved Problem 17.2.) **A**

*2.9 Hugo has a concave utility function of $U(W) = W^{0.5}$. His only asset is shares in an Internet start-up company. Tomorrow he will learn the stock's value. He believes that it is worth \$144 with probability $\frac{2}{3}$ and \$225 with probability $\frac{1}{3}$. What is his expected utility? What risk premium would he pay to avoid bearing this risk? (*Hint*: See Solved Problem 17.2.) **A**

2.10 Joanna is considering three possible jobs. The following table shows the possible incomes she might get in each job.

	Outcome A		Outcome B	
	Probability	Earnings	Probability	Earnings
Job 1	0.5	20	0.5	40
Job 2	0.3	15	0.7	45
Job 3	1	30		

For each job, calculate the expected value, the variance, and the standard deviation. If Joanna is averse to risk (as measured by variance), what can you predict about her job choice? What if she is risk neutral?

2.11 Lisa just inherited a vineyard from a distant relative. In good years (without rain or frost during the harvest season), she earns \$100,000 from the sale of grapes from the vineyard. If the weather is poor, she loses \$20,000. Lisa's estimate of the probability of good weather is 60%.

a. Calculate the expected value and the variance of Lisa's income from the vineyard.

b. Lisa is risk averse. Ethan, a grape buyer, offers Lisa a guaranteed payment of \$70,000 each year in exchange for her entire harvest. Will Lisa accept this offer? Explain.

c. Why might Ethan make such an offer? Give three reasons, and explain each. One of these reasons should refer to his attitude toward risk. Illustrate this reason using a diagram that shows the general shape of Ethan's utility function over income. **A**

2.12 Farrell et al. (2000) estimated that the elasticity of demand for lottery tickets is about −1. If the U.K. National Lottery is running its game to make money (it gets a percentage of the total revenues), is it running the lottery optimally? Explain your answer. **A**

2.13 The Casino de Monte-Carlo is a gambling and entertainment complex located in Monaco that attracts a large international clientele to its table games, slot machines, and special gaming events. It is likely that most of its visitors are risk averse. Provide at least three reasons why so many risk-averse people would travel there to gamble. (*Hint*: See the Application "Gambling.")

3. Reducing Risk

3.1 Lori, who is risk averse, has two pieces of jewelry, each worth \$1,000. She wants to send them to her sister in Thailand, but she is concerned about the safety of shipping them. She believes that the probability that the jewelry won't arrive is θ. Is her expected utility higher if she sends the articles together or in two separate shipments? Explain.

3.2 Lucy, the manager of the medical test firm, Dubrow Labs, worries that the firm may be sued for botching results from blood tests. The firm expects to earn a profit of 100 if it is not sued, but only 10 if it is successfully sued. Lucy believes the probability of a successful suit is 5%. If fair insurance is available and Lucy is risk averse, how much insurance will she buy? **A**

*3.3 Consider a household that possesses \$160,000 worth of valuables such as jewelry. This household faces a 0.2 probability of a burglary in which it loses \$70,000 worth of the valuables. Suppose it can buy an insurance policy for \$15,000 that would fully reimburse the amount of loss from burglary. The household's utility is given by $U(X) = 4X^{0.5}$.

a. Should the household buy this insurance policy?

b. What is the fair price for the insurance policy?

c. What is the most the household is willing to pay for this insurance policy that fully covers it against the loss? **A**

3.4 Vivaan wants to protect her valuables against damage and theft. If there is a 25% probability of burglary and she wishes to purchase insurance that would fully reimburse her for a loss of up to ₹2.5 million, what would be the cost of an insurance policy to her if the insurance company offered fair insurance? If the company charges Vivaan, who is risk neutral, a premium of ₹750,000 for the insurance policy, would she accept it? (*Hint*: See Solved Problem 17.3.)

3.5 Heavy rains caused serious flooding in Lake Ontario and the St. Lawrence River in spring 2017 and severe damage to properties on both sides of the border between Canada and the United States. Does the expectation that the government will subsidize the cost of rebuilding a home that is destroyed by flooding affect the probability that risk-averse people will buy insurance or the amount that they actually buy? To consider these questions, let the utility function for a risk-averse person be $U(W) = 2W^{0.5}$ and the probability that the person's house will be completely destroyed be 0.3%.

a. If there is no subsidy, the loss in wealth is C\$500,000. Will a prospective client buy insurance if the cost of insurance is C\$2,000?

b. If the government provides a subsidy equal to 25% of the cost of rebuilding the house so that the loss in wealth is reduced to C\$375,000, will the client buy insurance if the cost of insurance is still C\$2,000?

c. If the government provides a subsidy equal to 25% of the cost of rebuilding the house so that the loss in wealth is reduced to C\$375,000, will the client buy insurance if the cost of insurance is reduced to C\$1,000?

3.6 Using information from the Application "Flight Insurance," calculate the price of fair insurance if the probability were as high as 0.00000077, the frequency in 2001 when many people died in the 9/11 disasters. **A**

4. Investing Under Uncertainty

4.1 What is the difference—if any—between an individual's gambling at a casino and buying a stock? What is the difference for society?

*4.2 Andy and Kim live together. Andy may invest \$10,000 (possibly by taking on an extra job to earn the additional money) in Kim's education this year. This investment will raise Kim's future earnings by \$24,000 (in present value terms—see Chapter 16). If they stay together, they will share the benefit from the additional earnings. However, the probability is $\frac{1}{2}$ that they will split up in the future. If they were married and then split, Andy would get half of Kim's additional earnings. If they were living together without any legal ties and they split, then Andy would get nothing. Suppose that Andy is risk neutral. Will Andy invest in Kim's education? Does your answer depend on the couple's legal status? **A**

4.3 Use a decision tree to illustrate how a risk-neutral plaintiff in a lawsuit decides whether to settle a claim or go to trial. The defendants offer €60,000 to settle. If the plaintiff does not settle, she believes that the probability of winning at trial is 70%. If the plaintiff wins, the amount awarded is V. How large can V be before the plaintiff refuses to settle? If the plaintiff was risk averse, would the result change? **A**

4.4 A medical doctor has a patient with coronary artery disease. The decreased flow of blood to the heart is causing chest pain and shortness of breath, which has lowered the patient's utility. The patient is presently taking drugs to increase the blood flow to the heart. An alternative is bypass surgery, which can correct the problem and return the patient's level of utility to a more normal level. However, there is a 4% probability that the patient will die as a result of the surgery. Use a decision tree to illustrate how the patient would make a decision about whether to have the operation. Assume the patient's normal level of utility is 100, present level of utility is 50, level of utility following a successful surgery is 80, and level of utility following an unsuccessful surgery is 0.

4.5 In Solved Problem 17.4, advertising increases the probability of high demand to 80%. If all the other information in the Solved Problem stays the same, what is the minimum probability of high demand resulting from advertising such that Gautam decides to invest and advertise?

4.6 Jian and Chen are each considering purchasing a certain stock at a price of ¥135 per share. They have the same financial analyst, and according to her, there is a 60% probability that the share price will rise to ¥190 over the next year and a 25% probability that it will fall to ¥60. Jian is risk loving. Chen, on the other hand, is risk averse and has the utility function $U(W) = W^{0.6}$, where W is wealth. Using decision trees, show that Jian is more likely to buy the stock than Chen is, all other factors being the same. (*Hint*: See Figure 17.4.)

4.7 Elections can be emotional events. After all, they may put into power a person or party whose policies may be different from those to which we've become accustomed. Political decisions affect economic policy and economic outcomes which, in turn, impact financial markets. Thus, political uncertainty is often reflected in falling stock prices and depressed investment as firms adopt a more cautious approach until after the election. Explain a possible economic reason for this result. Does your explanation require that the firms' managers are risk averse?

5. Behavioral Economics of Uncertainty

5.1 First answer the following two questions about your preferences:

a. You are given \$5,000 and offered a choice between receiving an extra \$2,500 with certainty or flipping a coin and getting \$5,000 more if heads or \$0 if tails. Which option do you prefer?

b. You are given \$10,000 if you will make the following choice: return \$2,500 or flip a coin and return \$5,000 if heads and \$0 if tails. Which option do you prefer?

Most people choose the sure \$2,500 in the first case but flip the coin in the second. Explain why this behavior is not consistent. What do you conclude about how people make decisions concerning uncertain events? **A**

5.2 Evan is risk seeking with respect to gains and risk averse with respect to losses. Louisa is risk seeking with respect to losses and risk averse with respect to gains. Illustrate both utility functions. Which person's attitudes toward risk are consistent with prospect theory? Which of these people would you expect to be susceptible to framing effects?

5.3 Suppose a person's utility function is $U(W) = 3W^3 - 40W^2 + 180$, where W is wealth. Their initial wealth is 6. If invested, that amount could rise or fall.

a. If it is equally likely that wealth could remain unchanged or fall to 2, compare the expected utility of the loss with the utility of having certain wealth equal to the expected wealth. Is the person risk averse?

b. If it is equally likely that wealth could remain unchanged or rise to 10, compare the expected utility of the gain with the utility of having certain wealth equal to the expected wealth. Is the person risk preferring?

c. What can you say about this person's risk preference? Draw their utility curve.

*5.4 Joe has lost a substantial amount gambling at a racetrack today. On the last race of the day, he decides to make a large enough bet on a longshot so that, if he wins, he will make up for his earlier losses and break even on the day. His friend Sue, who won more than she lost on the day, makes just a small final bet so that she will end up ahead for the day even if she loses the last race. This is typical race track behavior for winners and losers. Would you explain this behavior using overconfidence bias, prospect theory, or some other principle of behavioral economics?

5.5 What are the major differences between expected utility theory and prospect theory?

6. Challenge

6.1 Unfortunately, oil spills are not uncommon. For example, in August 2016, two tankers ran into each other off the coast of Japan, causing an oil spill. In January 2017, two different tankers ran into each other off the coast of India, causing another. The environmental consequences of these accidents, which can be very costly, include clean-up and site restoration costs, legal fees, punitive damages, and compensation for people whose livelihoods have been adversely affected. Suppose a firm (the contractee) contracts with an oil tanker company (the contractor) to transport refined petroleum products for €8 million. The contractor is fully liable for damages from any oil spills. The total cost of an oil spill is €20 million, and the probability of a spill is δ.

a. What is the value of δ that would make a risk-neutral contractor indifferent to accepting or rejecting the contract? For what values of δ would the contractor accept the contract?

b. If the true probability of a spill is 2%, the contractor is risk averse, and fair insurance is offered, how much insurance would the contractor buy?

c. If the contractee partially compensates the risk-neutral contractor so that the latter would have to pay only 40% of the cost of an oil spill, what value of δ would make the contractor indifferent to accepting or rejecting the contract? What does this say about the relationship between liability and acceptance?

d. If the true probability of a spill is 2%, the contractor is risk averse, fair insurance is offered, and the contractor is partially compensated as in part c, how much fair insurance will the contractor now buy?

Externalities, Open-Access, and Public Goods

18

I shot an arrow in the air and it stuck.

Challenge

Trade and Pollution

According to a 2016 estimate, 5.5 million people die annually from air pollution. Does free trade cause much of this pollution? That's what protesters in many countries allege. For years, these protesters have disrupted meetings of the World Trade Organization (WTO), which promotes free trade among its 161 member countries. The WTO forbids member countries from passing laws that unreasonably block trade, including environmental policies. The environmental protesters argue that when rich countries with relatively strong pollution laws import from poor countries without controls, world pollution rises. Even a country that only cares about its own welfare wants to know the answer to the question: Does exporting benefit a country if it does not regulate its domestic pollution?

property right
the exclusive privilege to use an asset

In this chapter, we show that if a **property right**—an exclusive privilege to use an asset—is not clearly assigned, a market failure is likely. By owning this book, you have a property right to read it and to stop others from taking it. But many goods have incomplete or unclear property rights.

Unclearly defined property rights may cause *externalities*, which occur when someone's consumption or production activities help or harm someone else outside of a market. A harmful externality occurs when a manufacturing plant spews pollution, injuring neighboring firms and individuals. When people lack a property right to clean air, factories, drivers, and others pollute the air rather than incur the cost of reducing their pollution.

Indeed, if no one holds a property right for a good or a bad (like pollution), it is unlikely to have a price. If you had a property right to be free from noise pollution, you could use the courts to stop your neighbor from playing loud music. Or you could sell your right, permitting your neighbor to play the music. If you did not have this property right, no one would be willing to pay you a positive price for it.

We start our analysis by examining pollution. Some of the most important bad externalities arise as a by-product of production (such as water pollution from manufacturing) and consumption (such as congestion or air pollution from driving). A competitive market produces more pollution than a market that is optimally regulated by the government, but a monopoly may not create as much of a pollution problem as a competitive market. Clearly defined property rights help reduce externality problems.

Next, we show that market failures due to externalities also occur if a good lacks exclusion. A good has *exclusion* if its owner has clearly defined property rights and can prevent others from consuming it. You have a legal right to stop anyone from eating your apple. However, a country's national defense cannot protect some citizens without protecting all citizens.

Market failures may also occur if a good lacks *rivalry*, where only one person can consume it, such as an apple. National defense lacks rivalry because my consumption does not prevent you from consuming it.

We look at three types of markets that lack exclusion or rivalry or both. An *open-access common property* is a resource, such as an ocean fishery, where *exclusion* of potential users is impossible. A *club good*, such as a swimming pool, is a good or service that allows for exclusion but is *nonrival*: One person's consumption does not use up the good—others can also consume it (at least until capacity is reached). A *public good*, such as national defense and clean air, is both nonexclusive and nonrival. Public goods may not have a market or the market undersupplies these goods.

When such market failures arise, government intervention may raise welfare. A government may regulate an externality such as pollution directly, or indirectly control an externality through taxation or laws that make polluters liable for the damage they cause. Similarly, a government may provide a public good.

In this chapter, we examine six main topics

1. **Externalities.** By-products of consumption and production may benefit or harm other people.
2. **The Inefficiency of Competition with Externalities.** A competitive market produces too much of a harmful externality.
3. **Regulating Externalities.** Taxation or regulation can reduce or prevent the overproduction of pollution and other externalities.
4. **Market Structure and Externalities.** With a harmful externality, a noncompetitive market equilibrium may be closer to the socially optimal level than that of a competitive equilibrium.
5. **Allocating Property Rights to Reduce Externalities.** Clearly assigning property rights allows exchanges that reduce or eliminate externality problems.
6. **Rivalry and Exclusion.** If goods lack rivalry or exclusion, competitive markets suffer from a market failure.

18.1 Externalities

externality
the direct effect of the actions of a person or firm on another person's well-being or a firm's production capability rather than an indirect effect through changes in prices

An **externality** occurs when a person's well-being or a firm's production capability is directly affected by the actions of other consumers or firms rather than indirectly through changes in prices. A firm whose production process lets off fumes that harm its neighbors is creating an externality, which is not traded in a market. In contrast, the firm is not causing an externality when it harms a rival by selling extra output that lowers the market price.

Externalities may either help or harm others. An externality that harms someone is called a *negative externality*. Your neighbors harm you if they keep you awake by screaming at each other late at night. A chemical plant spoils a lake's beauty when it dumps its waste product into the water, harming a firm that rents boats for use on that waterway. Government officials in Sydney, Australia, used loud Barry Manilow music to drive away late-night revelers from a suburban park—and in the process drove local residents out of their minds.[1]

A *positive externality* benefits others. A 2016 report by the U.S. Forest Service estimated that trees lining Californian streets and boulevards provide $1 billion in benefits to municipalities and residents, including carbon storage, the removal of air pollutants, the interception of rainfall, energy savings in heating and cooling homes, as well as aesthetics.

A single action may confer positive externalities on some people and negative externalities on others. The smell of pipe smoke pleases some people and annoys others. Some people think that their wind chimes please their neighbors, whereas anyone with an ounce of sense knows that those chimes are annoying! It was reported that efforts to clean the air in Los Angeles, while helping people breathe more easily, caused radiation levels to increase far more rapidly than if the air had remained dirty.

Application

Negative Externalities from Spam

Spam—unsolicited bulk email messages—inflicts a major negative externality on businesses and individuals around the world by forcing people to waste time removing it, by inducing people to reveal private information unintentionally, and infecting computers with malware, which is malicious software. Spammers take advantage of the open-access nature of email. A spammer targets people who might be interested in the information provided in the spam message. This target group is relatively small compared to the vast majority of recipients who do not want the message and who incur the costs of reading and removing it. (Moreover, many spam messages are scams.) In 2016, people around the world receive 400 billion spam messages daily, which is 86% of global email traffic according to Talos.

The worldwide cost of spam and malware is enormous. Global spending on cyber-security technology alone exceeded $84 billion in 2015. Firms incur large costs to delete spam by installing spam filters and using employees' labor. A study at a German university found that the working time losses caused by spam were approximately 1,200 minutes or $2\frac{1}{2}$ days per employee per year (Caliendo et al., 2012). Various estimates of the cost range from $20 billion to $50 billion per year. Yahoo! researchers, Rao and Reiley (2012), concluded that society loses $100 for every $1 of profit to a spammer, a rate that is "at least 100 times higher than that of automobile pollution."

18.2 The Inefficiency of Competition with Externalities

Competitive firms and consumers do not have to pay for the harms of their negative externalities, so they create excessive amounts. Similarly, because producers are not compensated for the benefits of a positive externality, too little of such externalities is produced.

To illustrate why externalities lead to nonoptimal production, we examine a (hypothetical) competitive market in which firms produce paper and by-products of the production process—such as air and water pollution—that harm people who live near paper mills. We'll call the pollution *gunk*. Producing an extra ton of paper

[1]"Manilow Tunes Annoy Residents," **cnn.com**, July 17, 2006.

increases the amount of gunk. The only way to decrease the volume of gunk is to reduce the amount of paper manufactured. No less-polluting technologies are available, and it is not possible to locate plants where the gunk bothers no one.

private cost
the cost of production only, not including *externalities*

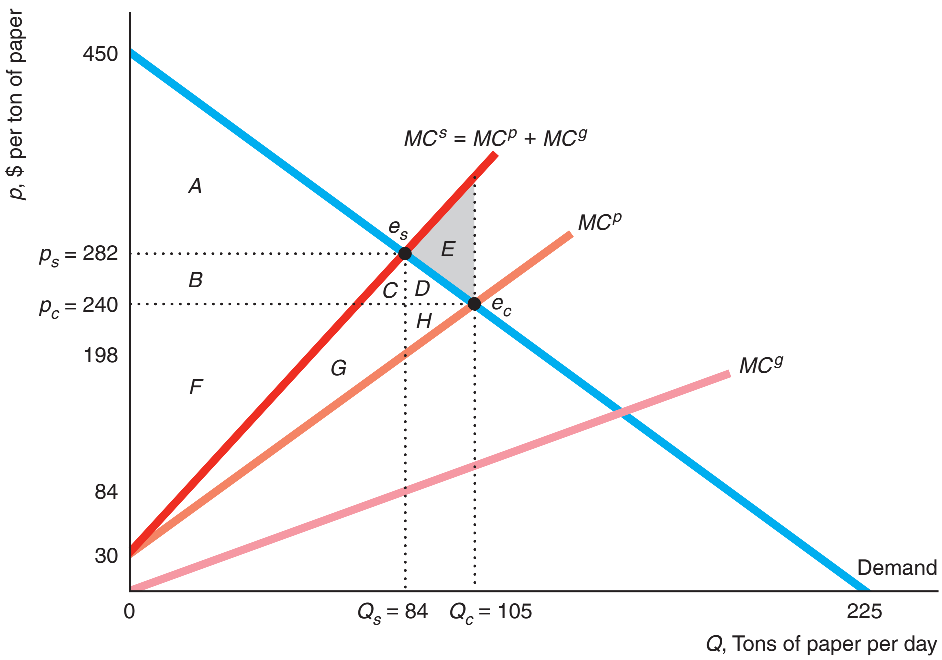
social cost
the private cost plus the cost of the harms from *externalities*

Paper firms do not have to pay for the harm from the pollution they cause. As a result, each firm's **private cost**—the cost of production only, not including externalities—includes its direct costs of labor, energy, and wood pulp but not the indirect costs of the harm from gunk. The true **social cost** is the private cost plus the cost of the harms from externalities.

The paper industry is the major industrial source of water pollution. We use a supply-and-demand diagram for the paper market in Figure 18.1 to illustrate that *a competitive market produces excessive pollution because the firms' private cost is less than their social cost.*[2] In the competitive equilibrium, the firms consider only

Figure 18.1 Welfare Effects of Pollution in a Competitive Market MyLab Economics Video

The competitive equilibrium, e_c, is determined by the intersection of the demand curve and the competitive supply or private marginal cost curve, MC^p, which ignores the cost of pollution. The social optimum, e_s, is at the intersection of the demand curve and the social marginal cost curve, $MC^s = MC^p + MC^g$, where MC^g is the marginal cost of the pollution (gunk). Private producer surplus is based on the MC^p curve, and social producer surplus is based on the MC^s curve.



	Social Optimum	Private	Change
Consumer Surplus, CS	A	$A + B + C + D$	$B + C + D$
Private Producer Surplus, PS_p	$B + C + F + G$	$F + G + H$	$H - B - C$
Externality Cost, C_g	$C + G$	$C + D + E + G + H$	$D + E + H$
Social Producer Surplus, $PS_s = PS_p - C_g$	$B + F$	$F - C - D - E$	$-B - C - D - E$
Welfare, $W = CS + PS_s$	$A + B + F$	$A + B + F - E$	$-E = DWL$

[2]Appendix 18A uses algebra to analyze this model and derives the numbers in the figure. These numbers are not based on actual estimates.

their private costs in making decisions and ignore the harms of the pollution externality they inflict on others. The market supply curve is the aggregate *private marginal cost* curve, MC^p, which is the horizontal sum of the private marginal cost curves of each of the paper manufacturing plants.

The intersection of the market supply curve and the market demand curve for paper determines the competitive equilibrium, e_c. The competitive equilibrium quantity is $Q_c = 105$ tons per day, and the competitive equilibrium price is $p_c = \$240$ per ton.

The firms' *private producer surplus* is the producer surplus of the paper mills based on their *private marginal cost* curve: the area, $F + G + H$, below the market price and above MC^p up to the competitive equilibrium quantity, 105. The competitive equilibrium maximizes the sum of consumer surplus and private producer surplus (Chapter 9). If the market had no externality, the sum of consumer surplus and private producer surplus would equal welfare, so competition would maximize welfare.

Because of the pollution, however, the competitive equilibrium does *not* maximize welfare. Competitive firms produce too much gunk because they do not have to pay for the harm from the gunk. This *market failure* (Chapter 9) results from competitive forces that equalize the price and *private marginal cost* rather than *social marginal cost*, which includes both the private costs of production and the externality damage.

For a given amount of paper production, the full cost of one more ton of paper to society, the *social marginal cost* (MC^s), is the cost of manufacturing one more ton of paper to the paper firms plus the additional externality damage to people in the community from producing this last ton of paper. Thus, the height of the social marginal cost curve, MC^s, at any given quantity equals the vertical sum of the height of the MC^p curve (the private marginal cost of producing another ton of paper) plus the height of the MC^g curve (the marginal externality damage) at that quantity.

The social marginal cost curve intersects the demand curve at the socially optimal quantity, $Q_s = 84$. At smaller quantities, the price—the value consumers place on the last unit of the good sold—is higher than the full social marginal cost. The gain to consumers of paper exceeds the cost of producing an extra unit of output (and hence an extra unit of gunk). At larger quantities, the price is below the social marginal cost, so the gain to consumers is less than the cost of producing an extra unit.

Welfare is the sum of consumer surplus and *social producer surplus*, which is based on the *social marginal cost* curve rather than the *private marginal cost* curve. *Welfare is maximized where price equals social marginal cost*. At the social optimum, e_s, welfare equals $A + B + F$: the area between the demand curve and the MC^s curve up to the optimal quantity, 84 tons of paper.

Welfare at the competitive equilibrium, e_c, is lower: $A + B + F - E$, the areas between the demand curve and the MC^s curve up to 105 tons of paper. The area between these curves from 84 to 105, $-E$, is a deadweight loss because the social cost exceeds the value that consumers place on these last 21 tons of paper. *A deadweight loss results because the competitive market equates price with private marginal cost instead of with social marginal cost*.

Welfare is higher at the social optimum than at the competitive equilibrium because the gain from reducing pollution from the competitive to the socially optimal level more than offsets the loss to consumers and producers of the paper. The cost of the pollution to people who live near the factories is the area under the MC^g curve between zero and the quantity produced. By construction, this area is the same as the area between the MC^p and the MC^s curves. The total damage from the gunk is $-C - D - E - G - H$ at the competitive equilibrium and only $-C - G$ at the social optimum. Consequently, the extra pollution damage from producing

Good news. Production's up 13%!

the competitive output rather than the socially optimal quantity is $-D - E - H$.

The main beneficiaries from producing at the competitive output level rather than at the socially optimal level are the paper buyers, who pay \$240 rather than \$282 for a ton of paper. Their consumer surplus rises from A to $A + B + C + D$. The corresponding change in private producer surplus is $H - B - C$, which is negative in this figure.

The figure illustrates two main results with respect to negative externalities. First, *a competitive market produces excessive negative externalities*. Because the price of the pollution to the firms is zero, which is less than the marginal cost that the last unit of pollution imposes on society, an unregulated competitive market produces more pollution than is socially optimal.

Second, *the optimal amount of pollution is greater than zero*. Even though pollution is harmful and we'd like to have none of it, we cannot wipe it out without eliminating virtually all production and consumption. Making paper, dishwashers, and televisions creates air and water pollution. Fertilizers used in farming pollute the water supply. Delivery people pollute the air by driving to your home.

Application

Global Warming

A 2014 United Nations (UN) report concluded that global warming is occurring and the situation is becoming more serious. This report reflected the work of 800 scientists on a UN panel, which had previously won the Nobel Peace Prize for its work on the environment.

According to the UN report, human activity—pollution—is causing temperatures to rise. At least 97% of actively publishing climate scientists agree that the climate-warming trends over the past century are likely due to human activities, as do the Academies of Science from 80 different countries. Nonetheless, some non-scientists are skeptical. A 2015 Pew Research Center poll of U.S. adults found that 68% say that solid evidence exists that the planet has been warming over the last several decades, while 25% think that such evidence does not exist.

The UN report found that continued emission of greenhouse gases will cause further warming, increasing the likelihood of "severe, pervasive and irreversible impacts for people and ecosystems." Island nations and coastal cities face inundation due to rising sea levels. Pal and Eltahir (2015) predicted that, by the end of this century, areas of the Persian Gulf could be hit by severe waves of heat and humidity that would be "intolerable to humans." The global volume of weather-related insurance losses has more than tripled since the 1980s. A 2016 World Bank report predicted that more than 100 million people could be driven into extreme poverty by 2030, unless actions are taken to protect the world's poor from climate change catastrophes such as crop failures, natural disasters, and waterborne diseases.

18.3 Regulating Externalities

Because competitive markets produce too many negative externalities, government intervention may provide a social gain. In 1952, London suffered from unusually thick "peasouper" fog—pollution so dense that people had trouble finding their way home—that killed an estimated 4,000 to 12,000 people. Those dark days prompted the British government to pass its first Clean Air Act, in 1956.[3] The United States and Canada passed Clean Air Acts in 1970.

Now virtually the entire world is concerned about pollution. Carbon dioxide (CO_2), which is primarily produced by burning fossil fuels, is a major contributor to global warming, damages marine life, and causes additional harm. China and the United States are by far the largest producers of CO_2 from industrial production, as Table 18.1 shows. The amount of CO_2 per person is extremely high in Australia, Canada, Russia, and the United States. China and Russia have a very high pollution to gross domestic product (GDP) ratio. The last column of the table shows that China and India at least doubled their production of CO_2 since 1990, while only a few countries—such as France, Germany, Russia, and the United Kingdom—reduced their CO_2 production.

China produces 27% of the world's CO_2, the United States spews out 17%, and India and Russia are each responsible for 5%. Thus, these four countries are responsible for half of the world's CO_2.[4]

Developing countries spend little on controlling pollution, and many developed countries' public expenditures on pollution regulation have fallen in recent years. In response, various protests have erupted. China and India now face regular pollution protests.

Nonetheless, politicians in countries around the world disagree about how and whether to control pollution. Most U.S. Congressional Democrats favor stronger pollution controls but most Republicans call for reducing such regulations. Australia imposed a tax on carbon in 2012, repealed it in 2014, and may reinstate it. Similar fights occur in Canada and European nations. Clearly, pollution control will be a major bone of contention throughout the world for the foreseeable future. The one bright spot is that

Table 18.1 Industrial CO_2 Emissions, 2011

	CO_2, Million Metric Tons	CO_2 Tons per Capita	CO_2 kg per \$100 GDP	Percentage Change in CO_2 Since 1990
China	9,020	6.6	65	267
United States	5,306	16.8	34	10
India	2,074	1.7	35	200
Russian Federation	1,808	12.6	56	−13[a]
Japan	1,188	9.3	27	9
Germany	729	8.8	21	−22[b]
Canada	485	14.1	36	12
United Kingdom	448	7.2	19	−19
Mexico	467	3.9	25	48
Australia	369	16.2	40	40
France	339	5.3	14	−10

[a]Since 1992; [b]Since 1991.

Source: CO_2 emissions in metric tons (CDIAC): **http://mdgs.un.org/unsd/mdg/Data.aspx** (viewed July 17, 2016).

---

[3]King Edward I established an air pollution commission in 1286 to reduce London's smog. At the commission recommendation, he banned burning coal in the city, with a punishment of torture or death. **https://www.epa.gov/aboutepa/londons-historic-pea-soupers**.

[4]**http://www.ucsusa.org/**.

the 195 countries that attended an international meeting on climate change in Paris in 2015 agreed to national goals to restrict emissions. The agreement was ratified in 2016. However, with the election of President Trump, the participation of the United States is in doubt.

Suppose that a government wants to regulate pollution and it has full knowledge about the marginal damage from pollution, the demand curve, costs, and the production technology. The government could optimally control pollution directly by restricting the amount of pollution that firms may produce or by taxing the pollution they create. A limit on the amount of air or water pollution that may be released is called an *emissions standard*. A tax on air pollution is an *emissions fee*, and a tax on pollution discharges into air or water is an *effluent charge*.

Frequently, however, a government controls pollution indirectly, through quantity restrictions or taxes on outputs or inputs. Whether the government restricts or taxes outputs or inputs may depend on the nature of the production process. It is generally better to regulate pollution directly than to regulate output, because direct regulation of pollution encourages firms to adopt efficient, new technologies to control pollution (a possibility we ignore in our example).

Regulation can effectively reduce pollution. Shapiro and Walker (2015) observe that emissions of the most common air pollutants from U.S. manufacturing fell by 60% between 1990 and 2008 even though U.S. manufacturing output increased substantially. They estimated that at least 75% of the reduction was due to environmental regulation.

Emissions Standard

We use the paper mill example in Figure 18.1 to illustrate how a government may use an *emissions standard* to reduce pollution. Here the government can achieve the social optimum by forcing the paper mills to produce no more than 84 units of paper per day. (Because output and pollution move together in this example, regulating either reduces pollution in the same way.)

Unfortunately, the government usually does not know enough to regulate optimally. For example, to set quantity restrictions on output optimally, the government must know how the marginal social cost curve, the demand for paper curve, and pollution vary with output. The ease with which the government can monitor output and pollution may determine whether it sets an output restriction or a pollution standard.

Even if the government knows enough to set the optimal regulation, it must enforce this regulation to achieve the desired outcome. The U.S. Environmental Protection Agency (EPA) tightened its ozone standard to 0.075 parts per million in 2008. As of 2015, 36 areas were marginally out of compliance with this rule, three moderately, three severely, and two extremely (the Los Angeles-South Coast Air Basin and the San Joaquin Valley, California).[5]

Emissions Fee

The government may impose costs on polluters by taxing their output or the amount of pollution produced. (Similarly, a law could make a polluter liable for damages in court.) In our paper mill example, taxing output works as well as taxing the pollution directly because the relationship between output and pollution is fixed. However, if firms can vary the output-pollution relationship by varying inputs or adding pollution-control devices, then the government should tax pollution.

---

[5]See **http://www3.epa.gov/airquality/greenbook** for details on noncompliance with EPA standards, and go to **http://scorecard.goodguide.com/** to learn about environmental risks in your area.

In our example, if the government knows the marginal cost of the gunk, MC^g, it can set the output tax equal to this marginal cost curve: $T(Q) = MC^g$. We write this tax as $T(Q)$ to show that it varies with output, Q. Figure 18.2 illustrates the manufacturers' after-tax marginal cost, $MC^s = MC^p + T(Q)$.

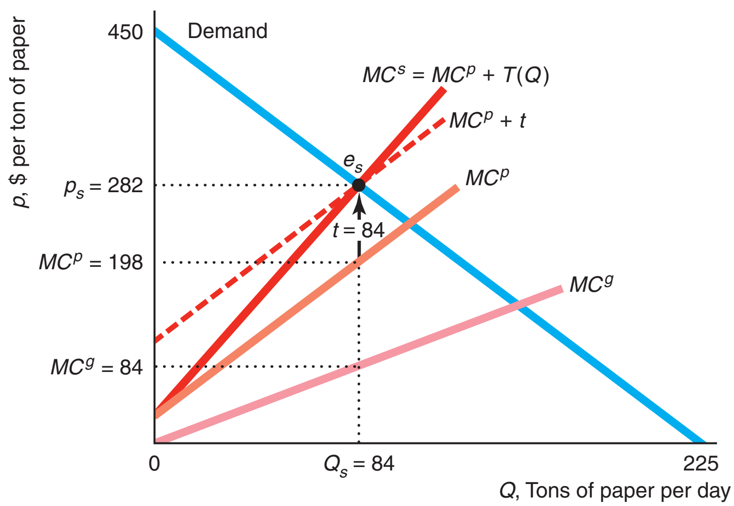
internalize the externality to bear the cost of the harm that one inflicts on others (or to capture the benefit that one provides to others)

The output tax causes a manufacturer to **internalize the externality**: to bear the cost of the harm that one inflicts on others (or to capture the benefit that one provides to others). The after-tax private marginal cost or supply curve is the same as the social marginal cost curve. As a result, the after-tax competitive equilibrium is the social optimum.

Usually, the government sets a specific tax rather than a tax that varies with the amount of pollution, as MC^g does. As Solved Problem 18.1 shows, applying an appropriate specific tax results in the socially optimal level of production.

Figure 18.2 Taxes to Control Pollution

Placing a tax on the firms equal to the harm from the gunk, $T(Q) = MC^g$, causes them to internalize the externality, so their private marginal cost is the same as the social marginal cost, MC^s. As a result, the competitive after-tax equilibrium is the same as the social optimum, e_s. Alternatively, applying a specific tax of t = \$84 per ton of paper, which is the marginal harm from the gunk at $Q_s = 84$, also results in the social optimum.



Solved Problem 18.1

For the market with pollution in Figure 18.1, what constant, specific tax, t, on output could the government set to maximize welfare?

Answer

Set the specific tax equal to the marginal harm of pollution at the socially optimal quantity. At the socially optimal quantity, $Q_s = 84$, the marginal harm from the gunk is \$84, as Figure 18.2 shows. If the specific tax is t = \$84, the after-tax private marginal cost (after-tax competitive supply curve), $MC^p + t$, equals the social marginal cost at the socially optimal quantity. Consequently, the after-tax competitive supply curve intersects the demand curve at the socially optimal quantity. By paying this specific tax, the firms internalize the cost of the externality at the social optimum. All that is required for optimal production is that the tax equals the marginal cost of pollution at the optimum quantity; it need not equal the marginal cost of pollution at other quantities.

Application

Why Tax Drivers

Driving causes many externalities including pollution, congestion, and accidents. Taking account of pollution from producing fuel and driving, Hill et al. (2009) estimated that burning one gallon of gasoline (including all downstream effects) causes a carbon dioxide-related climate change cost of 37¢ and a health-related cost of conventional pollutants associated with fine particulate matter of 34¢

A driver imposes delays on other drivers during congested periods. Parry et al. (2007) estimated that this cost is $1.05 per gallon of gas on average across the United States.

Edlin and Karaca-Mandic (2006) measured the accident externality from additional cars by the increase in the cost of insurance. These externalities are big in states with a high concentration of traffic but not in states with low densities. In California, with lots of cars per mile, an extra driver raises the total statewide insurance costs of other drivers by between $1,725 and $3,239 per year, and a 1% increase in driving raises insurance costs 3.3% to 5.4%. While the state could build more roads to lower traffic density and hence accidents, it's cheaper to tax the externality. A tax equal to the marginal externality cost would raise $66 billion annually in California—more than the $57 billion raised by all existing state taxes—and over $220 billion nationally. As of 2015, Germany, Austria, Slovakia, the Czech Republic, Poland, Hungary, and Switzerland have some form of a *vehicle miles traveled tax* (VMT), which is more clearly targeted at preventing congestion and accidents.

Vehicles are inefficiently heavy because owners of heavier cars ignore the greater risk of death that they impose on other drivers and pedestrians in accidents (Anderson and Auffhammer, 2014). Raising the weight of a vehicle that hits you by 1,000 pounds increases your chance of dying by 47%. The higher externality risk due to the greater weight of vehicles since 1989 is 26¢ per gallon of gasoline and the total fatality externality roughly equals a gas tax of between 97¢ and $2.17 per gallon.

Taking account of both carbon dioxide emissions and accidents, Sheehan-Connor (2015) estimates that the optimal flat tax is $1.14 per gallon. In 2014, the International Monetary Fund (IMF) estimated the optimal tax for the United States as $1.60 per gallon for gasoline and $2.10 for diesel.

To reduce the negative externalities of driving, governments have taxed gasoline, cars, and the carbon embodied in gasoline. However, such taxes have generally been much lower than the marginal cost of the externality and have not been adequately sensitive to vehicle weight or time of day.

Benefits Versus Costs from Controlling Pollution

The Clean Air Act of 1970 and the Clean Air Act Amendments of 1990 cleansed U.S. air. Between 1980 and 2015, the national average of sulfur dioxide (SO_2) plummeted 84%, carbon monoxide (CO) fell 84%, lead dove 92%, nitrogen dioxide (NO_2) tumbled 59%, and ozone dropped 32%.[6]

The EPA believes that the Clean Air Act saves over 160,000 lives a year, avoids more than 100,000 hospital visits, prevents millions of cases of respiratory problems, and saves 13 million lost workdays. The EPA (2011) estimated the costs of complying with the Clean Air Act were $53 billion, but the benefits were $1.3 trillion in 2010. Thus, the benefits outweighed costs by nearly 25 to 1.

Application

Protecting Babies

Some policy changes raise benefits and *lower* costs. For example, E-ZPass reduces congestion and pollution and increases babies' health. The E-ZPass, an electronic toll collection system on toll ways in New Jersey, Pennsylvania, and 12 other states, allows vehicles to pay a toll without stopping at a tollbooth.

[6]According to **https://www.epa.gov/air-trends** (viewed July 21, 2016).

Idling cars waiting to pay a toll create extra pollution and waste drivers' time. E-ZPass reduces delays at toll plazas by 85% and lowers NO_2 emissions from traffic by about 6.8%. Introducing E-ZPass reduced premature births by 11% and led to 12% fewer low-birthweight babies of mothers who lived within 2 kilometers (km) of a toll plaza relative to those who lived 2 to 10 km from a toll plaza (Currie and Walker, 2011).[7] Similarly, Knittel et al. (2016) found that lowering the amount of particulate matter by one unit (the average was 29 micrograms per cubic meter of air) saved 10 lives per 100,000 births, a decrease in the weekly infant mortality rate of about 4%.

18.4 Market Structure and Externalities

Two of our main results concerning competitive markets and negative externalities—that too much pollution is produced and that a tax equal to the marginal social cost of the externality solves the problem—do not hold for other market structures. Although a competitive market always produces too many negative externalities, a noncompetitive market may produce more or less than the optimal level of output and pollution. If a tax is set so that firms internalize the externalities, a competitive market produces the social optimum, whereas a noncompetitive market does not.

Monopoly and Externalities

We use the paper-gunk example to illustrate these results. In Figure 18.3, the monopoly equilibrium, e_m, is determined by the intersection of the marginal revenue, MR, and private marginal cost, MC^p, curves. Like the competitive firms, the monopoly ignores the harm its pollution causes, so it considers just its direct, private costs in making decisions.

Output is only 70 tons in the monopoly equilibrium, e_m, which is less than the 84 tons at the social optimum, e_s. Thus, this figure illustrates that *the monopoly outcome may be less than the social optimum even with an externality*.

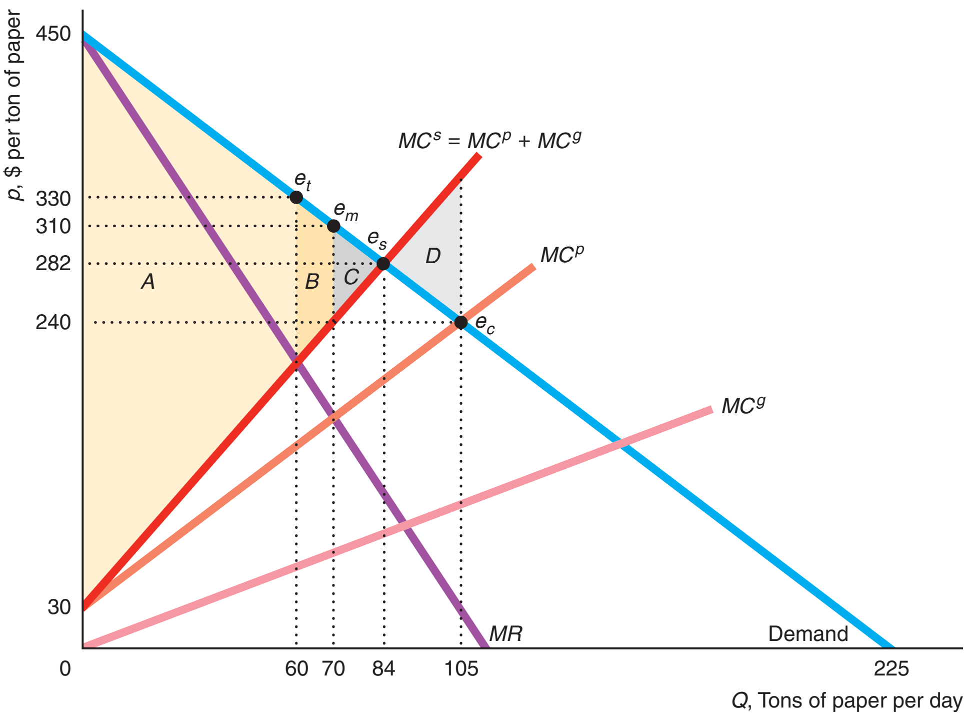
Although the competitive market with an externality always produces more output than the social optimum, a monopoly may produce more than, the same as, or less than the social optimum. The reason that a monopoly may produce too little or too much is that it faces two offsetting effects: The monopoly tends to produce too little output because it sets its price above its marginal cost, but the monopoly tends to produce too much output because its decisions depend on its private marginal cost instead of the social marginal cost.

Which effect dominates depends on the elasticity of demand for the output and on the extent of the marginal damage the pollution causes. If the demand curve is very elastic, the monopoly markup is small. As a result, the monopoly equilibrium is close to the competitive equilibrium, e_c, and greater than the social optimum, e_s. If extra pollution causes little additional harm—MC^g is close to zero at the equilibrium—the social marginal cost essentially equals the private marginal cost, and the monopoly produces less than the social optimum.

[7]Not only are babies at risk from auto pollution, so are older folks. Anderson (2015) estimated that doubling the percentage of time spent downwind of a highway increased the mortality rate of individuals 75 and older by 3.6% to 6.8%.

Figure 18.3 Monopoly, Competition, and Social Optimum with Pollution MyLab Economics Video

At the competitive equilibrium, e_c, more is produced than at the social optimum, e_s. As a result, the deadweight loss in the competitive market is *D*. The monopoly equilibrium, e_m, is determined by the intersection of the marginal revenue and the private marginal cost, MC^p, curves. The social welfare (based on the marginal social cost, MC^s, curve) under monopoly is $A + B$. Here the deadweight loss of monopoly, *C*, is less than the deadweight loss under competition, *D*.



Monopoly Versus Competitive Welfare with Externalities

In the absence of externalities, welfare is greater under competition than under an unregulated monopoly (Chapter 11). However, with an externality, welfare may be greater with monopoly than with competition.[8]

If both monopoly and competitive outputs are greater than the social optimum, welfare must be greater under monopoly because the competitive output is larger than the monopoly output. If the monopoly produces less than the social optimum, we need to check which distortion is greater: the monopoly's producing too little or the competitive market's producing too much.

Welfare is lower at monopoly equilibrium, area $A + B$, than at the social optimum, $A + B + C$, in Figure 18.3. The deadweight loss of monopoly, *C*, results from the monopoly's producing less output than is socially optimal.

In the figure, the deadweight loss from monopoly, *C*, is less than the deadweight loss from competition, *D*, so welfare is greater under monopoly. The monopoly produces only slightly too little output, whereas competition produces excessive output—and hence far too much gunk.

[8]Pennsylvania and sixteen other states have a wholesale or retail state monopoly over some or all types of alcoholic beverages. By charging high prices, they may reduce the externalities created by alcohol consumption, such as drunk driving.

Solved Problem 18.2

MyLab Economics Solved Problem

In Figure 18.3, what is the effect on output, price, and welfare of taxing the monopoly an amount equal to the marginal harm of the externality?

Answer

1. *Show how the tax affects the monopoly equilibrium*. A tax equal to the marginal cost of the pollution causes the monopoly to internalize the externality and to view the social marginal cost as its private cost. The intersection of the marginal revenue, MR, curve and the social marginal cost, MC^s, curve determines the taxed-monopoly equilibrium, e_t. The tax causes the equilibrium quantity to fall from 70 to 60 and the equilibrium price to rise from \$310 to \$330.

2. *Determine how this equilibrium movement affects the deadweight loss of monopoly*. The sum of consumer and producer surplus is only A after the tax, compared to $A + B$ before the tax. Thus, welfare falls. The difference between A and welfare at the social optimum, $A + B + C$, is $-(B + C)$, which is the deadweight loss from the taxed monopoly. The tax exacerbates the monopoly's tendency to produce too little output. The deadweight loss increases from C to $B + C$. The monopoly produced too little before the tax; the taxed monopoly produces even less.

Taxing Externalities in Noncompetitive Markets

Many people recommend that the government tax firms an amount equal to the marginal harm of pollution because such a tax achieves the social optimum in a competitive market. Solved Problem 18.2 shows that such a tax may lower welfare if applied to a monopoly. The tax definitely lowers welfare if the untaxed monopoly was producing less than the social optimum. If the untaxed monopoly was originally producing more than the social optimum, a tax may cause welfare to increase.

If the government has enough information to determine the social optimum, it can force either a monopolized or a competitive market to produce it. If the social optimum is greater than the unregulated monopoly output, however, the government has to subsidize (rather than tax) the monopoly to get it to produce as much output as is desired.

In short, trying to solve a negative externality problem is more complex in a noncompetitive market than in a competitive market. To achieve a social optimum in a competitive market, the government only has to reduce the externality, possibly by decreasing output. In a noncompetitive market, the government must eliminate problems arising from both externalities *and* the exercise of market power. Thus, the government needs more information to regulate a noncompetitive market optimally and may require more tools, such as a subsidy. To the degree that the problems arising from market power and pollution are offsetting, however, the failure to regulate a noncompetitive market is less harmful than the failure to regulate a competitive market.

18.5 Allocating Property Rights to Reduce Externalities

Instead of controlling externalities directly through emissions fees and emissions standards, the government may take an indirect approach by assigning a *property right*: an exclusive privilege to use an asset. If no one holds a property right for a good or a bad, the good or bad is unlikely to have a price. If you had a property right that assured you of the right to be free from air pollution, you could go to court to stop a nearby factory from polluting the air. Or you could sell your right, permitting the factory to pollute. If you did not have this property right, no one would be willing to pay you a positive price for it. Because of this lack of a price, a polluter's private marginal cost of production is less than the full social marginal cost.

Coase Theorem

Before Ronald Coase published his classic paper in 1960, economists, like other people, suffered from a

> **Common Confusion:** A polluter will necessarily pollute more if the government grants it the right to pollute than if the government grants the victim of pollution the right to be free from pollution.

According to the Coase Theorem, regardless of which party gets clearly defined property rights, a polluter and its victim may achieve the optimal levels of pollution if they can bargain freely.

To illustrate the Coase Theorem, we consider two adjacent firms, Alice's Auto Body Shop and Theodore's Tea House. The noise from the auto body shop hurts the tea house's business, as Table 18.2 illustrates. As the auto body shop works on more cars per hour, its profit increases, but the resulting extra noise reduces the tea house's profit. The last column shows the total profit of the two firms. Having the auto body shop work on one car at a time maximizes their joint profit: the socially optimal solution.

No Property Rights Initially, no one has clearly defined property rights concerning noise. Alice won't negotiate with Theodore. After all, why would she reduce her output and the associated noise, if Theodore has no legal right to be free of noise? Why would Theodore pay Alice to reduce the noise if he harbors the hope that the courts will eventually declare that he has a right to be free from noise pollution? Thus, Alice's shop works on two cars per hour, which maximizes her profit at 400. The resulting excessive pollution drives Theodore out of business, so their joint profit is 400.

Property Right to Be Free of Pollution Now, suppose that the courts grant Theodore the right to silence. He can force Alice to shut down, so that he makes 400 and their joint profit is 400. However, if Alice works on one car, her gain is 300, while Theodore's loss is 200. They should be able to reach an agreement where she pays him between 200 and 300 for the right to work on one car. As a result, they maximize their joint profit at 500.

Why doesn't Alice buy the rights to work on two cars instead of one? Her gain of 100 from working on the second car is less than Theodore's loss of 200, so they cannot reach a deal to let her work on the second car.

Property Right to Pollute Alternatively, suppose that the court says that Alice has the right to make as much noise as she wants. Unless Theodore pays her to reduce

Table 18.2 Daily Profits Vary with Production and Noise MyLab Economics Video

Auto Body Shop's Output, Cars per Hour	Profit, $		
	Auto Body Shop	Tea House	Total
0	0	400	400
1	300	200	500
2	400	0	400

the noise, he has to shut down. The gain to Theodore of 200 from Alice working on one rather than two cars is greater than the 100 loss to Alice. They should be able to reach a deal in which Theodore pays Alice between 100 and 200, she works on only one car, and they maximize their joint profit at 500.

Summary This example illustrates the three key results of the Coase Theorem:

1. Without clearly assigned property rights, one firm pollutes excessively and joint profit is not maximized.
2. Clearly assigning property rights results in the social optimum, maximizing joint profit, regardless of who gets the rights.
3. However, who gets the property rights affects how they split the joint profit. Because the property rights are valuable, the party without the property right compensates the party with the property rights.

Problems with the Coase Approach To achieve the socially optimal outcome, the two sides must bargain successfully with each other. However, the parties may not be able to bargain successfully for at least three important reasons.

First, if transaction costs are very high, it might not pay for the two sides to meet. For example, if a manufacturing plant pollutes the air, thousands or even millions of people may be affected. The cost of getting all of them together to bargain is prohibitive.

Second, if firms engage in strategic bargaining behavior, the firms may not be able to reach an agreement. For instance, if one party says, "Give me everything I want" and will not budge, reaching an agreement may be impossible.

Third, if either side lacks information about the costs or benefits of reducing pollution, the outcome is likely not to be optimal. It is difficult to know how much to offer the other party and to reach an agreement if you do not know how the polluting activity affects the other party.

For these reasons, Coasian bargaining is likely to occur in relatively few situations. Where bargaining cannot occur, the allocation of property rights affects the amount of pollution. Thus, the Coase Theorem is not a practical solution to most pollution problems. Rather, its value is that it demonstrates that a lack of clearly defined property rights is the root of the externality problem.

Application

Buying a Town

When the Environmental Protection Agency (EPA) stated that James Gavin American Electric Power was violating the Clean Air Act by polluting Cheshire, Ohio, the EPA effectively gave the residents the right to be free from pollution. To avoid the higher cost of litigation, installing new equipment, and other actions to reduce pollution at its plant, the company bought the town for $20 million, inducing the residents to pack up and leave. Thus, once clear property rights are established, a firm may find it less expensive to purchase those rights from others rather than incur endless litigation and pollution-reduction costs.

Markets for Pollution

If high transaction costs preclude bargaining, society may be able to overcome this problem by using a market, which facilitates exchanges between individuals. Starting in the early 1980s, the U.S. federal government, some state governments, and many governments around the world introduced cap-and-trade systems. Under a *cap-and-trade* system, the government distributes a fixed number of permits that allow firms to produce a specified amount of pollution. These permits not only create a property right to pollute, but they also limit or cap the total amount of pollution. Firms can trade these permits in a market, often by means of an auction. Firms that do not use all their permits sell them to other firms that want to pollute more—much as sinners bought indulgences in the Middle Ages.

Firms whose products are worth a lot relative to the harm from pollution they create buy rights from firms that make less valuable products. Suppose that the cost in terms of forgone output from eliminating each ton of pollution is $200 at one firm and $300 at another. If the government reduces the permits it gives to each firm so that each must reduce its pollution by 1 ton, the total cost is $500. With tradable permits, the first firm can reduce its pollution by 2 tons and sell one permit to the second firm, so the total social cost is only $400. The trading maximizes the value of the output for a given amount of pollution damage, thus increasing efficiency.

If the government knew enough, it could assign the optimal amount of pollution to each firm, and trading would be unnecessary. By using a market, the government does not have to collect this type of detailed information to achieve efficiency. It only has to decide how much total pollution to allow.

Application

Acid Rain Cap-and-Trade Program

The purpose of the Acid Rain Program, which was part of the 1990 U.S. Clean Air Act, was to reduce 10 million tons of sulfur dioxide (SO_2) and 2 million tons of nitrogen oxides (NO_x), the primary components of acid rain. It reduced the SO_2 level between 1990 and 2015 by 81%.

Under the law, the EPA issues SO_2 permits to firms that collectively equal an aggregate emission cap, which the EPA lowers over time. Each permit allows a firm to emit 1 ton of SO_2 annually. The government fines a firm $2,000 per ton on emissions above its allowance. If a company's emissions are less than its allowance, it may sell the extra permits to another firm, thus providing the firm with an incentive to reduce its emissions.

The EPA holds an annual spot auction for permits that firms may use in the current year and an advanced auction for permits effective in seven years. Anyone can purchase allowances. In some years, environmental groups, such as the Acid Rain Retirement Fund, the University of Tampa Environmental Protection Coalition, University of Tampa Environmental Protection Coalition, and Bates College Environmental Economics classes purchased permits and withheld them from firms to reduce pollution further. (You can see the outcome of the annual auctions at **http://www.epa.gov/airmarkets/so2-allowance-auctions**.)

According to some estimates, pollution reduction under this market program costs about a quarter to a third less than it would cost if permits were not tradable—a savings on the order of $225 to $375 million per year. Moreover, the EPA calculated the Acid Rain Program's annual benefits in 2010 at approximately $359 billion (in 2016 dollars), at an annual cost of about $8.4 billion, or a 43-to-1 benefit-to-cost ratio.

Markets for Positive Externalities

A market cannot solve the harms from negative externalities without regulation, taxation, or government intervention that clearly defines property rights. However, markets solve many positive externality problems without additional government intervention because property rights are usually clearly defined.

Bees pollinate oranges, almonds, avocados, and many other types of produce. If an orange orchard is located next to a commercial beekeeping firm, the beekeeper might not maintain enough bees to pollinate the oranges optimally because the beekeeper does not capture the full value from the bees. However, markets can solve this potential problem because beekeepers have a clearly defined property right to their bees. Orange farms hire beekeepers to bring hives to the farms during pollination season. This market transaction results in optimal pollination.

18.6 Rivalry and Exclusion

rival good
good that is used up as it is consumed

exclusion
others can be prevented from consuming a good

Until now, we've focused on *private goods*, which have the properties of rivalry and exclusion. A **rival good** is used up as it is consumed. If Jane eats an orange, that orange is gone so that no one else can consume it. **Exclusion** means that others can be prevented from consuming a good. If Jane owns an orange, she can easily prevent others from consuming that orange by locking it in her home. Thus, an orange is subject to rivalry and exclusion.

If a good lacks rivalry, everyone can consume the same good, such as clean air or national defense. If a market charges a positive price for that good, a market failure occurs because the marginal cost of providing the good to one more person is zero.

If the good lacks exclusion, such as clean air, no one can be prevented from consuming it because no one has an exclusive property right to the good. Consequently, a market failure may occur when people who don't have to pay for the good overexploit it, as when they pollute the air. If the market failure is severe—as it often is for open-access common resources and for public goods—governments may play an important role in provision or control of the good. For example, governments usually pay for streetlights.

We can classify goods by whether they exhibit rivalry and exclusion. Table 18.3 outlines the four possibilities: private good (rivalry and exclusion); open-access common property (rivalry, no exclusion); club good (no rivalry, exclusion); and public good (no rivalry, no exclusion).

Table 18.3 Rivalry and Exclusion

	Exclusion	*No Exclusion*
Rivalry	*Private good*: apple, pencil, computer, car	*Open-access common property*: fishery, freeway, park
No Rivalry	*Club good*: cable television, concert, tennis club	*Public good*: national defense, clean air, lighthouse

Open-Access Common Property

open-access common property
a resource that is nonexclusive and rival

An **open-access common property** is a resource that is nonexclusive and rival. Everyone has free access and an equal right to exploit this resource.

Many fisheries are open-access common properties. Fish are rival. Anyone can fish in an open-access fishery. Each fisher wants to land a fish to gain the property right to that fish, so that others can't claim it. The lack of clearly defined property rights while the fish are in the water leads to overfishing. Fishers have an incentive to catch more fish than they would if the fishery were private property.

Like polluting manufacturers, fishing boat owners look at only their private costs. In calculating these costs, they include the cost of boats, other equipment, a crew, and supplies. They do not include the cost that they impose on future generations by decreasing the stock of fish today, which reduces the number of fish in the sea next year. The fewer fish, the harder it is to catch any, so reducing the population today raises the cost of catching fish both now and in the future.

The social cost of catching a fish is the private cost plus the *externality cost* from reduced current and future populations of fish. Thus, the market failure arising from open-access common property is a negative externality.

In contrast, if each fisher owns a private pond, no externality occurs because the property rights are clearly defined. Each owner is careful not to overfish in any one year to maintain the stock (or number) of fish in the future.

Other important examples of open-access common property are petroleum, water, and other fluids and gases that firms extract from a *common pool*. Owners of wells drawing from a common pool compete to remove the substance most rapidly, to gain ownership of the good. This competition creates an externality by lowering fluid pressure, which makes further pumping more difficult. Iraq claimed that it invaded Kuwait (which led to the Persian Gulf War in 1991) because Kuwait was overexploiting common pools of oil underlying both countries.[9]

If many people try to access a single website at one time, congestion may slow traffic to a crawl. Similar problems occur on roads and freeways. If you own a car, you have a property right to drive that car but public roads and freeways are common property. Because you lack an exclusive property right to the highway on which you drive, you cannot exclude others from driving on the highway and must share it with them. However, each driver claims a temporary property right in a portion of the highway by occupying it, preventing others from occupying the same space. Competition for space on the highway leads to congestion, a negative externality that slows every driver.

To prevent overuse of a common resource, a government can clearly define property rights, restrict access, or tax users. Many developing countries over the past century have broken up open-access, common agricultural land into smaller private farms with clearly defined property rights. Governments frequently grant access to a resource on a first-come, first-served basis, such as at some popular national parks.

Alternatively, the government can impose a tax or fee to use the resource. Only those people who value the resource most gain access. For example, governments often charge an entrance fee to a park or a museum. By applying a tax or fee equal to the externality harm that each individual imposes on others (such as the value of increased congestion on a highway), the government forces each person to internalize the externality.

Application

Road Congestion

Roads and freeways belong to all of us. Anyone can drive on them for free. When many people try to use them at the same time, the resulting congestion—a negative externality—harms all drivers. That is, roads are a common resource that are overexploited or congested.

[9]Similarly, the state of Alaska proposed leasing land next to the federal Alaska National Wildlife Reserve, which would allow the leasing companies to drill and potentially drain oil from the Reserve, where drilling is prohibited. Taylor, Phil, "Alaska Unveils Plan to 'Drain' Federal Crude from ANWR," *E&E News*, June 30, 2011.

The cost in time, gasoline, and pollution from road congestion is extremely high in most developed nations. In the United States in 2015, commuters wasted 8 billion hours in traffic, which exceeds the time it would take to drive to Pluto (given the existence of a road and no congestion).

The typical U.S. consumer lost 50 hours, while those in the 10 worst corridors suffered from an annual 92 hours of delay. Belgium ranked second at 44 hours, followed by the Netherlands at 39, Germany at 38, Luxembourg at 33, Switzerland and the United Kingdom at 30, and France at 28. The worst city was London, where drivers lost an average of 101 hours—four days—in gridlock.

Charging to use roads, such as tolls on some bridges and highways, reduces the number of drivers, reducing congestion. For example, people must pay a toll to drive on the Golden Gate Bridge into San Francisco or through the Callahan Tunnel into Boston. London has a congestion charge for cars driving in Central London.

Club Goods

club good
a good that is nonrival but is subject to exclusion

A **club good** is a good that is nonrival but is subject to exclusion. Formal clubs, such as swim clubs or golf clubs, provide some club goods. These clubs exclude people who do not pay membership fees, but the services they provide, swimming or golfing, are nonrival: An extra person can swim or golf without reducing the enjoyment of others until these facilities become congested as capacity is reached.

However, the most significant club goods do not involve actual clubs. An important example is cable television. A cable television company, such as Comcast, can provide service to additional consumers at almost no additional cost (provided they have the cable in place). The service lacks rivalry as adding one more viewer does not impair the viewing experience of other viewers (and the marginal cost is nearly zero).[10] However, a cable television company can easily exclude people.

Only people who pay for the service receive the signal and can view the channel. The positive price that the cable company charges to view a channel exceeds the zero marginal cost of providing it. As a result, some cable subscribers who are willing to pay to view a channel (but less than the current price) cannot watch it, causing a deadweight loss to society.

Although club goods create a market failure, government intervention is rare because it is difficult for the government to help. As with regulation, an attempt to eliminate deadweight loss by forcing a cable television company to charge a price equal to its near-zero marginal cost would be self-defeating, as the service would not be produced and even more total surplus would be lost. A government could cap the cable TV price at average cost, which would reduce but not eliminate the deadweight loss.

Application

Software Piracy

One of the most important examples of a good that is not rival but does allow for exclusion is computer software, such as Microsoft Word. Software is nonrival. At almost no extra cost, Microsoft can provide a copy of the software program to another consumer. A market failure results because Microsoft charges a (high) positive price so it sells too few units.

[10] A club good with a zero marginal cost may be a natural monopoly Chapter 11.

However, if Microsoft cannot enforce its property right by preventing *pirating* of its software (use of software without paying for it), an even greater market failure may result: It may stop producing the product altogether. In countries where the cost of excluding nonpaying users is high, computer software is pirated and widely shared, which reduces the profitability of producing and selling software. The Business Software Alliance (BSA) estimated that the share of pirated software in 2015 was 39%. The rate was very high in many developing countries: 90% in Libya and Zimbabwe; between 80% and 90% in Armenia, Bangladesh, Indonesia, Iraq, Nicaragua, Pakistan; 70% in China; 64% in Russia; and 58% in India. The rate was between 21% and 32% for most EU countries, Australia, and Canada; 18% in Japan and New Zealand; and 17% in the United States. In 2013, Microsoft reported that 45% of counterfeit software came from the Internet.

Public Goods

public good
a good that is nonrival and nonexclusive

A **public good** is nonrival and nonexclusive. Clean air is a public good. One person's enjoyment of clean air does not stop other people from enjoying clean air as well, so clean air is nonrival. If we clean up the air, we cannot prevent others who live nearby from benefiting, so clean air is nonexclusive.

A public good is a special type of externality. If a firm reduces the amount of pollution it produces, cleaning the air, it provides a positive externality to its neighbors.

free riding
benefiting from the actions of others without paying

Free Riding Unfortunately, markets undersupply public goods due to a lack of clearly defined property rights. Because people who do not pay for the good cannot be excluded from consuming it, the provider of a public good cannot exercise property rights over the services provided by the public good. This problem is due to **free riding**: benefiting from the actions of others without paying. That is, free riders want to benefit from a positive externality. Consequently, it is very difficult for firms to provide a public good profitably because few people want to pay for the good no matter how valuable it is to them.

We illustrate the free rider problem using an example in which a market underprovides a public good. Two families, Family 1 and Family 2, live at the end of a road outside of town. They would both benefit equally from streetlights along their road. They must pay for the lights themselves.

Figure 18.4 shows the demand curves for the two families, D^1 and D^2. Each family's demand curve reflects its willingness to pay for a given number of lights. Family 1's demand curve, D^1, lies below Family 2's demand curve, D^2. For example, to install six streetlights, Family 1 is willing to pay \$80 each (point a) and Family 2 is willing to pay \$160 each (point b).

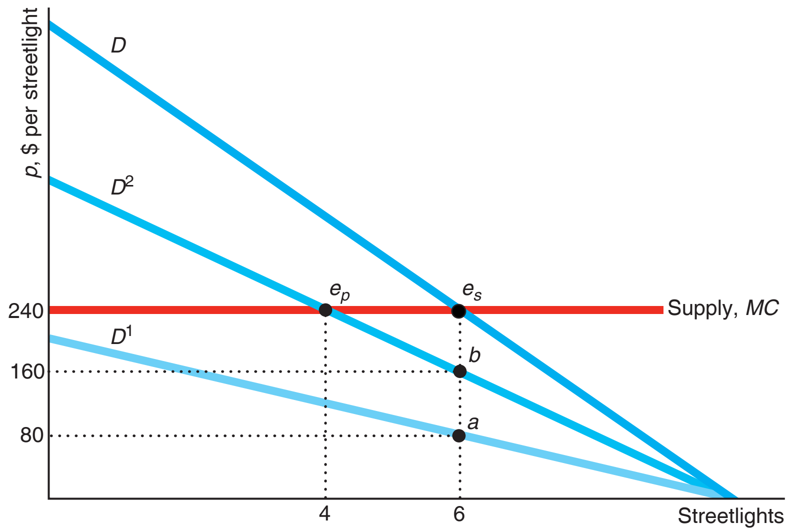
The figure also shows the market or social demand curve, D. The market demand curve for a public good differs from that for a private good.

The social marginal benefit of a private good is the same as the marginal benefit to the individual who consumes that good. Thus, the market demand or social marginal benefit curve for private goods is the *horizontal* sum of the demand curves of each individual (Chapter 2).

In contrast, the social marginal benefit of a public good is the sum of the marginal benefit to each person who consumes the good. Because a public good lacks rivalry, all people can get pleasure from the same unit of output. Consequently, the *social demand curve* or *social willingness-to-pay curve* for a public good is the *vertical* sum of the demand curves of each individual. For example, the social willingness to pay for six streetlights is \$240 (point e_s), which is the sum of the amounts each of the two families are willing to pay, \$80 + \$160.

Figure 18.4 Inadequate Provision of a Public Good MyLab Economics Video

Only two families live at the end of a road outside of town. They must pay to install streetlights. The social demand curve D is the vertical sum of the families' individual demand curves D^1 and D^2. The supply curve is horizontal at the marginal cost of \$240 per streetlight. The social optimum occurs where the social demand curve intersects the supply curve at e_s, where the families pay for six streetlights. If the families act independently, Family 2 buys four streetlights at e_p, where D^2 intersects the supply curve. The supply curve is everywhere above D^1, so Family 1 buys no streetlights and free rides on Family 2's four streetlights. Thus, the families buy too few streetlights if they act independently.



The competitive supply curve is horizontal at \$240 per streetlight, which is the social marginal cost. In the social optimum, e_s, the social marginal cost equals the social willingness to pay. That is, the social optimum, e_s, occurs where the competitive supply curve intersects the social demand curve. In the social optimum, the families buy six streetlights.

However, if the families act independently, they buy only four streetlights at the market equilibrium, e_p. The supply curve intersects D^2 at e_p, where Family 2 buys four streetlights. It is above D^1 everywhere, so Family 1 does not want to buy any streetlights: Family 1 free rides on Family 2's streetlights. Family 1 benefits from the streetlights without paying because the streetlights are a public good. Thus, the competitive market provides fewer streetlights, four, than the socially optimal six.

In more extreme cases, no public good is provided because people who don't pay cannot be stopped from consuming the good. Usually, if the government does not provide a nonexclusive public good, no one provides it.

Solved Problem 18.3

The only two stores in a mall decide whether to hire one guard or none—extra guards provide no extra protection. The guard patrolling the mall provides a service without rivalry, simultaneously protecting both stores. A guard costs 20 per hour. The benefit to each store is 16 per hour. The stores play a game in which they act independently. The table shows their payoffs. What is the outcome of this game? What is the social optimum?

		Store 1	
		Hire	*Do Not Hire*
Store 2	*Hire*	–4 –4	16 –4
	Do Not Hire	–4 16	0 0

Answer

1. *Use a best-response analysis* (Chapter 14) *to determine the Nash equilibrium to this game.*

   If Store 1 hires a guard, Store 2's payoff is −4 if it hires a guard and 16 if it does not, so its best response is to not hire. The light-green triangle in the lower-left cell shows this choice. Similarly, if the Store 1 does not hire a guard, Store 2's payoff is −4 if it hires and 0 if it does not, so its best response is to not hire. Thus, Store 2 has a dominant strategy of not hiring. Using a similar analysis, Store 1 also has a dominant strategy of not hiring (as the dark-green triangles show). The only Nash equilibrium to this game is for neither store to hire.

2. *Calculate the benefits and costs of hiring a guard to determine the social optimum.* The cost of a guard is 20, but the payoff to the two stores combined is 32, so it pays to hire a guard.

Comment: Acting independently, the stores do not achieve the social optimum because each firm tries to free ride. This game is an example of the prisoners' dilemma (Chapter 14).

Application

Free Riding on Measles Vaccinations

Measles vaccination is a public good. A person who gets a vaccination provides a positive externality to other free riding people by helping to limit the spread of the disease. Immunizing most of the population against measles reduces the risk of exposure for everyone in the community, including people who refuse vaccinations: The populace has *herd immunity*. That is, a vaccinated person provides a positive externality to others, lowering their probability of getting the disease.

In contrast, when too many people free ride by forgoing vaccination, the entire herd becomes more vulnerable. A person with the disease inflicts a negative externality on others. Measles is so contagious that 90% of people who are exposed become infected. One sick person typically infects 12 to 18 others who lack immunity.

The disease rapidly spreads in areas that lack herd immunity. Before the introduction of the measles vaccine in 1963, measles infected 90% of Americans by the time they were 15. The vaccine has prevented an estimated 35 million cases since 1963. The United States declared measles eliminated in 2000. However, travelers from other countries continue to import the disease.

The best estimate is that herd immunity requires at least a 92% to 94% vaccination rate. Vaccination of school kids is mandatory in Mississippi, where 99.7% of kindergarten students receive vaccinations, so the state has herd immunity. However, California, six other states, and Washington, D.C., have vaccination rates less than 90%, so they lack herd immunity. As a result, California had an outbreak of 136 cases of measles from December 2014 through April 2015.

Reducing Free Riding One solution to the free riding problem is for the government to provide the good. Governments provide public defense, roads, and many other public goods.

Alternatively, governmental or other collective actions can reduce free riding. Methods that may be used include social pressure, mergers, privatization, and compulsion.

Social pressure may reduce or eliminate free riding, especially for a small group. Such pressure may cause most firms in a mall to contribute "voluntarily" to hire security guards. The firms may cooperate in a repeated prisoners' dilemma game, especially if the market has relatively few firms.

A direct way to eliminate free riding by firms is for them to *merge* into a single firm, which internalizes the positive externality. The sum of the benefit to the individual stores equals the benefit to the single firm, so it makes the socially optimal decision.

If the independent stores sign a contract that commits them to share the cost of the guards, they achieve the practical advantage from a merger. However, the question remains as to why they would agree to sign the contract, given the prisoners' dilemma problem (Chapter 14).

Privatization—exclusion—eliminates free riding. A good that would be a public good if anyone could use it becomes a private good if access to it is restricted. An example is clean water, which water utilities can monitor and price using individual meters.

Another way to overcome free riding is through *mandates*. Some outside entity such as the government may mandate (dictate) a solution to a free-riding problem. For example, the management of a mall with many firms may require tenants to sign a rental contract committing them to pay fees to hire security guards that are determined through tenants' votes. If the majority votes to hire guards, all must share the cost. Although a firm might be unwilling to pay for the guard service if it has no guarantee that others will also pay, it may vote to assess everyone—including itself—to pay for the service.

Application

What's Their Beef?

Under U.S. federal law, agricultural producers can force all industry members to contribute to public goods if the majority of firms agrees. Under the Beef Promotion and Research Act, all beef producers must pay a \$1-per-head fee on cattle sold in the United States. The \$80 million raised by this fee annually finances research, educational programs on mad cow disease, and collective advertising, such as its original 2015 and 2016 "Beef: It's What's for Dinner" campaign and its 2012 campaign, "Stay Home. Grill Out." Supporters of this collective advertising estimate that producers receive \$5.67 in additional marginal revenue for every dollar they contribute.

Valuing Public Goods To ensure that a nonexclusive public good is provided, a government usually produces it or compels others to do so. Issues faced by a government when it provides such a public good include whether to provide it at all and, if so, how much of the good to provide. When grappling with these questions, the government needs to know the cost—usually the easy part—and the value of the public good to many individuals—the hard part.

Through surveys or voting results, the government may try to determine the value that consumers place on the public good. A major problem with these methods is that most people do not know how much a public good is worth to them. How much would you pay to maintain the National Archives? How much does reducing air pollution improve your health? How much better do you sleep at night knowing that the armed forces stand ready to protect you?

Even if people know how much they value a public good, they have an incentive to lie on a survey. Those who highly value the good and want the government to provide it may exaggerate its value. Similarly, people who place a low value on it may report *too low* a value—possibly even a negative one—to discourage government action.

Rather than relying on surveys, a government may ask its citizens to vote on public goods. Suppose that citizens use a separate, majority-rule vote to decide whether to install a traffic signal—a public good—at each of several street corners. If the majority vote to install the signal, all voters share the cost equally. An individual will vote to install a signal if the value of the signal to that voter is at least as much as the tax that each voter must pay for the signal.

Whether the majority votes for the signal depends on the preferences of the *median voter*: the voter with respect to whom half the populace values the project less and half values the project more. If the median voter wants to install a signal, then at least half the voters agree, so the vote carries. Similarly, if the median voter is against the project, at least half the voters are against it, so the vote fails.

It is *efficient* to install the signal if the value of the signal to society is at least as great as its cost. Does majority voting result in efficiency? The following examples illustrate that voting does not always result in an efficient outcome.

Each signal costs $300 to install. Three people vote and will be taxed to pay for the signal if it is approved. Thus, each individual votes for the signal only if that person puts a value on the signal of at least $100, which is the tax share each person pays for an installed signal. Table 18.4 shows the value that each voter places on installing a signal at each of three intersections.

For each of the proposed signals, Hayley is the median voter, so her views "determine" the outcome. If Hayley, the median voter, likes the signal, then she and Asa, a majority, vote for it. Otherwise, Nancy and Hayley vote against it. The majority favors installing a signal at corners *A* and *C* and is against doing so at corner *B*. It would be efficient to install the signal at corner *A*, where the social value is $300, and at corner *B*, where the social value is $375, because each value equals or exceeds the cost of $300.

At corner *A*, the citizens vote for the signal, and that outcome is efficient. The other two votes lead to inefficient outcomes. No signal is installed at corner *B*, where society values the signal at more than $300, but a signal is installed at corner *C*, where voters value the signal at less than $300.

The problem with yes–no votes is that they ignore the intensity of preferences. A voter indicates only whether the project is worth more than a certain amount. Thus, such majority voting fails to value the public good fully and hence does not guarantee that the public good is provided efficiently.[11]

Table 18.4 Voting on $300 Traffic Signals

	Value to Each Voter, $				
Signal Location	**Nancy**	**Hayley**	**Asa**	**Value to Society, $**	**Outcome of Vote***
Corner *A*	50	100	150	300	Yes
Corner *B*	50	75	250	375	No
Corner *C*	50	100	110	260	Yes

*An individual votes to install a signal at a particular corner if and only if that person thinks that the signal is worth at least $100, the tax that individual must pay if the signal is installed.

[11]Although voting does not reveal how much a public good is worth, Tideman and Tullock (1976) and other economists have devised taxing methods that can sometimes induce people to reveal their true valuations. However, these methods are rarely used.

Challenge Solution

Trade and Pollution

In the Challenge at the beginning of the chapter, we asked whether free trade benefits a country if it does not regulate its domestic pollution. This issue is increasingly important as nations move toward free trade and trade expands.

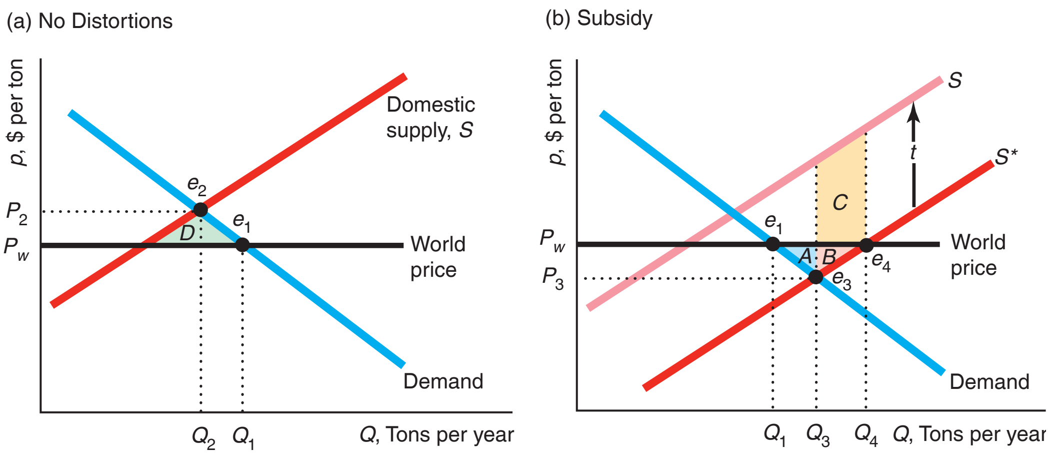
The United States has signed free-trade agreements (FTA) that eliminate or reduce tariffs and quotas and liberalize rules on foreign investment to increase trade with Australia, Bahrain, Canada, Chile, Costa Rica, El Salvador, Guatemala, Honduras, Israel, Jordan, Mexico, Morocco, Nicaragua, Peru, Singapore, South Korea, and other countries. As of 2014, FTA countries accounted for 47% of U.S. exports and 35% of imports. One of the most hotly debated issues in 2016 was whether the United States should sign a new trans-Pacific trade deal.

Liberalized trade has expanded trade. Trade was 30% of the U.S. gross domestic product 2012, compared to slightly less than 10% in 1965. The share of trade is even greater in many other countries: 30% in the European Union, 31% in Japan, 48% in India, 55% in China, 59% in Mexico and in the United Kingdom, and 63% in Canada.

Everyone can gain from free trade if losers are compensated and if domestic markets are perfectly competitive and not distorted by taxes, tariffs, or pollution (Chapters 9 and 10). Business and jobs lost in one sector from free trade are more than offset by gains in other sectors. However, if an economy has at least two market distortions, correcting one of them may either increase or decrease welfare.[12] For example, if a country bars trade and has uncontrolled pollution, then allowing free trade without controlling pollution may not increase welfare.

What are the welfare effects of permitting trade if a country's polluting export industry is unregulated? To analyze this question, we couple the trade model from Chapter 9 with this chapter's pollution model.

Suppose that the country's paper industry is a price taker on the world paper market. The world price is p_w. Panel a of the figure shows the gain to trade in a market without pollution (or optimally regulated pollution). The domestic supply curve, S, is upward sloping, and the home country can import as much as it wants at the world price, p_w. In the free-trade equilibrium, e_1, the equilibrium quantity is Q_1 and the equilibrium price is the world price, p_w. With a ban on imports, the equilibrium is e_2, quantity falls to Q_2, and price rises to p_2. Consequently, the deadweight loss from the ban is area D. (See the discussion of Figure 9.8 for a more thorough analysis.)



[12]In the economics literature, this result is referred to as the *Theory of the Second Best*.

In panel b, we include pollution in the analysis. The supply curve S^* is the sum of the firms' private marginal cost curves where the firms do not bear the cost of the pollution (similar to curve MC^p in Figure 18.1.) If the government imposes a specific tax, t, that equals the marginal cost of the pollution per ton of paper, then the firms internalize the cost of pollution, and the resulting supply curve is S (similar to MC^s in Figure 18.1).

If the government does not tax or otherwise regulate pollution, the private supply curve, S^*, lies below the social supply curve, which results in excess domestic production. If the government bans trade, the equilibrium is e_3, with a larger quantity, Q_3, than in the free-trade equilibrium, e_1, and a lower consumer price, p_3. Deadweight loss results because the true marginal cost (the height of the S curve at Q_3) exceeds the consumer price.

With free trade, the Theory of the Second Best tells us that welfare does not necessarily rise, because the country still has the pollution distortion. The free-trade equilibrium is e_4. Firms sell all their quantity, Q_4, at the world price, with Q_1 going to domestic consumers and $Q_4 - Q_1$ to consumers elsewhere. The private gain to trade—ignoring the government's cost of providing the subsidy—is area $A + B$. However, the expansion of domestic output increases society's cost due to excess pollution from producing Q_4 rather than Q_3, which is area $B + C$. The height of this area is the distance between the two supply curves, which is the marginal and average costs of the pollution damage (t), and the length is the extra output sold ($Q_4 - Q_3$). Thus, if area C is greater than area A, trade causes a net welfare loss. As the diagram is drawn, C is greater than A, so allowing trade lowers welfare if pollution is not taxed.

Should the country prohibit free trade? No, the country should allow free trade and regulate pollution to maximize welfare.

Summary

1. **Externalities.** An externality occurs when a consumer's well-being or a firm's production capabilities are directly affected by the actions of other consumers or firms rather than indirectly affected through changes in prices. An externality that harms others is a negative externality, and one that helps others is a positive externality. Some externalities benefit one group while harming another.

2. **The Inefficiency of Competition with Externalities.** Because producers do not pay for a negative externality such as pollution, the private costs are less than the social costs. Consequently, competitive markets produce more negative externalities than are optimal. If the only way to cut externalities is to decrease output, the optimal solution is to set output where the marginal benefit from reducing the externality equals the marginal cost to consumers and producers from less output. It is usually optimal to have some negative externalities, because eliminating all of them requires eliminating desirable outputs and consumption activities as well. If the government has sufficient information about demand, production cost, and the harm from the externality, it can use taxes or quotas to force the competitive market to produce the social optimum. It may tax or limit the negative externality, or it may tax or limit output.

3. **Regulating Externalities.** Governments may use emissions fees (taxes) or emissions standards to control externalities. If the government has full knowledge, it can set a fee equal to the marginal harm of the externality that causes firms to internalize the externality and produce the socially optimal output. Similarly, the government can set a standard that achieves the social optimum. However, if the government lacks full information, whether it should use a tax or fee depends on a number of factors.

4. **Market Structure and Externalities.** Although a competitive market produces excessive output and negative externalities, a noncompetitive market may produce more or less than the optimal level. With a negative externality, a noncompetitive equilibrium may be closer than a competitive equilibrium to the social optimum. A tax equal to the marginal social harm of a negative externality—which results in the social optimum when applied to a competitive market—may lower welfare when applied to a noncompetitive market.

5. **Allocating Property Rights to Reduce Externalities.** Externalities arise in the absence of clearly defined property rights. According to the Coase Theorem, allocating property rights to *either* of two parties results in an efficient outcome if the parties can bargain. The assignment of the property rights, however, affects income distribution, as the rights are valuable. Unfortunately, bargaining is usually not practical, especially when many people are involved. In such cases, markets for permits to produce externalities may overcome the externality problem. Markets often solve potential positive externality problems because property rights are clearly defined.

6. **Rivalry and Exclusion.** A private good is subject to rivalry—after the owner consumes a unit of the good, others cannot consume it—and to exclusion—the owner of the good can stop others from consuming it. Some goods lack one or both of these properties. Open-access common property, such as a fishery, is nonexclusive, but is subject to rivalry. This lack of exclusion causes overfishing because users of the fishery do not take into account the costs they impose on others (forgone fish) when they go fishing. A club good is nonrival but exclusive. For example, a swimming club lacks rivalry up to capacity but can exclude nonmembers. A market failure occurs if a firm charges a positive price for a good that it produces at zero marginal cost because it sells too few units. A public good such as public defense is both nonrival and nonexclusive. The lack of exclusion causes a free rider problem in a market: People use the good without paying for it. Therefore, potential suppliers of such goods receive inadequate compensation, so they underprovide the good. Because private markets tend to underprovide nonprivate goods, governments often produce or subsidize such goods.

Questions

Select questions are available on MyLab Economics*; * = answer appears at the back of this book;* **A** *= algebra problem;* **C** *= calculus problem.*

1. Externalities

1.1 According to a study in the *New England Journal of Medicine*, your friendships or “social networks” are more likely than your genes to make you obese. If it is true that people who have overweight friends are more likely to be overweight all else the same, is that an example of a negative externality? Why? (*Hints*: Is this relationship a causal one, or do heavier people choose heavier friends? Also remember that people with thinner friends may be thinner.)

1.2 Teamwork is a high priority for many employers. The ability to work efficiently in a group can foster creativity and trust, combine the strengths of individual teammates, reduce conflict, and promote a wider sense of ownership. In this context, would you consider productivity gains from cooperating with colleagues to be an example of a positive externality? Explain.

1.3 The November 2016 Market Report of the Australian Clean Energy Finance Corporation identifies facilities that turn urban waste into electricity as a major investment opportunity. About 23 million tons of urban waste per year is sent to landfills in Australia. Energy from waste projects can divert some of this landfill and produce reliable renewable energy. These projects are eligible for government support, and their costs of production are subsidized by state landfill levies. Is the urban waste a positive or a negative externality? Why is a market solution practical here?

1.4 A company has a research arm that works to advance technological information and knowledge and create new products or processes for its own benefit. The new technologies it has developed have often been implemented by other firms in the industry, often without compensation, and improved their productivity. Can social welfare be increased if the research arm of the company conducts additional research? Could the firms in the industry that benefit from these new technologies work together to solve the externality?

1.5 Zebra mussels are native to freshwater lakes and rivers of western Russia. Canal building to facilitate shipping around the turn of the eighteenth century allowed them to spread throughout Europe and Britain. In the late 1980s, Zebra mussels were introduced to Lake Erie in North America, likely through ballast water discharged from transoceanic ships. Within a few years, they had spread throughout the Great Lakes and into inland waterways, where they clog the water intake structures for power plants and water treatment plants, alter ecosystems, and threaten fish and wildlife. Did shipping provide an externality in this respect? What type?

2. The Inefficiency of Competition with Externalities

2.1 Why is zero pollution not the best solution for society? Can society have too little pollution? Why or why not?

2.2 In Figure 18.1, explain why area $D + E + H$ is the externality cost difference between the social optimum and the private equilibrium.

2.3 While genetically modified wheat is not yet being grown commercially, its production could help meet the growing worldwide demand for wheat. Proponents of genetic modification in wheat argue that it would increase yields by increasing resistance to parasites, infection, and adverse environmental conditions and could be engineered to carry additional nutrients. Opponents argue that it could adversely affect human health through increased allergens and resistance to antibiotics and reduce biodiversity. If genetically modified wheat is introduced to the market for wheat and it has both positive effects on yield and negative effects on human health and the environment, what would be the effects on social welfare?

3. Regulating Externalities

3.1 Various countries including Australia, Ireland, and the United States require that firms sell more fuel-efficient light bulbs (such as compact fluorescent bulbs) instead of incandescent light bulbs. The objective of these restrictions is to reduce carbon and global warming. What alternative approaches could countries use to achieve the same goals? What are the advantages and disadvantages of a ban relative to the alternatives?

*3.2 In the paper market example in this chapter, what are the optimal emissions fee and the optimal tax on output (assuming that only a single, constant fee or tax is applied)? (*Hint*: See Solved Problem 18.1.)

3.3 In Figure 18.2, the government may optimally regulate the paper market using a tax on output. A technological change drives down the private marginal cost of production. Discuss the welfare implications if the output tax is unchanged. (*Hint*: See Solved Problem 18.1.)

3.4 In Figure 18.1, could the government use a price ceiling or a price floor to achieve the optimal level of production?

3.5 According to scientists, one of the implications of climate change is rising ocean temperatures, which will reduce the survival rates for spawning salmon species and threaten the economic viability of this industry. Human activities that increase the amounts of greenhouses gases in the atmosphere are contributing to rising temperatures. What is the optimal policy for controlling air pollution and hence temperature change? What policy instruments could governments use to achieve the optimal policy?

*3.6 Using the numerical example in Appendix 18A, determine the social optimum if the marginal harm of gunk is $MC^g = \$84$ (instead of Equation 18A.3). Can you think of a shortcut that would allow you to solve this problem without algebra? **A**

3.7 Markowitz et al. (2012) found that limiting the number of liquor stores reduces crime. To maximize welfare taking into account the harms associated with alcohol sales, how should a regulatory agency set the number of liquor licenses? Should the profit-maximizing owner of a liquor store lobby for or against tighter restrictions on licenses?

3.8 Does texting while driving create a negative externality? If so, should the government implement a no-texting tax that would lead to a socially desirable level of texting while driving or should it ban texting while driving altogether?

3.9 As part of its commitment to greenhouse gas reduction, the Government of Ontario, Canada, offers incentives to increase the adoption of electric vehicles. For example, new electric vehicles with a battery capacity of over 16kWh and a sales price of under C\$150,000 are eligible for a rebate of up to C\$14,000. The Ford Focus Electric qualifies as an eligible vehicle. Its 2017 sales price starts at about C\$29,000 compared to about C\$16,000 for the gasoline-powered model. The Ontario government also offers a rebate of up to C\$1,000 for the purchase and installation of a Level 2 home charging station (full charge in 4 to 6 hours). Is the Ontario government's approach likely to be a cost-effective way to control pollution?

3.10 Employees in the United Kingdom who are sick for at least four days in a row are entitled to statutory sick leave of £89.35 per week for up to 28 weeks. Sick leave is paid by employers and is a taxable benefit (GOV.UK). As no amount is payable for the first three qualifying days, the average employee who is off sick for one week receives a statutory sick leave entitlement of under 9% of normal salary. (Of course, companies with sick-pay schemes would provide more.) In contrast, workers are paid 100% of their salary for at least a month in, for example, Luxembourg, Germany, Austria, Belgium, and Norway. Not surprisingly, the rate of absence from work for reasons of sickness is relatively low in Britain (*The Telegraph*, 2016). Evaluate the efficiency and welfare implications of these different sick-leave policies, taking account of externalities.

3.11 Northern Victoria, Australia, imposed a vomit tax on pubs in the Greater Shepparton area that remain open between 3:00 A.M. and 6:00 A.M. The tax was to be used to pay for cleaning up the mess left by drunks who get sick in the street. Pub owners objected that politicians assume that their customers are responsible for the mess. Discuss the pros and cons of using such a tax to deal with this externality.

3.12 Many jurisdictions strictly limit sales of hard liquor (liquor with a significantly higher alcohol content than wine) in an effort to limit the associated negative externalities (de Mello et al., 2013). One approach is to impose a high tax on sales of such products. Another approach is to require sellers to obtain licenses and to limit the number of licenses to the socially desirable number. Often the government sells these licenses to the highest bidder. However, in other jurisdictions, the government sets the price of licenses low enough that extensive excess demand for licenses occurs.

a. Under what circumstances would auctioning licenses be equivalent to a tax?

b. Why might regulators or politicians favor underpricing of liquor licenses? (*Hint*: Such licenses often end up in the hands of political donors or of friends and associates of donors.)

3.13 Countries around the world provide consumers with fossil fuel subsidies (Coady et al., 2016). In 2015, the subsidies reached $5.3 trillion (6.5% of global gross domestic product). Davis (2016) estimated that subsidies cause $4 billion in external costs annually. In a figure, show the effect of these subsidies and compare the post-subsidy equilibrium to the pre-subsidy equilibrium and to the social optimum, which accounts for pollution harms.

3.14 Houses that are built with high levels of insulation improve energy efficiency, help reduce the environmental costs of heating, and generate a positive externality. Many governments provide grants to accelerate investment programs of this type. Illustrate how subsidies affect a competitive market for energy-efficient homes.

3.15 Spam imposes significant negative externalities on email users (see the Application "Negative Externalities from Spam"). If the sender had to pay a small fee for each email, would the extent and impact of this negative externality decline? What effect would such a charge have on the net benefits arising from non-spam emails? How would the charge affect the proportion of spam in overall email traffic?

4. Market Structure and Externalities

4.1 Suppose that the only way to reduce pollution from paper production is to reduce output. The government imposes a tax equal to the marginal harm from the pollution on the monopoly producer. Show that the tax may raise welfare.

4.2 Suppose that the inverse demand curve for paper is $p = 200 - Q$, the private marginal cost (unregulated competitive market supply) is $MC^p = 80 + Q$, and the marginal harm from gunk is $MC^g = Q$.

a. What is the unregulated competitive equilibrium?

b. What is the social optimum? What specific tax (per unit of output or gunk) results in the social optimum?

c. What is the unregulated monopoly equilibrium?

d. How would you optimally regulate the monopoly? What is the resulting equilibrium? (*Hint*: See Solved Problems 18.1 and 18.2.) **A**

5. Allocating Property Rights to Reduce Externalities

5.1 List three specific examples where Coasian bargaining may result in the social optimum.

*5.2 In the example in Table 18.2, suppose that the two parties cannot negotiate. The government imposes a tax on the auto body shop equal to the marginal harm it does to the tea house. What is that tax schedule? Does it result in the welfare-maximizing outcome? How does the outcome change if the tax is imposed and they can bargain? (*Hint*: If Alice is paying the tax, then she has the right to produce at whatever level she wants.)

5.3 A man has a neighbor who owns a dog. The man has nothing against dogs as pets, but this one is yappy and barks incessantly. Let's say that the neighbor values the dog at its fair market value of £600 (we'll ignore any sentimental value the owner might have for the pet).

a. Suppose the man values peace and quiet at £1,000 but the neighbor has the legal right to keep a yappy dog. Can the man and his neighbor reach a mutually agreeable price for relocating the dog to a farm in the country? Explain.

b. If the man instead values peace and quiet at £400 but the neighbor has the legal right to keep a yappy dog, can the man and his neighbor reach a mutually agreeable price for resolving the situation? Explain.

c. If the man instead values peace and quiet at £400 and has the legal right to peace and quiet, can the man and his neighbor reach a mutually agreeable price for resolving the situation? Explain.

6. Rivalry and Exclusion

6.1 List three examples of goods that do not fit neatly into the categories in Table 18.3 because they are not strictly rival or exclusive.

6.2 A study of 13 main arterial roads in Delhi, India, in June 2017 by the Centre for Science and Environment found that "rush hour" is every day from 8 A.M. to 8 P.M., with average speeds not exceeding 30 km per hour 92% of the time. Are these roads common property resources? Can anything be done to mitigate the negative externalities they cause?

6.3 Are broadcast television and cable television public goods? Is exclusion possible? If either is a public good, why is it privately provided?

6.4 Do publishers sell the optimal number of intermediate microeconomics textbooks? Discuss in terms of public goods, rivalry, and exclusion.

6.5 To prevent overfishing, could one set a tax on fish or on boats? Explain and illustrate with a graph.

6.6 Two competitive retail stores are located next to one another. The demand curve for security services for the first store is $q_1 = 24 - 0.5p_1$ and for the second store is $q_2 = 24 - p_2$. The marginal cost of security services is $MC =$ €36 per hour. Compare the private equilibrium quantity of security services to the social equilibrium quantity of security services.

6.7 A city government is considering a proposal to build a footbridge over a waterway that will benefit the residents and firms situated in the neighborhoods adjacent to the footbridge on both sides of the river. If the demand for this public good for each of $n = 20{,}000$ identical individuals in the two neighborhoods is $Q = 500 - 0.5p_i$, where p_i is each individual's willingness to pay for the footbridge, what is the social demand curve, and how much would the residents be willing to pay collectively for the footbridge? **A**

6.8 A pandemic influenza (akin to the severe 1918 flu epidemic) that kills 2 million or more people could also cause annual income loss of 4% to 5% of global national income (see Fan et al., 2016). Using the concepts in this chapter (such as externalities, free riding, and public goods), explain what a pandemic is. Is government intervention necessary to help prevent pandemics, or can society rely solely on markets? (*Hint*: See the Application "Free Riding on Measles Vaccinations.")

6.9 Patent trolls are firms that buy patents hoping to bring patent infringement lawsuits against major firms rather than to producing goods themselves. Recently, 53 firms such as Google, Microsoft, Ford Motor Company, JP Morgan Chase, Solar City, and Uber have joined the LOT (License of Transfer) Network.[13] These companies control about 360,000 patents. If any fall into the hands of a troll, these companies automatically cross-license the patent to all members so that the troll cannot sue them. Ken Seddon, LOT executive director, says that this program is similar to the herd immunity conferred by vaccines. Moreover, the more companies that join, the more attractive it is for additional companies to join. Explain this reasoning. (*Hint*: See the Application "Free Riding on Measles Vaccinations.")

6.10 You and your roommate have a stack of dirty dishes in the sink. Either of you would wash the dishes if the decision were up to you; however, neither will do it in the expectation (hope?) that the other will deal with the mess. Explain how this example illustrates the problem of public goods and free riding.

*6.11 Purchasing a franchise provides access to an established company's brand name. It also means that franchisees can benefit from brand promotion. Accordingly, franchisees are typically required to pay an advertising fee, which may be fixed or a percentage of sales to the franchisor who then spends the funds on advertising programs for the benefit of the entire franchise. If the franchise fee is set at 2% of each franchisee's sales and it is estimated that marginal revenue for the franchisees increases by 3% as a result of the collective advertising, is the franchisor advertising optimally? Explain your answer.

6.12 Anna and Bess are assigned to write a joint paper within a 24-hour period about the Pareto optimal provision of public goods. Let t_A denote the number of hours that Anna contributes to the project and t_B the number of hours that Bess contributes. The numeric grade that Anna and Bess earn is a function, $23 \ln (t_A + t_B)$, of the total number of hours that they contribute to the project. If Anna contributes t_A, then she has $(24 - t_A)$ hours in the day for leisure. Anna's utility function is $U_A = 23 \ln (t_A + t_B) + \ln (24 - t_A)$, and Bess' utility function is $U_B = 23 \ln (t_A + t_B) + \ln (24 - t_B)$. If they choose the hours to contribute simultaneously and independently, what is the Nash equilibrium number of hours that each will provide? What is the number of hours each should contribute to the project that maximizes the sum of their utilities? **C**

6.13 In Solved Problem 18.3, suppose that the firms will split the cost of a guard if they both vote to hire one. Show the new payoff matrix. Do they hire a guard?

7. Challenge

*7.1 Redraw panel b of the Challenge Solution figure to show that it is possible for trade to increase welfare even when pollution is not taxed or otherwise regulated.

*7.2 In the Challenge Solution, in the absence of pollution as in panel a of the figure, how do we know that winners from trade can compensate losers and still have enough left over to benefit themselves?

[13] Carolyn Said, "Tech, Auto Companies Join Forces to Thwart Patent Trolls," *San Francisco Chronicle*, January 28, 2016.

Asymmetric Information

19

The buyer needs a hundred eyes, the seller not one. —George Herbert (1651)

Challenge

Dying to Work

In part because of the differing amounts that firms invest in safety, jobs in some firms are more dangerous than in others. In 2010, the United States had a particularly bad year for major disasters: A blast at a refinery in Washington State killed 7 workers, 11 workers died when BP's Deepwater Horizon oil rig exploded in the Gulf of Mexico, and 29 coal miners died in a disaster at Massey Energy's West Virginia mine explosion. These disasters are only the tip of the iceberg as thousands of workers are killed on the job every year—4,821 in 2014. Major disasters have occurred in other countries. An apparel factory collapse in Bangladesh killed 1,129 workers in 2013. A warehouse explosion in the port of Tianjin, China, killed 173 workers in 2015. A factory blast killed 23 workers in Bangladesh in 2016.

Prospective employees often do not know the injury rates at individual firms but they may know the *average* injury rate over an entire industry, in part because government reports such statistics. Injury rates vary dramatically by industry, ranging from only 0.9 fatal injuries per 100,000 workers in the financial services industry to 9.7 in construction and 12.4 in mining. Some occupations are particularly dangerous. Logging had a fatal injury rate of 91.3 per 100,000 workers; fishing 75.0; agriculture, 22.9; and driving a truck, 23.6. On the other hand, safe occupations include sales, 1.6, and educational services, 0.7 (although students sometimes risk dying of boredom).[1]

If people are rational and fear danger, they agree to work in a dangerous job only if that job pays a sufficiently higher wage than less-risky alternative jobs. Economists have found that workers receive *compensating wage differentials* in industries and occupations that government statistics show are relatively risky.

However, if workers are unaware of the greater risks at certain firms within an industry, they may not receive compensating wage differentials from more dangerous employers within that industry.[2] Workers are likely to have a sense of the risks associated with an industry: Everyone knows that mining is relatively risky—but they do not know which mining companies are particularly risky until a major accident occurs. For example, in the decade before Alpha Natural Resources acquired Massey Energy in 2011, 54 coal miners were killed in

[1]Government statistics also tell us that males have an accident rate, 6.0. That is nearly an order of magnitude greater than females, 0.7. Some of this difference is due to different occupations and some to different attitudes toward risk. How many women are injured after saying, "Hey! Watch this!"?

[2]Of course, a few good employers may gain a positive reputation. In 2012, Redhook Brewery started selling a special beer in memory of an employee who was killed when a keg exploded, with the proceeds going to his family.

Massey mines, a much higher rate than at other mines, yet no evidence shows that these workers received higher pay than did workers at other mining firms.[3]

Because workers do not know which firms are safer than others, each firm bears the full cost of its safety investments but does not get the full benefits. If workers are aware of the average risk in an industry, all firms benefit from one firm's safety investment because that investment improves the industry average. Thus, other firms share the benefit from one firm's investment in safety. Consequently, firms, when making the important strategic decision of how much they should invest in safety, take this spillover effect into account.

Does a firm underinvest in safety if it knows how dangerous a job is but potential employees do not? Can government intervention overcome such safety problems?

So far, we've examined models in which everyone is equally knowledgeable or equally ignorant about prices, product quality, and other factors relevant to a transaction. That is, they have *symmetric information*. Even in uncertainty models (Chapter 16), companies that sell insurance and the people who buy it may be equally uncertain about future events.

asymmetric information one party to a transaction has relevant information that another party lacks

In contrast, in this chapter, people have **asymmetric information**: One party to a transaction has relevant information that another party lacks. For example, the seller knows the quality of a product and the buyer does not.

hidden characteristic an attribute of a person or thing that is known to one party but unknown to others

Two important types of asymmetric information are *hidden characteristics* and *hidden actions*. A **hidden characteristic** is an attribute of a person or thing that is known to one party but unknown to others. For example, the owner of a property may know the mineral composition of the land, unlike a mining company that is considering buying the land.

hidden action an act by one party to a transaction that the other party cannot observe

A **hidden action** is an act by one party to a transaction that the other party cannot observe. An example is a firm's manager using a company jet for personal use without the owners' knowledge.

opportunistic behavior one party takes economic advantage of another when circumstances permit

When both parties to a transaction have equal information or equally limited information, neither has an advantage over the other. In contrast, asymmetric information leads to **opportunistic behavior**, where one party takes economic advantage of another when circumstances permit. Such *opportunistic behavior* due to asymmetric information leads to market failures, and destroys many desirable properties of competitive markets.

adverse selection occurs when one party to a transaction possesses information about a hidden characteristic that is unknown to other parties and takes economic advantage of this information

Two problems of opportunistic behavior arise from asymmetric information. One—*adverse selection*—is due to hidden characteristics, while the other—*moral hazard*—is associated with hidden actions. The problem of **adverse selection** arises when one party to a transaction possesses information about a hidden characteristic that is unknown to other parties and takes economic advantage of this information. For example, if a roadside vendor sells a box of oranges to a passing motorist and only the vendor knows that the oranges are of low quality, the vendor may allege that the oranges are of high quality and charge a premium price for them. That is, the seller seeks to benefit from an informational asymmetry due to a hidden

[3]Massey had 515 violations in 2009, and the U.S. Mine Safety and Health Administration issued Massey 124 safety-related citations in 2010 prior to the April 2010 accident at Massey's Upper Big Branch mine in West Virginia that killed 29 workers. Mine Safety and Health Administration safety officials concluded in 2011 that Massey could have prevented the 2010 explosion that took 29 lives. The former head of security at the mine was convicted of two felonies and sentenced to 36 months in prison. In 2016, Don Blankenship, the former Massey Energy head, was convicted and given a one-year sentence.

characteristic, the quality of the oranges. If potential buyers worry about such opportunistic behavior, they may be willing to pay only low prices or may forgo purchasing the oranges entirely.

moral hazard
an informed party takes an action that the other party cannot observe and that harms the less-informed party

The primary problem arising from hidden action is **moral hazard**, which occurs when an informed party takes an action that the other party cannot observe and that harms the less-informed party. If you pay a mechanic by the hour to fix your car, and you do not actually watch the repairs, then the time spent by the mechanic on your car is a hidden action.[4] Moral hazard occurs if the mechanic bills you for excessive hours.

This chapter focuses on adverse selection and unobserved characteristics. Adverse selection often leads to markets in which some desirable transactions do not take place or even the market as a whole cannot exist. We also discuss possible solution to adverse selection problems. Chapter 20 concentrates on moral hazard problems due to unobserved actions and on the use of contracts to deal with them.

In this chapter, we examine five main topics

1. **Adverse Selection.** Adverse selection may prevent desirable transactions from occurring, possibly eliminating a market.
2. **Reducing Adverse Selection.** To reduce the harms from adverse selection, government actions or contracts between involved parties may equalize information between the parties or restrict the informed party's ability to exploit the uninformed party.
3. **Price Discrimination Due to False Beliefs About Quality.** If some consumers incorrectly think that quality varies across identical products, a firm may price discriminate.
4. **Market Power from Price Ignorance.** Consumers' ignorance about the price that each firm charges gives firms market power.
5. **Problems Arising from Ignorance When Hiring.** Attempts to eliminate information asymmetries in hiring may raise or lower social welfare.

19.1 Adverse Selection

One of the most important problems associated with adverse selection is that consumers may not make purchases to prevent better-informed sellers from exploiting them. As a result, not all desirable transactions occur, and potential consumer and producer surplus is lost. Indeed, in the extreme case, adverse selection may prevent a market from operating at all. We illustrate this idea using two important examples of adverse selection problems: insurance and products of varying quality.

Insurance Markets

Hidden characteristics and adverse selection are very important in the insurance industry. Were a health insurance company to provide fair insurance by charging everyone a rate for insurance equal to the average cost of health care for the entire population, the company would lose money due to adverse selection. Many unhealthy people—people who expect to incur health care costs that are higher than

[4] A lawyer dies in an accident and goes to heaven. A host of angels greets him with a banner that reads, "Welcome Oldest Man!" The lawyer is puzzled: "Why do you think I'm the oldest man who ever lived? I was only 47 when I died." One of the angels replied, "You can't fool us; you were at least 152 when you died. We saw how many hours you billed!"

average—would view this insurance as a good deal and would buy it. In contrast, unless they were very risk averse, healthy people would not buy it because the premiums would exceed their expected health care costs. Given that a disproportionately large share of unhealthy people would buy the insurance, the market for health insurance would exhibit adverse selection, and the insurance company's average cost of medical care for covered people would exceed the population's average.

Adverse selection results in an inefficient market outcome, which does not maximize the sum of producer and consumer surpluses. The loss of potential surplus occurs because some potentially beneficial sales of insurance to relatively healthy individuals do not occur. These consumers would be willing to buy insurance at a lower rate that was closer to the fair rate for them given their superior health. The insurance company would be willing to offer such low rates only if it could be sure that these individuals were relatively healthy.

Products of Unknown Quality

Anagram for General Motors: or great lemons

Adverse selection often arises because sellers of a product have better information about the product's quality—a hidden characteristic—than do buyers. Used cars that appear to be identical on the outside often differ substantially in the number of repairs they will need. Some cars—*lemons*—are cursed. They have a variety of insidious problems that become apparent to the owner only after driving the car for a while. Thus, the owner knows from experience if a used car is a lemon, but a potential buyer does not.

If buyers have the same information as sellers, no adverse selection problem arises. However, when sellers have more information than do buyers, adverse selection is likely to occur. In this case, many people believe that

> **Common Confusion:** If consumers know that some goods are low quality and others are high quality, but not which is which, all the goods will sell for the average price of the two types of goods.

As intuitively appealing this belief is, it is generally untrue.

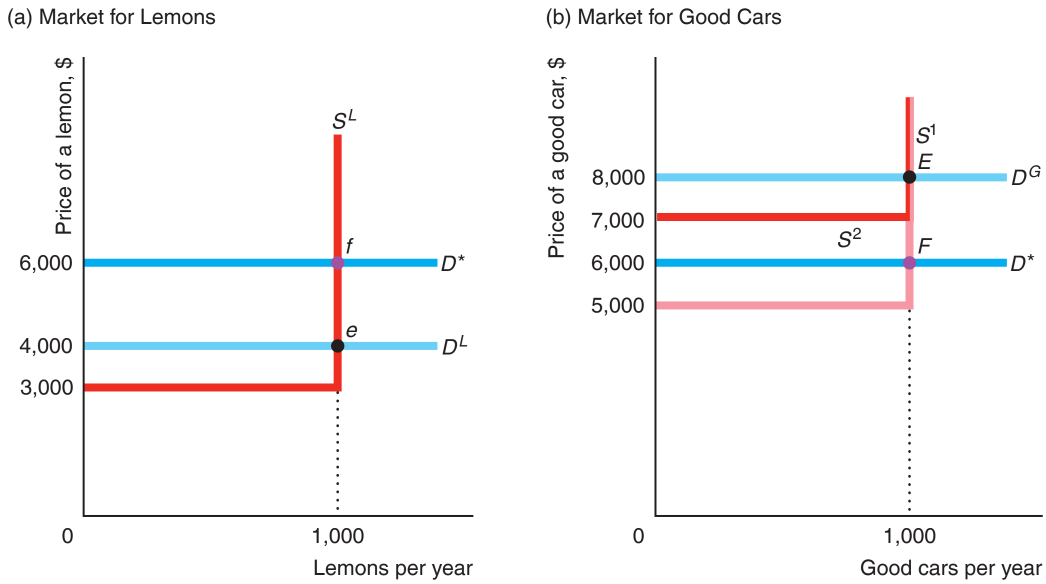
If consumers don't know the quality before they buy a good, the presence of low-quality goods in the market may drive high-quality products out of the market (Akerlof, 1970). Why? Used car buyers worry that a used car might be a lemon. As a result, they will not pay as high a price as they would if they knew the car was of good quality. They will only buy if the price is low enough to reflect the possibility of getting a lemon. Given that sellers of excellent used cars do not want to sell their cars for that low a price, they do not enter the market. Adverse selection has driven the high-quality cars out of the market, leaving only the lemons.

In the following example, we assume that sellers cannot alter the quality of their used cars and that the number of potential used car buyers is large. All potential buyers are willing to pay \$4,000 for a lemon and \$8,000 for a good used car: The demand curve for lemons, D^L, is horizontal at \$4,000 in panel a of Figure 19.1, and the demand curve for good cars, D^G, is horizontal at \$8,000 in panel b.

Although the number of potential buyers is virtually unlimited, only 1,000 owners of lemons and 1,000 owners of good cars are willing to sell. The *reservation price* of lemon owners—the lowest price at which

Figure 19.1 Markets for Lemons and Good Cars MyLab Economics Video

If everyone has full information, the equilibrium in the lemons market is *e* (1,000 cars sold for \$4,000 each), and the equilibrium in the good-car market is *E* (1,000 cars sold for \$8,000 each). If buyers can't tell quality before buying but assume that equal numbers of the two types of cars are for sale, their demand in both markets is D^*, which is horizontal at \$6,000. If the good-car owners' reservation price is \$5,000, the supply curve for good cars is S^1, and 1,000 good cars (point *F*) and 1,000 lemons (point *f*) sell for \$6,000 each. If their reservation price is \$7,000, the supply curve is S^2. No good cars sell; 1,000 lemons sell for \$4,000 each (point *e*).



they will sell their cars—is \$3,000. Consequently, the supply curve for lemons, S^L in panel a, is horizontal at \$3,000 up to 1,000 cars, where it becomes vertical (no more cars are for sale at any price). The reservation price of owners of high-quality used cars is v, which is less than \$8,000. Panel b shows two possible values of v. If $v = \$5{,}000$, the supply curve for good cars, S^1, is horizontal at \$5,000 up to 1,000 cars and then becomes vertical. If $v = \$7{,}000$, the supply curve is S^2.

Market Equilibrium with Symmetric Information If both sellers and buyers know the quality of all the used cars before any sales take place—they have full, symmetric information—all 2,000 cars sell, and the good cars sell for more than lemons. In panel a of Figure 19.1, the intersection of the lemons demand curve D^L and the lemons supply curve S^L determines the equilibrium at *e* in the lemons market, where 1,000 lemons sell for \$4,000 each. Regardless of whether the supply curve for good cars is S^1 or S^2 in panel b, the equilibrium in the good-car market is *E*, where 1,000 good cars sell for \$8,000 each.

If everyone has the same information, all cars sell. As a result, *the market is efficient because the goods go to the people who value them the most*. All current owners, who value the cars less than do the potential buyers, sell their cars.

It does not matter whether all buyers and sellers have full information or all lack information—it's the equality (or symmetry) of information that matters. However, *the amount of information they have affects the price at which the cars sell*. With full information, good cars sell for \$8,000 and lemons for \$4,000.

If no one can tell a lemon from a good car at the time of purchase, both types of cars sell for the same price. Suppose that everyone is risk neutral (Chapter 17) and no one can identify the lemons: Buyers *and* sellers are equally ignorant. A buyer has an equal chance of buying a lemon or a good car. The expected value (Chapter 17) of a used car is

$$\$6{,}000 = \left(\tfrac{1}{2} \times \$4{,}000\right) + \left(\tfrac{1}{2} \times \$8{,}000\right).$$

A risk-neutral buyer would pay \$6,000 for a car of unknown quality. Because sellers cannot distinguish between the cars either, sellers accept this amount and sell all the cars.[5] Thus, this market is efficient because the cars go to people who value them more than their original owners.

If only lemons sell, they would sell for \$4,000. The presence of good-quality cars raises the price received by sellers of lemons to \$6,000. Similarly, if only good cars sell, they would sell for \$8,000. The presence of lemons lowers the price that sellers of good cars receive to \$6,000. Thus, *sellers of good-quality cars are effectively subsidizing sellers of lemons*.

Market Equilibrium with Asymmetric Information If sellers know the quality but buyers do not, this market may be inefficient: The better-quality cars may not sell even though buyers value good cars more than sellers do. The equilibrium in this market depends on whether the value that the owners of good cars place on their cars, v, is greater or less than the expected value of buyers, \$6,000. The two possible equilibria are (1) all cars sell at the average price, or (2) only lemons sell at a price equal to the value that buyers place on lemons.

Initially, we assume that the sellers of good cars value their cars at $v = \$5{,}000$, which is less than the buyers' expected value of the cars (\$6,000), so that transactions occur. The equilibrium in the good-car market is determined by the intersection of S^1 and D^* at F in panel b of Figure 19.1, where 1,000 good cars sell at \$6,000. Similarly, owners of lemons, who value their cars at only \$3,000, are happy to sell them for \$6,000 each. The new equilibrium in the lemons market is f.

Thus, all cars sell at the same price. Consequently, here *asymmetric information does not cause an efficiency problem, but it does have equity implications*. Sellers of lemons benefit and sellers of good cars suffer from consumers' inability to distinguish quality. Consumers who buy the good cars get a bargain, and buyers of lemons have a sour taste in their mouths.

Now suppose that the sellers of good cars place a value of $v = \$7{,}000$ on their cars and thus are unwilling to sell them for \$6,000. As a result, in this case, the *lemons drive good cars out of the market*. Buyers realize that they can buy only lemons at any price less than \$7,000. Consequently, in equilibrium, the 1,000 lemons sell for the expected (and actual) price of \$4,000, and no good cars change hands. *This equilibrium is inefficient* because high-quality cars remain in the hands of people who value them less than potential buyers do.

In summary, if buyers have less information about product quality than sellers do, the result might be a lemons problem in which high-quality cars do not sell even though potential buyers value those cars more than their current owners do. If so, the asymmetric information causes a competitive market to lose its desirable efficiency and welfare properties. The lemons problem does not occur if the information is symmetric. If buyers and sellers of used cars know the quality of the cars, each car

[5]Risk-neutral sellers place an expected value of $\left(\tfrac{1}{2} \times \$3{,}000\right) + \tfrac{1}{2}v = \$1{,}500 + \tfrac{1}{2}v$ on a car of unknown quality. If $v = \$7{,}000$, this expected value is \$1,500 + \$3,500 = \$5,000. If $v = \$5{,}000$, the expected value is only \$4,000. In either case, sellers would be happy to sell their cars for \$6,000.

sells for its true value in a perfectly competitive market. If, as with new cars, neither buyers nor sellers can identify lemons, both good cars and lemons sell at a price equal to the expected value rather than at their (unknown) true values.

Solved Problem 19.1

MyLab Economics Solved Problem

Suppose that everyone in our used car example is risk neutral, and potential car buyers value lemons at \$4,000 and good used cars at \$8,000. The reservation price of lemon owners is \$3,000 and the reservation price of owners of high-quality used cars is \$7,000. The share of current owners who have lemons is θ (in our previous example, the share was $\theta = \frac{1}{2}$). For what values of θ do all potential sellers sell their used cars? Describe the equilibrium.

Answer

1. *Determine how much buyers are willing to pay if all cars sell.* Because buyers are risk neutral, if they believe that the probability of getting a lemon is θ, the most they are willing to pay for a car of unknown quality is the average price,

$$p = \$4{,}000\theta + (\$8{,}000 \times [1 - \theta]) = \$8{,}000 - \$4{,}000\theta. \qquad (19.1)$$

For example, $p = \$6{,}000$ if $\theta = \frac{1}{2}$, and $p = \$7{,}000$ if $\theta = \frac{1}{4}$.

2. *Solve for the values of* θ *such that all the cars sell, and describe the equilibrium.* All owners will sell if the market price equals or exceeds the reservation price of good car owners, \$7,000. Using Equation 19.1, we know that the market (equilibrium) price is \$7,000 or more if a quarter or fewer of the used cars are lemons, $\theta \le \frac{1}{4}$. Thus, for $\theta \le \frac{1}{4}$, all the cars sell at the price in Equation 19.1.

Varying Quality with Asymmetric Information Most firms can adjust their product's quality. If consumers cannot identify high-quality goods before purchase, they pay the same for all goods regardless of quality. Because the price that firms receive for top-quality goods is the same as they receive for low-quality items, they do not produce top-quality goods. Such an outcome is inefficient if consumers are willing to pay sufficiently more for top-quality goods.

This unwillingness to produce high-quality products is due to an externality (Chapter 18): *A firm does not completely capture the benefits from raising the quality of its product.* By selling a better product than what other firms offer, a seller raises the average quality in the market, so buyers are willing to pay more for all products. As a result, the high-quality seller shares the benefits from its high-quality product with sellers of low-quality products by raising the average price to all. *The social value of raising the quality*, as reflected by the increased revenues shared by all firms, *is greater than the private value*, which is only the higher revenue received by the firm with the good product.

Solved Problem 19.2

MyLab Economics Solved Problem

It costs \$10 to produce a low-quality wallet and \$20 to produce a high-quality wallet. Consumers cannot distinguish between the products before purchase, they do not make repeat purchases, and they value the wallets at the cost of production. The five firms in the market produce 100 wallets each. Each firm produces only high-quality or only low-quality wallets. Consumers pay the expected value of a wallet. Do any of the firms produce high-quality wallets?

Answer

1. *Calculate the expected value of wallet.* If all five firms make a low-quality wallet, consumers pay $10 per wallet. If one firm makes a high-quality wallet and all the others make low-quality wallets, the probability that a consumer buys a high-quality wallet is $\frac{1}{5}$. Thus, the expected value per wallet to consumers is

$$\left(\$10 \times \tfrac{4}{5}\right) + \left(\$20 \times \tfrac{1}{5}\right) = \$12.$$

2. *Show that it does not pay for a single firm to make high-quality wallets if the other firms make low-quality wallets due to asymmetric information.* If one firm raises the quality of its product, all firms benefit because the wallets sell for $12 instead of $10. The high-quality firm receives only a fraction of the total benefit from raising quality but bears the entire cost. It gets $2 extra per high-quality wallet sold, which is less than the extra $10 it costs to make the better wallet. (In contrast, the other firms benefit from the higher price without incurring the extra cost.) Therefore, due to asymmetric information, no firm produces high-quality goods even though consumers are willing to pay for the extra quality.

19.2 Reducing Adverse Selection

Because adverse selection results from one party exploiting asymmetric information about a hidden characteristic, the two main methods for solving adverse selection problems are to *equalize information* among the parties and to *restrict the ability of the informed party to take advantage of hidden information.* Responses to adverse selection problems increase welfare in some markets, but may do more harm than good in others.

Equalizing Information

Providing information to all parties may reduce or eliminate adverse selection problems. Either informed or uninformed parties can eliminate information asymmetries. Three methods for reducing informational asymmetries are:

screening
an action taken by an uninformed person to determine the information possessed by informed people

signaling
an action taken by an informed person to send information to a less-informed person

1. An uninformed party (such as an insurance company) can use **screening**: an action taken by an uninformed person to determine the information possessed by informed people.
2. An informed party (such as a person seeking to buy health insurance) can use **signaling**: an action taken by an informed person to send information to a less-informed person.
3. A *third party* (such as a firm or a government agency) not directly involved in the transaction may collect information and sell it or give it to the uninformed party.

Screening Uninformed people may eliminate their disadvantage by screening to gather information on the hidden characteristics of informed people. Life insurance companies reduce adverse selection problems by requiring medical exams. Based on this information, a firm may decide not to insure high-risk individuals or to charge them a higher premium as compensation for the extra risk.

It is costly to collect information on a person's health or dangerous habits such as smoking, drinking, or skydiving. As a result, insurance companies collect information

Congratulations!
You qualify for our auto insurance.

only up to the point at which the marginal benefit from the extra information they gather equals the marginal cost of obtaining it. Over time, insurance companies have increasingly concluded that it pays to collect information about whether individuals exercise, have a family history of dying young, or engage in potentially life-threatening activities.

Consumers can use screening techniques, too. For example, a potential customer can screen a used car by test-driving it; by having an objective, trustworthy mechanic examine the car; or by paying a company such as CARFAX to check the history of the repairs on the vehicle. As long as the consumers' costs of securing information are less than the private benefits, they obtain the information, transactions occur, and markets function smoothly.

In some markets, consumers can avoid the adverse selection problem by buying only from a firm that has a *reputation* for providing high-quality goods. For example, consumers know that a used car dealer that expects repeat purchases has a stronger incentive not to sell defective cars than does an individual.

Signaling An informed party may signal the uninformed party to eliminate adverse selection. However, signals solve the adverse selection problem only when the recipients view them as credible. Smart consumers may place little confidence in a firm's unsubstantiated claims. Do you believe that a used car runs well just because an ad tells you so?

If only high-quality firms find it worthwhile to send a signal, then a signal is credible. Producers of high-quality goods often try to signal to consumers that their products are of better quality than those of their rivals. If consumers believe their signals, these firms can charge higher prices for their goods.

But the signals must be credible to be effective. For example, a firm may distribute a favorable review of its product by an independent testing agency to try to convince buyers that its product is of high quality. Only if low-quality firms cannot obtain such a report from an independent testing agency, do consumers believe the signal.

A warranty may serve as both a signal and a guarantee. It is less expensive for the manufacturer of a reliable product to offer a warranty than it is for a firm that produces low-quality products. Consequently, if one firm offers a warranty and another does not, then a consumer may infer that the firm with the warranty produces a superior product. Of course, sleazy firms may try to imitate high-quality firms by offering a warranty that they do not intend to honor.

An applicant for life insurance can present an insurance company with a written statement from the doctor as a signal of good health. However, an insurance company may not trust such a signal if it is easy for people to find unscrupulous doctors who will falsely report that they are in good health. Here, screening by the insurance company using its own doctors may work better because the information is more credible.

Application

Discounts for Data

Are you healthy? Your insurance company wants to know—and it wants to keep you healthy. Life insurance companies in Australia, Europe, Singapore, and South Africa use the Internet to allow customers to signal that they're healthy.

When Andrew Thomas swipes his membership card when he arrives at his gym, his South African life insurance company, Vitality, receives instant information. The company checks whether he's still there 30 minutes later by tracking his location using his smart phone. In return for sharing medical and exercise information with his insurance company, Mr. Thomas earns points, which reduce his insurance premium by 9%.

In 2015, John Hancock became the first U.S. life insurance company to introduce a similar program. The company provides customers with Fitbit monitors that automatically upload their activity levels. The most active customers will earn a discount of up to 15% on their life-insurance premium, Amazon gift cards, and half-price stays at Hyatt hotels. As of 2016, Fitbit had partnership agreements with many health insurers, such as Anthem, Premera Blue Cross, Humana Vitality, and Optum/Unitedas.

Thus, life insurance companies are helping their customers signal that they're healthy, reducing the adverse selection problem. In addition, the companies are using this information to provide incentives for the people they insure to lead healthier lives so that their families have to wait longer to collect on their life insurance.

Third-Party Information In some markets, consumer groups, nonprofit organizations, and government agencies provide buyers with information about the quality of different goods and services. If this information is credible, it can reduce adverse selection by enabling consumers to avoid buying low-quality goods or paying less for poorer quality products.

For an outside organization to provide believable information, it must convince consumers that it is trustworthy. Consumers Union, which publishes the product evaluation guide *Consumer Reports*, tries to establish its trustworthiness by refusing to accept advertising or other payments from firms.

Auditing is another important example of third-party assessment, in which an independent accounting firm assesses the financial statements of a firm or other organization. Sometimes a firm obtains an audit voluntarily to enhance its reputation (a signal). Sometimes audits are required by a law or as a condition to participate in a particular transaction (screening).

Many local governments require that home sellers pay a professional to inspect the home and then to disclose all relevant facts about the home to potential buyers such as the age of the home and any known defects in the electrical work or plumbing. By doing so, these governments protect buyers against adverse selection due to undisclosed defects.

standard
a metric or scale for evaluating the quality of a particular product

certification
a report that a particular product meets or exceeds a given standard

Governments, consumer groups, industry groups, and others also provide information by establishing a **standard**: a metric or scale for evaluating the quality of a particular product. For example, the R-value of insulation—a standard—tells how effectively insulation works. Consumers learn of a brand's quality through **certification**: a report that a particular product meets or exceeds a given standard.

Many industry groups set their own standards and get an outside group or firm, such as Underwriters Laboratories (UL) or Factory Mutual Engineering Corporation (FMEC), to certify that their products meet specified standard levels. For example, by setting standards for the size of the thread on a screw, we ensure that screws work in products regardless of brand.

When standard and certification programs inexpensively and completely inform consumers about the relative quality of all goods in a market and do not restrict the goods available, the programs are socially desirable. However, some of these programs have harmful effects for two reasons.

First, standard and certification programs that provide imprecise information may mislead consumers. Some standards use only a high- versus low-quality rating even though quality varies continuously. Such standards encourage the manufacturing of products that have either the lowest possible quality (and cost of production) or the minimum quality level necessary to obtain the top rating.

Second, if standard and certification programs restrict salable goods and services to those that are certified, such programs may also have anticompetitive effects. For example, many governments license only professionals and craftspeople who meet some minimum standards. People without a license are not allowed to practice their profession or craft (See the Application "Occupational Licensing" in Chapter 2). The license drives up the price because the average quality of workers is higher and because the licensing law eliminates competition by lower-skilled workers. As a result, welfare may go up or down, depending on whether the increased-quality effect or the higher-price effect dominates. Whether government agencies set such restrictions properly and cost-effectively is widely debated.

Application

Adverse Selection and Remanufactured Goods

Because consumers can't see a good before buying it over the Internet, it's easy for a shady seller to misrepresent its quality. In the worst-case lemons-market scenario, low-quality goods drive out high-quality goods.

Adverse selection concerns are particularly strong for electronic goods sold on eBay, as consumers cannot screen by inspecting the product before purchase. That this market exists on eBay indicates that sellers have found ways to signal quality or consumers have found ways to screen.

Consumers know that selling on eBay creates an enforceable contract, so sellers' signals reduce adverse selection on eBay. Sellers have a variety of ways to signal. Some sellers offer money back guarantees or warranties. Some pay extra to eBay to post flashy displays. One important signal is whether the good is new, remanufactured, or used. For years, manufacturers of cameras, computers, mobile phones, MP3 players, and other consumer durables have refurbished or upgraded returned products before trying to sell them again. Even though these remanufactured products may be comparable to new ones, consumers do not perceive them that way, so they are willing to pay a premium for new ones.

Neto et al. (2016) studied sales on eBay of three types of iPods: the Classic, the Touch, and the Nano. They found, for example, that the average price of a used iPod Nano was 65% of a new one, while a remanufactured Nano was 82% of a new one. Thus, the signal of the good's type affects the price.

What about other signals such as positive descriptions? Consumers view used goods as varying more in quality than remanufactured or new products: the prices of used goods have greater variance than that of the others. As a result, quality claims may be more likely to affect the price of used goods than those of new and remanufactured goods. Neto et al. found that positive descriptions affect the price of most types of used iPods, but not new and remanufactured iPods.

Consumers also can screen using eBay's feedback (reputation) score: the percentage of positive ratings by past customers. A sleazy seller will have a bad feedback score. Subramanian and Subramanyam (2012) studied the sales of electronic goods on eBay. They found that a higher seller feedback score reduced the price differential between new and remanufactured goods. They also discovered that consumers pay higher prices for products remanufactured by the original manufacturer or their authorized factories than for those remanufactured by third parties.

Thus, third-party information in the form of ratings from previous customers, which consumers use to screen, and a variety of signals from firms help reduce the adverse selection problem for used and remanufactured goods.

Restricting Opportunistic Behavior

In addition to setting standards and certifying goods and services, governments and various organizations prevent opportunism in several ways. Three common examples are disclosure requirements, product liability laws, and universal coverage requirements.

Disclosure Requirements Governments often require that the informed party disclose all relevant information to the uninformed party. For example, many local governments require that sellers disclose all relevant facts about a house to potential buyers such as its age and any known defects in the electrical work or plumbing. By doing so, these governments protect buyers against adverse selection due to undisclosed defects.

Product Liability Laws In many countries, product liability laws protect consumers from being stuck with nonfunctional or dangerous products. Moreover, many U.S. state supreme courts have concluded that new products carry an implicit understanding that they will safely perform their intended functions. If they do not, consumers can sue the seller even in the absence of product liability laws. If consumers can rely on explicit or implicit product liability laws to force a manufacturer to compensate consumers for defective products, they do not need to worry about adverse selection. However, the transaction costs of going to court are very high.

Universal Coverage Health insurance markets have adverse selection because low-risk consumers do not buy insurance at prices that reflect the average risk. Some governments eliminate adverse selection by providing insurance to everyone or by mandating that everyone buy insurance. Canada, the United Kingdom, and many other countries provide basic health insurance to all residents, financed by a combination of mandatory premiums and taxes. The U.S. Patient Protection and Affordable Care Act ("Obamacare") has an "individual mandate," which requires virtually all Americans to obtain health care coverage.

Similarly, firms often provide mandatory health insurance as a benefit to all employees, rather than paying them a higher wage and allowing them to decide whether to buy such insurance on their own. By doing so, firms reduce adverse selection problems for their insurance carriers: Both healthy and unhealthy people are covered. As a result, firms can buy medical insurance for their workers at a lower cost per person than workers can obtain on their own.

19.3 Price Discrimination Due to False Beliefs About Quality

We've seen that bad products can drive out good products if consumers cannot distinguish lemons from good-quality products at the time of purchase. The market outcome also changes if consumers falsely believe that identical products differ in quality. Consumers pay more for a product that they believe is of higher quality.

If some consumers know that two products are identical while others believe that they differ in quality, a firm can profitably price discriminate. The firm takes advantage of the less-informed customers by charging them a high price for the allegedly superior product. The firm does not want to charge informed customers this same high price. Doing so would reduce profit because the resulting drop in sales would exceed the gain from the higher price.

Asymmetric information on the part of some, but not all, consumers makes price discrimination possible. However, if all customers are informed or all are uninformed about the quality of different products, firms charge a single price.

By intentionally increasing consumer uncertainty, a firm may be better able to exploit ignorant consumers and earn a higher profit (Salop, 1977). One way in which firms confuse consumers is to create *noise* by selling virtually the same product under various brand names. A *noisy monopoly* may be able to sell a product under its own brand name at a relatively high price and supply grocery or discount stores with a virtually identical product that sells at a lower price under a *private-label* (house or store) brand. For example, a single firm produces Prego spaghetti sauce and similar house brands for various grocery stores.

If some consumers know that two products are identical while others believe that their qualities differ, a firm can engage in a special type of price discrimination (Chapter 12). For example, a food manufacturer may take advantage of less-informed customers by charging a higher price for the allegedly superior national brand while informed customers buy a less expensive but equally good private-label brand.

Brand proliferation pays if the cost of producing multiple brands is relatively low and the share of consumers who are willing to buy the higher-price product is relatively large. Otherwise, the firm makes a higher profit by selling a single product at a moderate price than by selling one brand at a low price and another at a high price.

Application

Reducing Consumers' Information

By selling the same product under more than one brand name, firms can charge ignorant consumers higher prices. For decades, outside firms have manufactured products that Sears, Roebuck & Company sells under its house brand names—Kenmore, DieHard, and Craftsman. Whirlpool sells its own washers and driers, but Sears also markets these products under the Kenmore name. Sears also places its label on Amana, Caloric, GE, Gibson, Jenn-Air, Toshiba, and other products.[6]

Frequently, the Kenmore product is identical to or even superior to the brand-name product and costs less. Knowledgeable consumers realizing that the two brands are identical except for the label buy the Sears brand at the lower price. But customers who falsely believe that the name brand is better than the Kenmore product pay more for the name brand.

Over time, as consumers have become familiar with private-label brands and recognized their quality, private-label products have rapidly gained market share. According to the Private Label Manufacturers Association (PLMA) in 2016, the private-label share of units sold was 23% of supermarket goods and 17% of drug store goods. The PLMA also reports that the private-label market share was 52% in Switzerland, 50% in Spain, 46% in the United Kingdom, and 31% in Sweden. As consumers gain more knowledge about the quality of private-label brands, the advantage from maintaining multiple brands diminishes.

19.4 Market Power from Price Ignorance

We've just seen that consumer ignorance about quality can keep high-quality goods out of markets or lead to price discrimination. Consumer ignorance about how prices vary across firms has yet another effect: It gives firms market power. As a result, firms have an incentive to make it difficult for consumers to collect information about prices. For this reason, some stores won't quote prices over the phone, though the Internet has made hiding prices more difficult.

[6]Want to know which firm made the product you bought at Sears? Go to **http://www.applianceaid.com/searscodes.html**.

We now examine why asymmetric information about prices leads to noncompetitive pricing in a market that would otherwise be competitive. Suppose that many stores in a town sell the same good. If consumers have *full information* about prices, all stores charge the full-information, competitive price, p^*. If one store were to raise its price above p^*, the store would lose all its business. Each store faces a residual demand curve that is horizontal at the going market price and has no market power.

In contrast, if consumers have *limited information* about the price that firms charge for a product, one store can charge more than others and not lose all its customers. Customers who do not know that the product is available for less elsewhere keep buying from the high-price store.[7] Thus, each store faces a downward-sloping residual demand curve and has some market power.

Tourist-Trap Model

We now show that, if the market has a single price, it is higher than p^*. You arrive in a small town near the site of the discovery of gold in California. Souvenir shops crowd the street. Wandering by one of these stores, you see that it sells the town's distinctive snowy: a plastic ball filled with water and imitation snow featuring a model of the Donner Party. You instantly decide that you must buy at least one of these tasteful mementos—perhaps more if the price is low enough. Your bus will leave very soon, so you can't check the price at each shop to find the lowest price. Moreover, determining which shop has the lowest price won't be useful to you in the future because you do not intend to return anytime soon.

Let's assume that you and the many other tourists have a guidebook that reports how many souvenir shops charge each possible price for the snowy, but the guidebook does not state the price at any particular shop.[8] You and the other tourists have identical demand functions.

It costs each tourist c in time and expenses to visit a shop to check the price or buy a snowy. Thus, if the price is p, the cost of buying a snowy at the first shop you visit is $p + c$. If you go to two souvenir shops before buying at the second shop, the cost of the snowy is $p + 2c$.

When Price Is Not Competitive Will all souvenir shops charge the same price? If so, what price will they charge? We start by considering whether each shop charges the full-information, competitive price, p^*.

The full-information, competitive price is the equilibrium price only if no firm has an incentive to charge a different price. No firm would charge less than p^*, which equals marginal cost, because it would lose money on each sale.

However, a firm could gain by charging a higher price than p^*, so p^* is *not* an equilibrium price. If all other shops charge p^*, a firm can profitably charge $p_1 = p^* +$, where , a small positive number, is the shop's price markup. Suppose that you walk into this shop and learn that it sells the snowy for p_1. You know from your guidebook that all other souvenir shops charge only p^*. You say to yourself,

---

[7]A grave example concerns the ripping-off of the dying and their relatives. A cremation arranged through a memorial society—which typically charges a nominal enrollment fee of $10 to $25—often costs half or less than the same service arranged through a mortuary. Consumers who know about memorial societies—which get competitive bids from mortuaries—can obtain a relatively low price.

[8]We make this assumption about the guidebook to keep the presentation as simple as possible. This assumption is not necessary to obtain the following result.

"How unfortunate [or other words to that effect], I've wandered into the only expensive shop in town." Annoyed, you consider going elsewhere. Nonetheless, you do not go to another shop if this shop's markup, $\quad = p_1 - p^*$, is less than c, the cost of going to another shop.

As a result, it pays for this shop to raise its price by an amount that is just slightly less than the cost of an additional search, deviating from the proposed equilibrium where all shops charge p^*. Thus, *if consumers have limited information about price, an equilibrium in which all firms charge the full-information, competitive price is impossible.*

Monopoly Price We've seen that the market price cannot be lower than or equal to the full-information, competitive price. Is an equilibrium in which all stores charge the same price and that price is higher than the competitive price possible? In particular, can we have an equilibrium when all shops charge $p_1 = p^* + \quad$? No, shops would deviate from this proposed equilibrium for the same reason that they deviated from charging the competitive price. A shop can profitably raise its price to $p_2 = p_1 + \quad = p^* + 2\quad$. Again, it does not pay for a tourist who is unlucky enough to enter that shop to go to another shop as long as $\quad < c$. Thus, p_1 is not the equilibrium price. By repeating this reasoning, we can reject other possible equilibrium prices that are above p^* and less than the monopoly price, p_m.

However, the monopoly price may be an equilibrium price. No firm wants to raise its price above the monopoly level because its profit would fall due to reduced sales. When tourists learn the price at a particular souvenir shop, they decide how many snowies to buy. If the price is set too high, the shop's lost sales more than offset the higher price, so its profit falls. Thus, although the shop can charge a higher price without losing all its sales, it chooses not to do so.

The only remaining question is whether a shop would like to charge a lower price than p_m if all other shops charge that price. If not, p_m is an equilibrium price.

Should a shop reduce its price below p_m by less than c? If it does so, consumers don't benefit from searching for this low-price firm. The shop makes less on each sale, so its profits must fall. Thus, a shop should not deviate by charging a price that is only slightly less than p_m.

Does a shop gain by dropping its price below p_m by more than c? If the market has few stores, consumers may search for this low-price shop. Although the shop makes less per sale than the high-price shops, its profits may be higher because of greater sales volume. However, in a market with many stores, consumers do not search for the low-price shop because their chances of finding it are low. As a result, when the presence of a large number of shops makes searching for a low-price shop impractical, no firm lowers its price, so p_m is the equilibrium price. Thus, *when consumers have asymmetric information and when search costs and the number of firms are large, the only possible single-price equilibrium is at the monopoly price.*

If a firm charging a low price will break the single-price equilibrium at p_m, no single-price equilibrium is possible. Either no equilibrium exists or the equilibrium has prices that vary across shops (see Stiglitz, 1979, or Carlton and Perloff, 2005). Multiple-price equilibria are common.

Solved Problem 19.3

Initially, the market has many souvenir shops, each of which charges p_m (because consumers do not know the shops' prices), and buyers' search costs are c. If the government pays for half of consumers' search costs, is a single-price equilibrium at a price less than p_m possible?

Answer

Show that the argument we used to reject a single-price equilibrium at any price except the monopoly price did not depend on the size of the search cost. If all other stores charge any single price p, where $p^* \le p < p_m$, a firm profits from raising its price. As long as it raises its price by no more than $c/2$ (the new search cost to a consumer), unlucky consumers who stop at this deviant store do not search further. This profitable deviation shows that the proposed single-price equilibrium is not an equilibrium. Again, the only possible single-price equilibrium is at p_m.[9]

Advertising and Prices

The U.S. Federal Trade Commission (FTC), a consumer protection agency, opposes groups that want to forbid price advertising; the FTC argues that advertising about price benefits consumers. If a firm informs consumers about its unusually low price, it may be able to gain enough extra customers that its profit rises. If low-price stores advertise their prices and attract many customers, they can break the monopoly-price equilibrium that occurs when consumers must search store by store for low prices. The more successful the advertising, the larger these stores grow and the lower the average price in the market. If enough consumers become informed, all stores may charge the low price. Thus, without advertising, no store may find it profitable to charge low prices, but with advertising, all stores may charge low prices. See MyLab Economics, Chapter 19, "Advertising Lowers Prices."

19.5 Problems Arising from Ignorance When Hiring

Asymmetric information is frequently a problem in labor markets. Prospective employees may have less information about working conditions than firms do—a question we raised in the Challenge at the beginning of the chapter and address at the end. Conversely, firms may have less information about potential employees' abilities than the workers do.

Information asymmetries in labor markets lower welfare below the full-information level. Workers may signal and firms may screen to reduce the asymmetry in information about workers' abilities. In this section, we consider situations in which workers have more information about their ability than firms do. We look first at inexpensive signals sent by workers, then at expensive signals sent by workers, and finally at screening by firms.

Cheap Talk

Honesty is the best policy—when there is money in it. —Mark Twain

cheap talk unsubstantiated claims or statements

When an informed person voluntarily provides information to an uninformed person, the informed person engages in **cheap talk**: unsubstantiated claims or statements (see Farrell and Rabin, 1996). People use cheap talk to distinguish themselves or their

[9]If the search cost is low enough, a firm may break the single-price equilibrium at p_m by profitably charging a low price. If that happens, only a multiple-price equilibrium is possible. If the search cost falls to zero, consumers have full information, so the only possible equilibrium is at the full-information, competitive price.

attributes at low cost. Even though informed people may lie when it suits them, it is often in their and everyone else's best interest for them to tell the truth. Nothing stops me from advertising that I have a chimpanzee for sale, but doing so serves no purpose if I actually want to sell my DVD player. One advantage of cheap talk, if it is effective, is that it is a less expensive method of signaling ability to a potential employer than paying to have that ability tested.

Suppose that a firm plans to hire Cyndi to work at either a demanding job or an undemanding job. The demanding job requires someone with high ability. Someone with low ability can better perform the undemanding job because the job bores more able people, so they perform poorly.

Cyndi knows whether her ability level is high or low, but the firm is unsure. It initially thinks that either level is equally likely. Panel a of Table 19.1 shows the payoffs to Cyndi and the firm under various possibilities.[10] If Cyndi has high ability, she enjoys the demanding job: Her payoff is 3. If she has low ability, she finds the demanding job too stressful—her payoff is only 1—but she can handle the undemanding job. The payoff to the firm is greater if Cyndi is properly matched to the job: She is given the demanding job if she has high ability and the undemanding job if she has low ability.

We can view this example as a two-stage game. In the first stage, Cyndi tells the firm something. In the second stage, the firm decides which job she gets.

Cyndi could make many possible statements about her ability. For simplicity, we assume that she says either "My ability is high" or "My ability is low." This two-stage game has an equilibrium in which Cyndi tells the truth, and the firm, believing her, assigns her to the appropriate job. If she claims to have high ability, the firm gives her the demanding job.

Table 19.1 Employee–Employer Payoffs MyLab Economics Video

(a) When Cheap Talk Works

		Job That the Firm Gives to Cyndi	
		Demanding	*Undemanding*
Cyndi's Ability	*High*	3, 2	1, 1
	Low	1, 1	2, 4

(b) When Cheap Talk Fails

		Job That the Firm Gives to Cyndi	
		Demanding	*Undemanding*
Cyndi's Ability	*High*	3, 2	1, 1
	Low	3, 1	2, 4

[10]In Chapter 14, we used a 2×2 matrix to show a simultaneous-move game in which both parties choose an action at the same time. Here only the firm can make a move. In Table 19.1, Cyndi does not take an action, because she cannot choose her ability level.

If the firm reacts to her cheap talk in this manner, Cyndi has no incentive to lie. If she did lie, the firm would make a mistake, and a mistake would be bad for both parties. Cyndi and the firm want the same outcomes, so cheap talk works.

However, in many other situations, cheap talk does not work. Given the payoffs in panel b, Cyndi and the firm do not want the same outcomes. The firm still wants Cyndi in the demanding job if she has high ability and in the undemanding job otherwise. But Cyndi wants the demanding job regardless of her ability. So she claims to have high ability regardless of the truth. Knowing her incentives, the firm views her statement as meaningless babbling—her statement does not change the firm's view that her ability is equally likely to be high or low.

Given that belief, the firm gives her the undemanding job, for which its expected payoff is higher. The firm's expected payoff is $\left(\frac{1}{2} \times 1\right) + \left(\frac{1}{2} \times 4\right) = 2.5$ if it gives her the undemanding job and $\left(\frac{1}{2} \times 2\right) + \left(\frac{1}{2} \times 1\right) = 1.5$ if it assigns her to the demanding job. Thus, given the firm's asymmetric information, the outcome is inefficient if Cyndi has high ability.

When the interests of the firm and the individual diverge, cheap talk does not provide a credible signal. Here an individual has to send a more expensive signal to be believed. We now examine such a signal.

Application

Cheap Talk in eBay's Best Offer Market

In addition to auctions, eBay allows a seller to offer to sell a good for a specified price. The buyer may allow a potential buyer to respond with a *best offer* of a lower price. The seller may accept the best-offer bid, decline it, or make a counteroffer. The transaction is completed when a buyer or the seller accept the other side's offer.[11]

Backus et al. (2015) suggested that some sellers use cheap talk by posting an initial price that is a multiple of \$100. Items listed in multiples of \$100 receive offers that are 5% to 8% lower, arrive 6 to 11 days sooner, and are 3% to 5% more likely to sell than are items listed at similar "precise" prices such as \$109. Thus, these round numbers may provide information that helps both parties: The seller makes a quick sale, and the customer buys at a low price.

Education as a Signal

No doubt you've been told that one good reason to go to college is to get a good job. Going to college may get you a better job because you obtain valuable training. Another possibility is that a college degree may land you a good job because it serves as a signal to employers about your ability. If high-ability people are more likely to go to college than low-ability people, schooling signals ability to employers (Spence, 1974).

To illustrate how such signaling works, we'll make the extreme assumptions that graduating from an appropriate school serves as the signal and that schooling provides no training that is useful to firms (Stiglitz, 1975). High-ability workers are θ share of the workforce, and low-ability workers are $1 - \theta$ share. The output that a high-ability worker produces for a firm is worth w_h, and that of a low-ability worker is w_l (over their careers). If competitive employers knew workers' ability levels, they would pay this value of the marginal product to each worker, so a high-ability worker receives w_h and a low-ability worker earns w_l.

[11]Because this game (Chapter 13) is very complicated, participants have difficulty devising optimal strategies.

pooling equilibrium
an equilibrium in which dissimilar people are treated (paid) alike or behave alike

separating equilibrium
an equilibrium in which one type of people takes actions (such as sending a signal) that allows them to be differentiated from other types of people

We assume that employers cannot directly determine a worker's skill level. For example, when production is a group effort—such as in an assembly line—a firm cannot determine the productivity of a single employee.

Two types of equilibria are possible, depending on whether or not employers can distinguish high-ability workers from others. If employers have no way of telling workers apart, the outcome is a **pooling equilibrium**: Dissimilar people are treated (paid) alike or behave alike. Employers pay all workers the average wage:

$$\overline{w} = \theta w_h + (1 - \theta)w_l. \tag{19.2}$$

Risk-neutral, competitive firms expect to break even because they underpay high-ability people by enough to offset the losses from overpaying low-ability workers.

We assume that high-ability individuals can get a degree by spending c to attend a school and that low-ability people cannot graduate from the school (or that the cost of doing so is prohibitively high). If high-ability people graduate and low-ability people do not, a degree is a signal of ability to employers. Given such a clear signal, the outcome is a **separating equilibrium**: One type of people takes actions (such as sending a signal) that allow them to be differentiated from other types of people. Here a successful signal causes high-ability workers to receive w_h and the others to receive w_l, so wages vary with ability.

We now examine whether a pooling or a separating equilibrium is possible. We consider whether anyone would want to change behavior in an equilibrium. If no one wants to change, the equilibrium is feasible.

Separating Equilibrium In a separating equilibrium, high-ability people pay c to get a degree and are employed at a wage of w_h, while low-ability individuals do not get a degree and work for a wage of w_l. The low-ability people have no choice, as they can't get a degree. High-ability individuals have the option of not going to school. Without a degree, however, they are viewed as low ability once hired, and they receive w_l. If they go to school, their net earnings are $w_h - c$. Thus, it pays for a high-ability person to go to school if

$$w_h - c > w_l.$$

Rearranging terms in this expression, we find that a high-ability person chooses to get a degree if

$$w_h - w_l > c. \tag{19.3}$$

Equation 19.3 says that the benefit from graduating, the extra pay $w_h - w_l$, exceeds the cost of schooling, c. If Equation 19.3 holds, no worker wants to change behavior, so a separating equilibrium is feasible.

Suppose that $c = \$15{,}000$ and that high-ability workers are twice as productive as are others: $w_h = \$40{,}000$ and $w_l = \$20{,}000$. Here the benefit to a high-ability worker from graduating, $w_h - w_l = \$20{,}000$, exceeds the cost by \$5,000. Thus, no one wants to change behavior in this separating equilibrium.

Pooling Equilibrium In a pooling equilibrium, employers pay all workers the average wage from Equation 19.2, $\overline{w}$. Again, because low-ability people cannot graduate, they have no choice. A high-ability person must decide whether to go to school. Without a degree, that individual earns the average wage. With a degree, the worker is paid w_h. It does not pay for the high-ability person to graduate if the benefit from graduating, the extra pay $w_h - \overline{w}$, is less than the cost of schooling:

$$w_h - \overline{w} < c. \tag{19.4}$$

Thus, if Equation 19.4 holds, no worker wants to change behavior, so a pooling equilibrium persists.

For example, if $w_h = \$40{,}000$, $w_l = \$20{,}000$, and $\theta = \frac{1}{2}$, then

$$\overline{w} = \left(\tfrac{1}{2} \times \$40{,}000\right) + \left(\tfrac{1}{2} \times \$20{,}000\right) = \$30{,}000.$$

If the cost of going to school is $c = \$15{,}000$, the benefit to a high-ability person from graduating, $w_h - \overline{w} = \$10{,}000$, is less than the cost, so a high-ability individual does not want to go to school. As a result, a pooling equilibrium occurs.

Solved Problem 19.4

For what values of θ is a pooling equilibrium possible in general? In particular, if $c = \$15{,}000$, $w_h = \$40{,}000$, and $w_l = \$20{,}000$, for what values of θ is a pooling equilibrium possible?

Answer

1. *Determine the values of θ for which it pays for a high-ability person to go to school.* From Equation 19.4, we know that a high-ability individual does not go to school if $w_h - \overline{w} < c$. Using Equation 19.2, we substitute for $\overline{w}$ in Equation 19.4 and rearrange terms to find that high-ability people do not go to school if $w_h - [\theta w_h + (1 - \theta)w_l] < c$, or

$$\theta > 1 - \frac{c}{w_h - w_l}. \tag{19.5}$$

   If almost everyone has high ability, so θ is large, a high-ability person does not go to school. The intuition is that, as the share of high-ability workers, θ, gets large (close to 1), the average wage approaches w_h (Equation 19.2), so the benefit, $w_h - \overline{w}$, of going to school is small.

2. *Solve for the possible values of θ for the specific parameters.* If we substitute $c = \$15{,}000$, $w_h = \$40{,}000$, and $w_l = \$20{,}000$ into Equation 19.5, we find that high-ability people do not go to school—a pooling equilibrium is possible—if $\theta > \frac{1}{4}$.

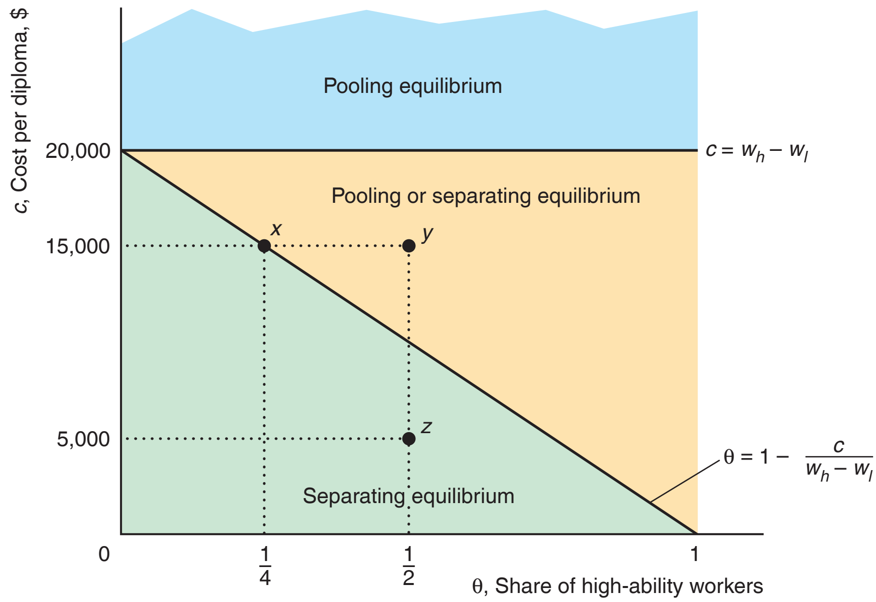
Unique or Multiple Equilibria Depending on differences in abilities, the cost of schooling, and the share of high-ability workers, only one type of equilibrium may be possible or both may be possible. In the following examples, using Figure 19.2, $w_h = \$40{,}000$ and $w_l = \$20{,}000$.

Only a pooling equilibrium is possible if schooling is very costly: $c > w_h - w_l = \$20{,}000$, so Equation 19.3 does not hold. A horizontal line in Figure 19.2 shows where $c = w_h - w_l = \$20{,}000$. Only a pooling equilibrium is feasible above that line, $c > \$20{,}000$, because it does not pay for high-ability workers to go to school.

Equation 19.5 demonstrates that, with few high-ability people (relative to the cost and earnings differential), only a separating equilibrium is possible. The figure shows a sloped line where $\theta = 1 - c/(w_h - w_l)$. Below that line, where $\theta < 1 - c/(w_h - w_l)$, relatively few people have high ability, so the average wage, $\overline{w}$, is low. A pooling equilibrium is not possible because high-ability workers would want to signal. Thus, below this line, only a separating equilibrium is possible. Above this line, Equation 19.5 holds, so a pooling equilibrium is possible. (The answer to Solved Problem 19.4

Figure 19.2 Pooling and Separating Equilibria MyLab Economics Video

If firms know workers' abilities, high-ability workers are paid w_h = \$40,000 and low-ability workers get w_l = \$20,000. The type of equilibrium depends on the cost of schooling, c, and the share of high-ability workers, θ. If $c >$ \$20,000, only a pooling equilibrium, in which everyone gets the average wage, is possible. If the market has relatively few high-ability people, $\theta < 1 - c/\$20{,}000$, only a separating equilibrium is possible. Between the horizontal and sloped lines, either type of equilibrium may occur.



shows that no one wants to change behavior in a pooling equilibrium if c = \$15,000 and $\theta > \frac{1}{4}$, which are points to the right of x in the figure, such as y.)

Below the horizontal line, where the cost of signaling is less than \$20,000 and above the sloped line, with relatively many high-ability workers, either equilibrium may occur. For example, where c = \$15,000 and $\theta = \frac{1}{2}$, Equations 19.3 and 19.4 (or equivalently, Equation 19.5) hold, so both a separating equilibrium and a pooling equilibrium are possible. In the pooling equilibrium, no one wants to change behavior, so this equilibrium is possible. Similarly, no one wants to change behavior in a separating equilibrium.

A government could ensure that one or the other of these equilibria occurs. It achieves a pooling equilibrium by banning schooling (and other possible signals). Alternatively, the government creates a separating equilibrium by subsidizing schooling for some high-ability people. Once some individuals start to signal, so that firms pay either a low or a high wage (not a pooling wage), it pays for other high-ability people to signal.

Efficiency In our example of a separating equilibrium, high-ability people get an otherwise useless education solely to show that they differ from low-ability people. An education is privately useful to the high-ability workers if it serves as a signal that gets them higher net pay. In our extreme example, education is socially inefficient because it is costly and provides no useful training.

Signaling changes the distribution of wages: Instead of everyone getting the average wage, high-ability workers receive more pay than low-ability workers.

Nonetheless, the total amount that firms pay is the same, so firms make zero expected profits in both equilibria.[12] Moreover, everyone works in both the pooling and the screening equilibrium, so total output is the same.

Nonetheless, everyone may be worse off in a separating equilibrium. At point y in Figure 19.2 ($w_h = \$40{,}000$, $w_l = \$20{,}000$, $c = \$15{,}000$, and $\theta = \frac{1}{2}$), either a pooling equilibrium or a separating equilibrium is possible. In the pooling equilibrium, each worker is paid $\overline{w} = \$30{,}000$ and no wasteful signaling occurs. In the separating equilibrium, high-ability workers make $w_h - c = \$25{,}000$ and low-ability workers make $w_l = \$20{,}000$.

Here high-ability people earn less in the separating equilibrium, \$25,000, than they would in a pooling equilibrium, \$30,000. Nonetheless, if anyone signals, all high-ability workers will want to send a signal to prevent their wage from falling to that of a low-ability worker. The reason socially undesirable signaling happens is that the private return to signaling—high-ability workers net an extra \$5,000 [$= (w_h - c) - w_l = \$25{,}000 - \$20{,}000$]—exceeds the net social return to signaling. The gross social return to the signal is zero—the signal changes only the distribution of wages—and the net social return is negative because the signal is costly.

This inefficient expenditure on education is due to asymmetric information and the desire of high-ability workers to signal their ability. Here the government can increase total social wealth by banning wasteful signaling (eliminate schooling). Both low-ability and high-ability people benefit from such a ban.

In other cases, however, high-ability people do not want a ban. At point z (where $\theta = \frac{1}{2}$ and $c = \$5{,}000$), only a separating equilibrium is possible without government intervention. In this equilibrium, high-ability workers earn $w_h - c = \$35{,}000$ and low-ability workers make $w_l = \$20{,}000$. If the government bans signaling, both types of workers earn \$30,000 in the resulting pooling equilibrium, so high-ability workers lose \$5,000 each. So even though the ban raises efficiency (wasteful signaling is eliminated), high-ability workers oppose the ban.

In this example, a ban on signaling increases efficiency because signaling is unproductive. However, some signaling is socially efficient because it increases total output. Education may raise output because its signal results in a better matching of workers and jobs or because it provides useful training as well as serving as a signal. Education also may make people better citizens. In conclusion, *total social output falls with signaling if signaling is socially unproductive but may rise with signaling if signaling also raises productivity or serves some other desirable purpose*.

Empirical evidence on the importance of signaling is mixed. Tyler et al. (2000) found that, for the least-skilled high school dropouts, passing the General Educational Development (GED) equivalency credential (the equivalent of a high school diploma) increases white dropouts' earnings by 10% to 19% but has no statistically significant effect on minority dropouts.

Screening in Hiring

Firms screen prospective workers in many ways. An employer may base hiring on an individual's characteristic that the employer believes is correlated with ability, such as how a person dresses or speaks, or a firm may use a test. Further, some employers engage in *statistical discrimination*, believing that an individual's gender, race, religion, or ethnicity is a proxy for ability.

[12]Firms pay high-ability workers more than low-ability workers in a separating equilibrium, but the average amount they pay per worker is $\overline{w}$, the same as in a pooling equilibrium.

Most societies accept the use of interviews and tests by potential employers. Firms commonly use interviews and tests as screening devices to assess abilities. If such screening devices are accurate, the firm benefits by selecting superior workers and assigning them to appropriate tasks. However, as with signaling, these costly activities are inefficient if they do not increase output. In the United States, courts may reject the use of hiring tests if the employer cannot demonstrate that the tests accurately measure skills or abilities required on the job.

If employers think that people of a certain gender, race, religion, or ethnicity have higher ability on average than others, they may engage in *statistical discrimination* (Aigner and Cain, 1977) and hire only such people. Employers may engage in this practice even if they know that the correlation between these factors and ability is imperfect.

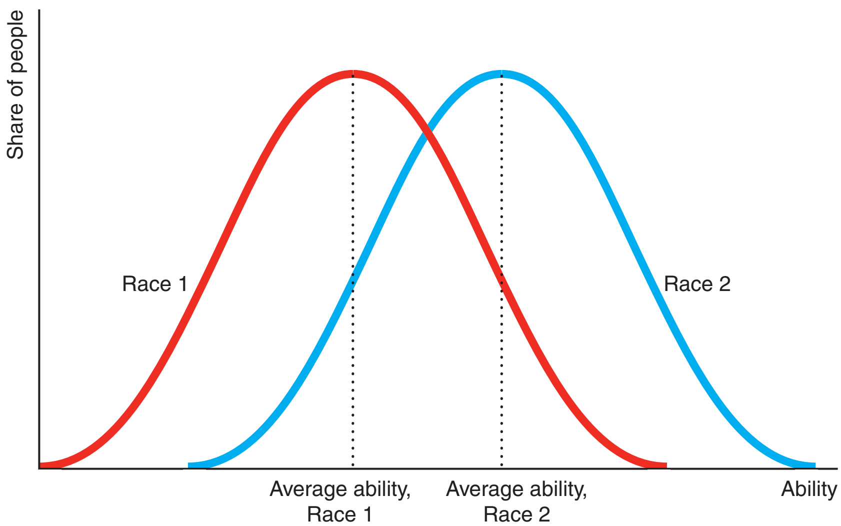
Figure 19.3 illustrates one employer's belief that members of Race 1 have, on average, lower ability than members of Race 2. The figure shows that the employer believes that some members of Race 1 have higher ability than do some members of the second race: Part of the Race 1 curve lies to the right of part of the Race 2 curve. Still, because the employer believes that a group characteristic, race, is an (imperfect) indicator of individual ability, the employer hires only people of Race 2 if enough of them are available.

The employer may claim not to be prejudiced but to be concerned only with maximizing profit.[13] Nonetheless, this employer's actions harm members of Race 1 as much as they would if they were due to racial hatred.

It may be very difficult to eliminate statistical discrimination even though ability distributions are identical across races. If all employers share the belief that members of Race 1 have such low ability that it is not worth hiring them, people of that race are never hired, so employers never learn that their beliefs are incorrect. Thus, false beliefs can persist indefinitely. Such discrimination lowers social output if it keeps skilled members of Race 1 from performing certain jobs.

Figure 19.3 Statistical Discrimination

This figure shows the beliefs of an employer who thinks that people of Race 1 have less ability on average than people of Race 2. This employer hires only people of Race 2, even though the employer believes that some members of Race 1 have greater ability than do some members of Race 2. Because this employer never employs members of Race 1, the employer may never learn that workers of both races have equal ability.



[13]Not all employment discrimination is due to statistical discrimination. Other common sources of discrimination are prejudice (Becker, 1971) and the exercise of monopsony power (Madden, 1973).

However, in some situations, statistical discrimination reflects true differences between groups. For example, insurance companies offer lower auto insurance rates to young women than to young men because young men are more likely, *on average*, to have an accident. The companies report that this practice lowers their costs of providing insurance by reducing moral hazard. Nonetheless, this practice penalizes young men who are unusually safe drivers and benefits young women who are unusually reckless drivers.

Challenge Solution

Dying to Work

In the Challenge at the beginning of the chapter, we asked two questions: Does a firm underinvest in safety if it knows how dangerous a job is but potential employees do not? Can the government intervene to improve this situation?

Consider an industry with two firms that are simultaneously deciding whether to make costly safety investments such as sprinkler systems in a plant or escape tunnels in a mine. Unlike the firms, potential employees do not know how safe it is to work at each firm. They know only how risky it is to work in this industry. If only Firm 1 invests, workers in the industry do not know that safety has improved at only Firm 1's plant. Because the government's accident statistics for the industry fall, workers realize that it is safer to work in the industry, so both firms pay lower wages. Thus, one firm's safety investment provides an externality to the other firm.

A prisoners' dilemma game (Chapter 14) illustrates this result. The profit table shows how the firms' profits depend on their safety investments. Firm 1 has a dominant strategy. If Firm 2 invests (compare profits in the cells in the right column), Firm 1's *no investment* strategy has a higher profit, 250, than its *investment* strategy, 225. Similarly, if Firm 2 does not invest (compare the cells in the left column), Firm 1's profit is higher if it doesn't invest, 200, than if it does. Thus, not investing is the dominant strategy, as is indicated by the horizontal red line through the dominated investing strategy. Because the game is symmetric, the same reasoning shows that not investing is the dominant strategy for Firm 2 as well.

		Firm 2	
		No Investment	*Investment*
Firm 1	*No Investment*	**200** (Firm 2) **200** (Firm 1)	100 (Firm 2) 250 (Firm 1)
	Investment	250 (Firm 2) 100 (Firm 1)	225 (Firm 2) 225 (Firm 1)

Because both firms have a dominant strategy of not investing, that combination of dominant strategies (the upper left cell) is the Nash equilibrium. Both firms receive an equilibrium profit of 200. If both firms invest in safety (the lower right cell), each earns 225, which is more than they earn at the Nash equilibrium. However, investment by both firms is not an equilibrium, because each firm can increase its profit from 225 to 250 by not investing if the other firm invests.

The firms underinvest in safety because each firm bears the full cost of its safety investments but derives only some of the benefits. In contrast, if workers knew how safe each firm is, only a firm that invests in safety would be able to pay a lower wage, which would change the profit table and increase the likelihood that the firms invest in safety.[14]

Thus, if the government or a union were to collect and provide workers with firm-specific safety information, the firms might invest. However, for the government or a union to provide this information, their cost of gathering the necessary information has to be relatively low.

Summary

Asymmetric information causes market failures when informed parties engage in opportunistic behavior at the expense of uninformed parties. Two types of problems arise from opportunism. Adverse selection occurs when someone with a hidden characteristic exploits this information to the detriment of someone who cannot observe this characteristic. Moral hazard occurs when an informed party takes advantage of a less-informed party through a hidden action.

1. **Adverse Selection.** Adverse selection creates problems in insurance markets because people with low risk do not buy insurance, which drives up the price for high-risk people. Due to adverse selection, not all desirable transactions take place. As a result, low-quality items tend to be overrepresented in transactions, as with the lemons problem associated with used cars and many other products. Bad products may drive good products out of the market.

2. **Reducing Adverse Selection.** Methods for dealing with adverse selection problems include signaling by firms (including establishing brand names and providing guarantees or warranties), consumer screening (such as by using experts or relying on firms' reputations), the provision of information by third parties such as government agencies or consumer groups, and laws that limit the ability of informed parties to exploit their private information.

3. **Price Discrimination Due to False Beliefs About Quality.** Firms may price discriminate if some consumers incorrectly think that quality varies across identical products. Because only some consumers collect information about quality, only those consumers know whether the quality differs between products in some markets. Firms can exploit ignorant consumers by creating noise: selling the same good under two different brand names at different prices.

4. **Market Power from Price Ignorance.** If consumers do not know how prices vary across firms, a firm can raise its price without losing all its customers. Consequently, consumers' ignorance about price creates market power. In a market that would be competitive with full information, consumer ignorance about price may lead to a monopoly price or a distribution of prices.

5. **Problems Arising from Ignorance When Hiring.** Companies use signaling and screening to try to eliminate information asymmetries in hiring. Where prospective employees and firms share common interests—such as assigning the right worker to the right task—everyone benefits from eliminating the information asymmetry by having informed job candidates honestly tell the firms—through *cheap talk*—about their abilities. When the two parties do not share common interests, cheap talk does not work. Potential employees may inform employers about their abilities by using expensive signals, such as a college degree. If these signals are unproductive (as when education serves only as a signal and provides no training), they may be privately beneficial but socially harmful. If the signals are productive (as when education provides training or leads to greater output due to more fitting job assignments), they may be both privately and socially beneficial. Firms may also screen. Job interviews, objective tests, and other screening devices that lead to a better matching of workers and jobs may be socially beneficial. However, screening by statistical discrimination is harmful to discriminated-against groups. Employers who discriminate based on their belief about a group's characteristics may never learn that their discrimination is based on a false belief because they never test the belief.

[14] Because this information is a public good (Chapter 18), others may obtain this information if the firm provides it to employees. The cost to the firm of having others, such as government regulators (who fine firms for accidents and transgressions), obtain this information may exceed the lower-wage benefit from providing it to workers.

Questions

Select questions are available on MyLab Economics*; * = answer appears at the back of this book;* **A** *= algebra problem.*

1. Adverse Selection

1.1 If a retailer pays a salesperson a set hourly wage and that salesperson puts forth less effort and takes longer breaks than if compensation had been tied to performance, is this an example of adverse selection or moral hazard?

1.2 Suppose that half the population is healthy and the other half is unhealthy. If an insured healthy person gets sick, the full cost to the insurance company is \$1,000. If an insured unhealthy person gets sick, the cost to the insurance company is \$10,000. In a given year, any one person (healthy or unhealthy) has a 40% chance of getting sick. People know whether they are healthy but the insurance company does not. The insurance company offers complete, actuarially fair insurance at the same price to everyone. The insurance company covers all medical expenses of its policyholders, and its expected profit is zero.

a. If everyone purchases insurance, what is the price of the insurance?

b. If only unhealthy people purchase insurance, what is the price of the insurance?

c. If each person has the option of buying insurance, explain why adverse selection might be expected unless healthy people are highly risk averse. **A**

1.3 You want to determine whether or not buying stock in a certain company is a good investment. Can you use the following information to help answer this question? If so, how would it help you determine if the purchase has a lemons problem?

a. The company's annual report

b. The past performance of the company's stock

1.4 People expect the value of houses they buy to appreciate over time. New cars, on the other hand, depreciate over time. The depreciation rate is highest during the first year (about 20% on average). Part of first-year depreciation is normal; just like any capital asset, cars wear out over time. Part of the depreciation reflects market power; the market for new cars is an oligopoly whereas the market for used cars is more competitive. Part of the depreciation reflects obsolescence; a new model with new functions and features will come out within the year. But could the high first-year depreciation rate also indicate a lemons problem? Explain.

1.5 Use the lemons model to explain why shops that cater to tourists might be likely to sell low-quality merchandise.

1.6 What are the answers to Solved Problem 19.1 if customers are willing to pay \$10,000 for a good used car?

*1.7 In its 2017 report on the second-hand economy, Kijiji estimated the value of used goods sales in Canada to be about C\$29 billion (or 1.4% of GDP). Most exchanged goods fall into the category of used clothing, shoes, and accessories. Suppose that buyers of good-quality used dresses value them at C\$80 but value lemons at C\$30. The reservation price of owners of good-quality used dresses is C\$70 and of owners of lemons is C\$20. If the share of lemons among used dresses is δ and everyone is risk neutral, under what conditions do all dresses sell? When do only lemons sell? Under what conditions (if any) do no dresses sell? (*Hint*: See Solved Problem 19.1.) **A**

1.8 Suppose that potential used-car buyers value high-quality used cars at €8,000 and low-quality used cars at €3,000. Owners of high-quality used cars have a reservation price of €6,500, while the reservation price for owners of low-quality used cars is €2,000. Everyone is risk neutral. The share of current owners who have low-quality cars is θ.

a. For what values of θ do all the potential sellers sell their used cars?

b. If each potential buyer incurs a transaction cost to purchase a used car (the value of the time spent searching) of t = €500, for what values of θ do all the potential sellers sell their used cars?

c. If θ equals the maximum value calculated in part a, would all potential sellers sell their used cars when transaction costs equal €500 per person? Would any of the potential sellers sell? **A**

1.9 In Solved Problem 19.2, show that, if all the other firms were producing a high-quality wallet, it would pay a firm to start producing a low-quality wallet. **A**

1.10 In Solved Problem 19.2, would any of the firms produce high-quality wallets if the cost of producing a higher-quality wallet is only \$11? Explain. **A**

1.11 It costs \$12 to produce a low-quality electric stapler and \$16 to produce a high-quality stapler. Consumers cannot distinguish good staplers from poor staplers when they make their purchases. Four firms produce staplers. Consumers value staplers at their cost of production and are risk neutral. Will any of the four firms be able to produce high-quality staplers without making losses? What happens if consumers are willing to pay $36 for high-quality staplers? (*Hint*: See Solved Problem 19.2.) **A**

1.12 In the world of French high cuisine, a three-star rating from the *Michelin Red Guide* is a widely accepted indicator of gastronomic excellence. French consumers consider Gault Milleau, another restaurant guide, not as authoritative as the Michelin guide because Gault Milleau, unlike Michelin, accepts advertising and its critics accept free meals.

a. Why are guides' ratings important to restaurant owners and chefs? Discuss the effect of a restaurant's rating on the demand for the restaurant.

b. Why do advertising and free meals taint the credibility of Gault Milleau? Discuss the moral hazard problem of Gault Milleau's ratings.

c. If advertising and free meals taint the credibility of Gault Milleau, why does the guide accept advertising and free meals?

2. Reducing Adverse Selection

2.1 It is illegal in Australia to drive without insurance to cover bodily injuries to third parties. Other types of insurance, for example, to cover damage to vehicles due to collision, fire, or theft, are optional. A person may have to pay higher car insurance premiums in Australia if they are under 25, male, a new driver, have a job that requires driving, live in a high-risk area, or typically drive longer, and if they have past claims and traffic violations. What would be the efficiency and equity implications of prohibiting the use of any one (or more) of these factors in setting car insurance rates?

*2.2 Earthquake insurance is not mandatory in Canada, nor is coverage for earthquakes typically included in a standard home insurance policy. However, some insurance companies require it for certain areas of the country where the risk of an earthquake causing significant damage is particularly high. When added to a home policy, the annual premium will rise by between 30% and 50% depending on where a person lives, and the deductible will be higher than for other perils. It generally covers loss or damage to a policyholder's home and belongings caused by the actual shaking of the earth as well as living expenses while a home is being repaired. What would be the implications if an insurance company ignored any important factor in setting the premium it charges? In the case of earthquake insurance, suppose the omitted factor is the relative stability of the soil under a house. Would such an omission be a moral hazard?

*2.3 A firm spends a great deal of money in advertising to inform consumers of the brand name of its mushrooms. Should consumers conclude that its mushrooms are likely to be of higher quality than unbranded mushrooms? Why or why not?

2.4 Life insurance premiums are typically lower for non-smokers than for smokers. Life insurance companies require a clear urine test to confirm a person's status. If you smoke cigarettes only occasionally (for example, at parties once a month), you are likely to be treated as a smoker. If you are a very occasional pipe or cigar smoker, on the other hand, you could be treated as a non-smoker. If a person quits smoking long enough to qualify for the non-smoker rates but intends to begin smoking once the policy is in force, is this adverse selection or moral hazard? If the insurance company does not follow-up to confirm the subsequent smoking status of its policyholders, is this adverse selection or moral hazard?

2.5 How does John Hancock Life Insurance use modern communications technology to reduce adverse selection? (*Hint:* See the Application "Discounts for Data.")

2.6 Grocery stores, hotels, and other firms give their customers free *loyalty cards*. A customer who uses the card receives a discount. Are these firms signaling, screening, price discriminating, or engaging in other activities?

3. Price Discrimination Due to False Beliefs About Quality

3.1 Explain how a monopoly firm can price discriminate by advertising sales in newspapers or magazines that only some of its customers see. Is it a noisy monopoly?

3.2 The Application "Reducing Consumers' Information" notes that a food manufacturer may sell a national brand product for more than an identical private-label product. Is the firm a noisy monopoly (or oligopoly)?

3.3 Multi-branding is marketing of two or more similar and competing products under different brand names by the same firm. Unilever, the Dutch-British transnational consumer goods company, produces several worldwide brands, including detergents, deodorants, soaps, and toothpastes. One of their washing powders in the United Kingdom is sold as both Surf at £7 (45 washes) and Persil at £8 (45 washes). Give an asymmetric information explanation as to why a firm might use pairs of brand names and why one product might sell for more than the other does.

4. Market Power from Price Ignorance

*4.1 In Solved Problem 19.3, if the vast majority of all consumers know the true prices at all stores and only a few shoppers have to incur a search cost to learn the prices, would firms set a single-price equilibrium price at the monopoly level, p_m?

4.2 The Bar Council of India, a statutory body created by the Indian Parliament, prescribes standards of professional conduct for lawyers in the country. One of these standards is that law firms and lawyers are prohibited from soliciting work or advertising their services in any manner (although a 2017 amendment allows them to provide information about themselves on websites). Could this standard be considered a restraint on trade and, if so, what effect would it have on equilibrium prices?

5. Problems Arising from Ignorance When Hiring

5.1 Suppose that high-ability workers have a 40% share of the workforce ($\theta = 0.4$), the output that a high-ability worker produces for a firm is worth w_h = £450,000 over a career, the output of a low-ability worker is worth w_l = £300,000, and the present value of the cost of a four-year university education (which signals high ability) is c. For what values of c are both a pooling equilibrium and a separating equilibrium possible, and when do high-ability workers have higher net earnings in a separating equilibrium compared to a pooling equilibrium? (*Hint*: See Solved Problem 19.4.) **A**

5.2 Education is a continuous variable, where e_h is the years of schooling of a high-ability worker and e_l is the years of schooling of a lower-ability worker. The cost per period of education for these types of workers is c_h and c_l, respectively, where $c_l > c_h$. The wages they receive if employers can tell them apart are w_h and w_l. Under what conditions is a separating equilibrium possible? How much education will each type of worker get?

5.3 In Question 5.2, under what conditions is a pooling equilibrium possible? **A**

5.4 In Questions 5.2 and 5.3, describe the equilibrium if $c_l \le c_h$. **A**

5.5 Universities in India tend to use different grading systems. For example, in one, a grade of more than 90% for distinction may have its minimum passing mark at 60%. In another, a grade of more than 70% may signify distinction, while the minimum passing mark is 45%. This makes a direct comparison of percentage grades among universities in India difficult. Why might this policy help or hurt students on the job market?

5.6 In the ability signaling model, suppose that firms can pay c^* for a test that determines a worker's ability. Does a firm benefit by testing workers?

5.7 When is statistical discrimination privately inefficient? When is it socially inefficient? Does it always harm members of the discriminated-against group?

5.8 All citizens of the European Union are entitled to equal treatment in employment and training irrespective of racial or ethnic origin, religion or belief, sexual orientation, gender, disability, or age. A 2010 OECD study of low-income workers found that signals such as old age and immigrant status were less important as they were overrepresented among applicants for low-skill positions (Bonoli and Hinrichs, 2010). More important was the way in which an applicant first makes contact with a prospective employer. Specifically, unsolicited applications and recommendations from already employed workers were seen to be positive, while referrals from a public employment office were associated with lower motivation. How does this hiring behavior differ from statistical discrimination by employers? Discuss the equity and efficiency implications of this practice.

6. Challenge

6.1 Can you change the payoffs in the table in the Challenge Solution so that the firms choose to invest in safety? Explain. **A**

6.2 Two manufacturing firms are debating whether or not to increase their spending on regular equipment inspections and training to improve safety and reduce the chances of injury to workers. They have the following payoff matrix. Does either firm have a dominant strategy? What is the Nash equilibrium in this game? What is the minimum fine that the government could levy on the firms that would lead to a Nash equilibrium in which both firms invest in safety? **A**

		Firm 2	
		Do not invest	Invest
Firm 1	Do not invest	140 160	120 190
	Invest	210 130	180 170

Contracts and Moral Hazards

20

The contracts of at least 33 major league baseball players had incentive clauses providing a bonus if that player were named the Most Valuable Player in a Division Series. Unfortunately, no such award is given for a Division Series.[1]

Challenge

Clawing Back Bonuses

A major cause of the 2007–2009 worldwide financial crisis was that managers and other employees of banks, insurance companies, and other firms took excessive risks. Looking back on the events that led to the financial meltdown, Goldman Sachs Chief Executive Lloyd Blankfein admitted that Wall Street firms, caught up in the pursuit of profits, had ignored risks, and that these firms needed to dramatically change compensation practices. As he said, "Decisions on compensation and other actions taken and not taken, particularly at banks that rapidly lost a lot of shareholder value, look self-serving and greedy in hindsight."

Bank managers threw accepted lending practices out the window, rewarding their mortgage brokers for bringing in large numbers of new mortgages regardless of risk. For example, they granted many loans to risky borrowers without requiring any down payment. A borrower who does not make a down payment is more likely to default—stop paying the mortgage—than one who makes a sizeable down payment.

In the San Francisco Bay area, 69% of families whose owner-occupied homes were in foreclosure in the first nine months of 2007 had put down 0% at the time of purchase, and only 10% had made the traditional 20% down payment. In the United States, foreclosures rose from 100,000 a month in the summer of 2006 to 250,000 in 2007, 300,000 in 2008, and 350,000 in 2009 and 2010, before falling to 225,000 in 2011, 200,000 in 2012, and to 95,000 in the first quarter of 2016—the lowest rate since 2006.

We're now tying annual executive bonuses to performance. You owe us $100,000.

Brokers and their managers issued these risky mortgages because they were rewarded based on the number of mortgages or annual (short-run) profits. They didn't worry about lower future profits that would occur if economic conditions turned bad and borrowers defaulted on their mortgages. The result was terrible losses for banks, borrowers, and the overall economy during the Great Recession.

One response called for in the 2010 Dodd-Frank Consumer Protection Act was that firms institute *clawback* provisions, which allow firms to *claw back* or reclaim some earlier bonus payments to managers if their past actions resulted in later losses. Many firms instituted such provisions voluntarily. While only 18% of Fortune 100 companies reported having a clawback policy in 1986, almost 90% had such a provision by 2013. In 2016,

[1]Tom FitzGerald, "Top of the Sixth," *San Francisco Chronicle*, January 31, 1997, C6.

the U.S. Securities and Exchange Commission proposed a rule that would require banks and a broad range of financial institutions to have seven-year clawback provisions for senior executive officers and "significant-risk"-taking employees. An alternative, similar policy to a clawback policy is for a firm to withhold bonuses and other compensation for an extended period (often several years) so that managers receive rewards only for the long-run success arising from their decisions.

In Solved Problem 20.1, we address the question: Why did executives at these banks take extra risks that resulted in major lost shareholder value? In the Challenge Solution, we analyze the question: Does evaluating a manager's performance over a longer period using delayed compensation or clawback provisions benefit shareholders?

A firm's manager takes extreme risks. A dentist caps your tooth, not because you need it, but because the dentist wants a new flat-screen TV. An employee cruises the Internet for jokes instead of working when the boss is not watching. A driver of a rental car takes it off the highway, risking ruining the suspension.

Each of these examples illustrates an inefficient use of resources due to a *moral hazard*, where an informed person takes advantage of a less-informed person, often through a *hidden action* (Chapter 19). In this chapter, we examine how to design contracts that *eliminate inefficiencies* due to moral hazard problems *without shifting risk to people who hate bearing risk*—or contracts that at least reach a good compromise between these two goals.

Insurance companies introduced the term *moral hazard* into common usage. Many types of insurance are highly vulnerable to hidden actions by insured parties that result in moral hazard problems. For example, Ralph, the owner of a clothing store, purchased a large quantity of designer jeans and stored the jeans in a warehouse. Following standard practice, he insured the merchandise for its original purchase price against such hazards as fire or theft. Unfortunately for Ralph, these jeans are now unfashionable and not selling. Because he faces a significant financial loss, he burns down the warehouse and makes an insurance claim. The hidden action is that of setting the warehouse on fire. Because most people view such an action as unethical or *immoral* (as well as being illegal), we use the term *moral* hazard.

A less extreme example of moral hazard arises when medical insurance covers the expense of doctor visits. If insured people do not have to pay for visits to their doctors, some insured people make "excessive" visits to the doctor—more than if they had to pay for the visits themselves. For example, some people visit a doctor in part because they are hypochondriacs or are lonely and want some company. Such behavior is not against the law and may not strike you as unethical. However, because it is costly, insurance companies take actions to reduce the number of visits.

To illustrate methods of controlling moral hazards and the trade-off between moral hazards and risk, we focus on contracts between a *principal*—such as an employer—and an *agent*—such as an employee. The principal contracts with the agent to take an *action* that benefits the principal. Until now, we have assumed that firms can produce efficiently. However, if a principal cannot practically monitor an agent constantly, the agent may steal, shirk—a moral hazard in which agents do not provide all the services they are paid to provide—or engage in other opportunistic behavior that lowers productivity.

Opportunistic behavior by an informed agent harms a less-informed principal. Sometimes the losses are so great that both parties would benefit if each had full information and if opportunistic behavior were impossible.

In this chapter, we examine five main topics

1. **The Principal-Agent Problem.** How much the agent produces depends on how the principal compensates the agent and the principal's ability to monitor the agent's actions.
2. **Using Contracts to Reduce Moral Hazard.** A principal and an agent may agree to a contract that eliminates the moral hazard or strikes a balance between reducing the moral hazard and optimal risk sharing.
3. **Monitoring to Reduce Moral Hazard.** Employees work harder if an employer monitors their behavior and makes it worthwhile for them to keep from losing their jobs.
4. **Checks on Principals.** As a restraint against taking advantage of employees, an employer may agree to contractual commitments that make it in the employer's best interest to tell employees the truth.
5. **Contract Choice.** By observing which type of contract an agent picks when offered a choice, a principal may obtain enough information to reduce moral hazards.

20.1 The Principal-Agent Problem

We refer to moral hazard in a principal-agent relationship as a *principal-agent problem* or an *agency problem*. If information is symmetric so that actions are not hidden, principal-agent relationships do not give rise to a moral hazard problem. When a building contractor (the principal) subcontracts with a house painter (the agent) and both work on the same building site, the contractor can directly observe how hard and how well the painter is working. Due to this close monitoring, the painter cannot engage in any hidden action, such as taking an hour-long coffee break or running personal errands during work hours. Consequently, no inefficiency arises from this principal-agent relationship.

In contrast, when you contract with people whose actions you cannot observe or evaluate, they may take advantage of you. If you (the principal) pay an auto mechanic (the agent) by the hour to fix your car, you do not know whether that person worked all the hours billed. If you retain a lawyer to represent you in a suit arising from an accident, you do not know whether the settlement the lawyer recommends is in your best interest or the lawyer's.

To illustrate the principal-agent problem, we consider an example where the payoffs to the owner of a firm (the principal) and the manager (the agent) depend on the agent's actions and the *state of nature*, such as weather (which affects demand) or input prices (which affect costs). The principal and the agent care about how they share risk and the allocation of payoffs.

Paul, the principal, owns many ice cream parlors across North America. He contracts with Amy, the agent, to manage his Miami shop. Her duties include supervising workers, purchasing supplies, and performing other necessary actions.

The shop's daily earnings depend on the local demand conditions and on how hard Amy works. Demand for ice cream varies with the weather, and is high half the time, and low otherwise.

Amy puts forth either normal or extra effort. She views herself as an honest person, and she would never steal from Paul. She is always at the shop during regular business hours and puts in at least a normal amount of effort, even if Paul cannot check on her. She politely, but impersonally, asks everyone who enters the shop, "May I help you?"

Nonetheless, Amy might not be working as hard as possible. She could put forth extra effort by enthusiastically greeting regular customers by name, serving customers rapidly, spending extra hours checking with nearby businesses to see if they would be interested in joint promotions, and improving the appearance of the shop. However, extra work is tiring and prevents Amy from spending time at the beach with friends, reading novels, watching "Dancing with the Stars," and engaging in other activities that she enjoys. She values her personal cost of this extra effort at 40 per day.

For any given level of demand, the shop sells more ice cream if Amy puts forth extra effort. The shop also sells more for a given level of Amy's effort if demand is high. The profit of the ice cream shop before Amy is paid—the combined payoff to Paul and Amy—for the four possible combinations of effort and demand are:

		Demand	
		Low	*High*
Amy's Effort	*Normal*	100	300
	Extra	300	500

If the demand is high and Amy puts forth normal effort, or if the demand is low and she works hard, the firm's daily profit is 300. The profit is 500 if the demand is high and Amy works hard but is only 100 if the demand is low and she applies only normal effort.

Efficiency

efficient contract an agreement in which neither party can be made better off without harming the other party

efficiency in production the principal's and agent's combined value (profits, payoffs) is maximized

efficiency in risk bearing risk sharing is optimal in that the person who least minds facing risk—the risk-neutral or less risk-averse person—bears more of the risk

Ideally, the principal, Paul, and the agent, Amy, agree to an **efficient contract**: an agreement in which neither party can be made better off without harming the other party. Using an efficient contract results in *efficiency in production* and *efficiency in risk sharing*.

Efficiency in production requires that the principal's and agent's combined value (profits, payoffs) is maximized. In our example, moral hazard hurts the principal by more than it helps the agent, so achieving efficiency in production requires preventing the moral hazard.

Efficiency in risk bearing requires that risk sharing is optimal in that the person who least minds facing risk—the risk-neutral or less-risk-averse person—bears more of the risk. Risk-averse people are willing to pay a risk premium to avoid risk, whereas risk-neutral people do not care if they face fair risk or not (Chapter 17). In our example, Paul owns many ice cream parlors across North America. He can pool the returns from all these stores, so he is risk neutral. Amy, like most people, is risk averse. Thus, risk bearing is efficient if Paul bears all of the risk and Amy bears none of it.

Symmetric Information

MyLab Economics Video

Moral hazard is not a problem if Paul lives in Miami and can directly supervise Amy. They could agree to a contract that specifies Amy receives 200 per day if she works extra hard, but loses her job if she doesn't. Because Amy's cost of working extra hard is 40, she nets 160 $(= 200 - 40)$ if she works hard, which is better than being fired and getting nothing. Even though the shop's profit varies with demand, Amy bears no risk: She receives 200 regardless of demand conditions.

Paul is the *residual claimant*: He receives the *residual profit*, which is the amount left over from the store's profit after Amy's wage is paid. Because Amy works hard (to avoid losing her job), Paul's residual profit varies only with demand. If demand is low, the shop earns 300, he pays Amy 200, and retains 100. If demand is high, the shop earns 500, so Paul keeps 300 after paying Amy 200. Paul's expected profit is the probability of low demand, 50%, times 100 plus the probability of high demand, 50%, times 300, or

$$\left(\tfrac{1}{2} \times 100\right) + \left(\tfrac{1}{2} \times 300\right) = 200.$$

Under this contract, Paul bears all the risk from the shop's uncertain earnings. The variance of Paul's earnings is large relative to his expected profit: $\frac{1}{2}(100 - 200)^2 + \frac{1}{2}(300 - 200)^2 = 10{,}000$.[2]

The first row of Table 20.1—*perfect monitoring*—summarize this result. The last two columns show that this contract is efficient because Paul, the risk-neutral party, bears all the risk, and their combined earnings are as high as possible because Amy works extra hard.

Asymmetric Information

Paul grows tired of warm weather and moves from Miami to Toronto, Canada, where he can no longer observe Amy's effort. Because Amy's effort is now a hidden action to Paul, he faces a moral hazard problem.

Table 20.1 Ice Cream Shop Outcomes

	Expected Payoffs				Efficiency	
Contract	Paul	Amy[a]	Paul + Amy	Amy's Variance	Risk Bearing[b]	Joint Payoff[c]
Symmetric Information						
Perfect monitoring	200	160	360	0	Yes	Yes
Asymmetric Information						
Fixed wage of 100	100	100	200	0	Yes	No
Licensing fee of 200	200	160	360	10,000	No	Yes
State-contingent fee of 100 or 300	200	160	360	0	Yes	Yes
50% profit share	200	160	360	2,500	No	Yes
Wage and bonus of 200; Amy is risk neutral	200	160	360	10,000	Yes	Yes
Wage and bonus of 200; Amy is very risk averse	100	100	200	0	Yes	No

[a]If Amy puts in extra work, her payoff is her earnings minus 40, which is the value she places on having to work harder.

[b]Because Amy is risk averse and Paul is risk neutral, risk bearing is efficient only if Paul bears all the risk, so that Amy's variance is zero.

[c]Production is efficient if Amy puts in extra work, so that the shop's expected payoff is 400 rather than 200.

[2]The variance (Chapter 17) is the probability of low demand, 50%, times the square of the difference between the payoff under low demand, 100, and the expected payoff, 200, plus the probability of high demand, 50%, times the square of the difference between the payoff under high demand, 300, and the expected payoff.

When Paul could monitor Amy's effort, he could make her wage contingent on hard work. Now, he pays her a wage that does not vary with her (hidden) effort. Initially, we assume that Paul and Amy's contract specifies that Paul pays Amy a *fixed wage* of 100 regardless of how much profit the shop earns. Such a contract is a special case of a *fixed-fee contract*, in which one party pays the other a constant payment or fee.

Because Amy receives the same amount no matter how hard she works, she chooses not to work hard, which is a moral hazard problem. If Amy works normally, she incurs no additional personal cost from extra effort and receives 100. On the other hand, if she provides extra effort, she receives a wage of 100 but incurs a personal cost of 40, so her net return is only 60.

Because Amy provides normal effort, the shop earns 100 with low demand, which is just enough to pay Amy with nothing left over for Paul, and the shop earns 300 with high demand, so Paul nets 200. Thus, Paul faces an uncertain profit with an expected value of $\left(\frac{1}{2} \times 0\right) + \left(\frac{1}{2} \times 200\right) = 100$. The variance of Paul's earnings remains high: $\frac{1}{2}(0 - 100)^2 + \frac{1}{2}(200 - 100)^2 = 10{,}000$.

The second row of Table 20.1 (*Fixed wage of 100*) summarizes the effects of this fixed-wage contract. Paul bears all the risk, so risk bearing is again efficient. However, their expected combined earnings, 200, are less than in the previous example with symmetric information, 360. Amy now makes 100. Paul has an expected value of 100, which is all he cares about because he is risk neutral. Both were better off with symmetric information: Amy netted 160 and Paul expected to earn 200. Because the moral hazard substantially reduces the shop's expected earnings, paying Amy a fixed wage is not optimal. In the next section, we examine how a well-designed contract reduces inefficiency due to moral hazard.

Solved Problem 20.1

MyLab Economics Solved Problem

Traditionally the Las Vegas Home Bank made only *prime* loans—providing mortgages only to people who were very likely to repay the loans. However, Leonardo, a senior executive at the bank, is considering offering *subprime* loans—mortgages to speculators and other less creditworthy borrowers. If he makes only prime loans, the bank will earn \$160 million. If he also makes subprime loans, the bank will make a very high profit, \$800 million, if the economy is good so that few people default. However, if the economy is bad, the large number of defaults will cause the bank to lose \$320 million. The probability that the economy is bad is 75%.

Leonardo will receive 1% of the bank's profit if it is positive. He believes that if the bank loses money, he can walk away from his job without repercussions but with no compensation. Leonardo and the bank's shareholders are risk neutral. Does Leonardo provide subprime loans if all he cares about is maximizing his personal expected earnings? What would the bank's stockholders prefer that Leonardo do (given that they know the risks involved)?

Answer

1. *Compare the bank's expected return on the two types of mortgages.* If the bank makes both prime and subprime loans, its expected return is $(0.25 \times 800) + (0.75 \times [-320]) = -40$ million dollars, an expected loss. That is substantially less than the certain profit of \$160 million the bank makes if it provides only prime mortgages.

2. *Compare the manager's expected profits on the two investments.* Leonardo earns 1% of \$160 million, or \$1.6 million, if he provides only prime loans. If he makes prime and subprime loans, he earns 1% of \$800 million, or \$8 million, with a

probability of 25%, and gets no compensation with a probability of 75%. Thus, he expects to earn $(0.25 \times 8) + (0.75 \times 0) = 2$ million dollars. Because Leonardo is risk neutral and does not care about the shareholders' returns, he makes both types of loans.

3. *Compare the shareholders' expected profits on the two types of mortgages*. If the bank provides only prime mortgages, the bank's shareholders earn 99% of the profit from the prime mortgages, or $0.99 \times \$160$ million $= \$158.4$ million. If the bank makes both prime and subprime loans, shareholders earn 99% of the \$800 million, \$792 million, if the economy is good. But in a bad economy, the shareholders bear the full loss, \$320 million. The expected return to shareholders is $(0.25 \times 792) + (0.75 \times [-320]) = -42$ million dollars, an expected loss. Thus, the shareholders would prefer that the bank make only prime loans.

Comment: Given that Leonardo has the wrong incentives (and ignores his responsibility to shareholders), he takes a hidden action—choosing to provide subprime loans—that is not in the shareholders' best interest. One possible solution to the problem of managers and shareholders' diverging interests is to change the manager's compensation scheme, as we discuss in the Challenge Solution.

Application

Honest Cabbie?

You arrive in a strange city and get in a cab. Will the driver take you to your destination by the shortest route, or will you get ripped off?

To find out, Balafoutas et al. (forthcoming) ran an experiment in Athens. Four native-speaking Greeks took 400 taxi trips. For each trip, they said, "I would like to get to [name of a destination]. Do you know where it is? I am not from Athens." A few seconds after the ride began, the passenger said either, "Can I get a receipt at the end of the ride?" or "Can I get a receipt at the end of the ride? I need it to have my expenses reimbursed by my employer."

The experimenters expected that fraudulent behavior would be less likely in the former (control) case than in the latter ("moral hazard") case, where passengers would have weaker incentives to control or report a longer than necessary trip or overcharging. Overcharging (mostly bonus surcharges) occurred 36.5% of the time for the moral hazard rides compared to 19.5% for the control rides. Overall, the fare for the moral hazard trips averaged 17% more than for the control trips.

20.2 Using Contracts to Reduce Moral Hazard

A verbal contract isn't worth the paper it's written on.

Many people fail to recognize the importance of pay incentives.

Common Confusion: It doesn't matter whether someone is paid a lump-sum, by the hour, a percentage of the revenue, or in other ways.

To the contrary, the payment contract greatly affects the outcome if moral hazard is possible. A skillfully designed contract that provides the agent with a strong incentive to achieve production efficiency may reduce or eliminate moral hazard problems.

In this section, we illustrate how several types of contracts increase efficiency in the Paul and Amy ice cream shop example. These contracts provide greater incentives for Amy, the agent, to work hard, but some require her to bear some risk even though she is more risk averse than Paul, the principal. Thus, some of these contracts create a trade-off, increasing production efficiency but reducing risk-bearing efficiency.

Fixed-Fee Contracts

We initially considered a fixed-fee contract in which Paul (the principal) pays Amy (the agent) a fixed wage, with the result that Paul bears all of the risk and Amy bears none. Alternatively, Amy could pay Paul a fixed amount so that she receives the residual profit: the profit left over after she pays Paul his fixed return. Amy is effectively paying a *license fee* to operate Paul's ice cream shop. With such a contract, Paul bears no risk as he receives a fixed fee, while Amy bears all the risk.[3]

As Amy receives the residual profit under such a licensing contract, she receives all the increase in expected profit from her extra effort. She is motivated to work hard.

To illustrate why, we suppose that Amy pays Paul a fixed licensing fee of 200 per day and keeps any residual profit. (Our analysis depends *only* on Amy paying Paul a fixed fee and *not* on the exact amount that she pays.) If she does not work hard, she makes 100 (= 300 − 200) with high demand, but suffers a loss, −100 (= 100 − 200), if demand is low. Her expected gain if she does not work hard is $0 = \left(\frac{1}{2} \times [-100]\right) + \left(\frac{1}{2} \times 100\right)$. If she works hard, she nets 60 (= 300 − 200 − 40) with low demand and 260 (= 500 − 200 − 40) with high demand, so that her expected gain from working hard is $160 = \left(\frac{1}{2} \times 60\right) + \left(\frac{1}{2} \times 260\right)$.

Her variance in earnings is $10{,}000 = \frac{1}{2}(-100 - 0)^2 + \frac{1}{2}(100 - 0)^2$ with low demand, which is the same as the variance with high demand, $10{,}000 = \frac{1}{2}(60 - 160)^2 + \frac{1}{2}(260 - 160)^2$. Thus, because her risk is the same with both levels of effort but her expected net earnings are higher if she puts forth high effort, it is in her best interest to work hard.

Consequently, Amy's and Paul's total expected earnings are higher if Amy pays a fixed fee to Paul than if Paul pays a fixed fee to Amy, because Amy works harder if she is the residual claimant and reaps all the benefits of working harder. As the third row (*Licensing fee of 200*) of Table 20.1 shows, when Amy pays Paul a license fee, the shop's expected earnings are 400 because Amy works hard. Paul makes 200 with certainty, and Amy expects a net gain of 160 after deducting her cost, 40, of providing high effort. Therefore, the expected sum of their payoffs is 360. In contrast, if Paul pays Amy a fixed wage (second row, *Fixed wage of 100*), Amy earns 100 and Paul expects to earn 100 for an expected total payoff of 200.

Although Amy paying Paul rather than the other way around increases their total earnings, it makes the risk-averse person, Amy, bear all the risk, while Paul, the risk-neutral person, bears no risk. Therefore, although this contract maximizes combined expected earnings, it does not provide for efficient risk bearing.

Which contract is better depends on how risk averse Amy is. If Amy is nearly risk neutral, the fixed payment to Paul is superior, because both parties have higher expected earnings and Amy is not very concerned about the risk. However, if Amy is

[3]In some businesses, both types of fixed-fee contracts are used. For example, in some hair salons, hairdressers rent a chair from the owner for a fixed fee and bear all the risk associated with variations in demand, while other hairdressers are paid an hourly rate, with the owner getting the residual profit from their activities.

extremely risk averse, she may prefer receiving a fixed wage even if that means giving up significant expected earnings.[4]

Contingent Contracts

Many contracts specify that the parties receive payoffs that are contingent on some other variable, such as the action taken by the agent; the state of nature; or the firm's profit, output, or revenue. For example, when Paul can monitor Amy's effort, he offers her a contract that makes her payoff contingent on her effort. She is paid only if she provides extra effort and loses her job otherwise. Such a contract is efficient, but it is not feasible if Paul cannot monitor Amy's effort. Paul can use a contingent contract when he cannot monitor her.

State-Contingent Contracts In a *state-contingent contract*, one party's payoff is contingent on *only* the state of nature. For example, suppose Amy pays Paul a license fee of 100 if demand is low and a license fee of 300 if demand is high and keeps any additional earnings. As the residual claimant, Amy has an incentive to provide high effort. With low demand the shop earns 300, Amy pays Paul 100, and Amy's residual profit is $160 = 300 - 100 - 40$, where 40 is the cost of her extra effort. With high demand, the shop earns 500, Amy pays Paul 300, and Amy's residual profit is $160 = 500 - 300 - 40$.

Paul's expected payoff is $200 = \left(\frac{1}{2} \times 100\right) + \left(\frac{1}{2} \times 300\right)$, as the fourth row (*State-contingent fee of 100 or 300*) of Table 20.1 shows. Because Amy earns 160 in both states of nature, she bears no risk, while Paul bears all the risk. This result is efficient because Paul is risk neutral and Amy is risk averse. This state-contingent contract is fully efficient even if Paul cannot monitor Amy's effort. However, it does require that both parties observe and agree on the state of nature, which may not be possible.

Application

Health Insurance and Moral Hazard

By early 2016, about 20 million uninsured people had obtained health insurance coverage under the U.S. Patient Protection and Affordable Care Act (ACA). Society benefits by shifting risk from these previously uncovered, risk-averse people to risk-neutral insurance companies (Chapter 17). However, many analysts argue that extending insurance coverage results in more moral hazard. For example, patients may use the medical system excessively, driving up costs to everyone.

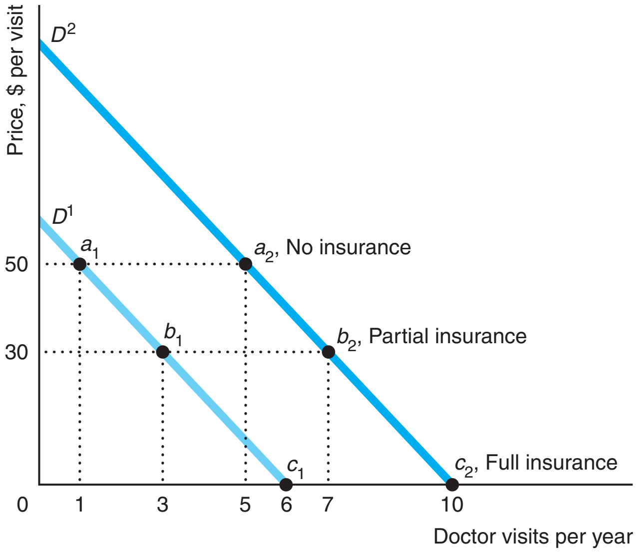
The adult dependent coverage provision of the ACA allows young adults up to age 26 to stay on their parents' health care policies. Prior to that, about one in three young adults ages 19–25 lacked insurance. Jhamb et al. (2015) estimated that this provision raised insurance coverage among young adults by 7.4% and the number of doctor visits by 3%. Do insured people excessive use health care services? One study, Kowalski (2015), examined an employer-sponsored health insurance for a large U.S. firm and estimated that the average deadweight losses from moral hazard substantially outweighed the average welfare gains from risk protection for this group.

[4]Amy might be more risk averse if, for example, she has no savings and would find it difficult to support herself during periods of low demand if she were the residual claimant.

Solved Problem 20.2

MyLab Economics Solved Problem

Gary's demand for doctor visits depends on his health. Half the time his health is good and his demand is D^1 in the figure. When his health is poor, his demand is D^2. Gary is risk averse. Without medical insurance, he pays \$50 a visit. With full insurance, he pays a fixed fee at the beginning of the year, and the insurance company pays the full cost of any visit. Alternatively, with a contingent contract, Gary pays a smaller premium at the beginning of the year, and the insurance company covers only \$20 per visit, with Gary paying the remaining \$30. How likely is a moral hazard problem to occur with each of these contracts? What is Gary's risk (the variance of his medical costs) with no insurance and with each of the two types of insurance? Compare the contracts in terms of the trade-offs between risk and moral hazard.



Answer

1. *Describe the moral hazard for each demand curve for each contract.* Given that Gary's health is good, if he does not have insurance, Gary pays the doctor \$50 a visit and goes to the doctor once, at point a_1 on D^1 in the figure. In contrast, with full insurance where he pays nothing per visit, he visits the doctor six times, at c_1. Similarly, if his health is poor, he goes to the doctor five times, a_2, without insurance, and 10 times, c_2, with full insurance. Thus, regardless of his health, he makes five extra visits a year with full insurance. These extra visits are the moral hazard.

   With a contingent contract, Gary pays \$30 a visit. He makes three visits if his health is good (at point b_1)—only two more than at a_1. If his health is poor, he makes seven visits, once again two more than if he were paying the full fee (five visits at point a_2). Thus, this contingent contract reduces the moral hazard problem: He makes only two extra visits.

2. *Calculate the variance of Gary's medical expenses for no insurance and for the two insurance contracts.* Without insurance, his average number of visits is $3 = \left(\frac{1}{2} \times 1\right) + \left(\frac{1}{2} \times 5\right)$, so his average annual medical cost is \$150. Thus, the variance of his medical expenses without insurance is

$$\begin{aligned}\sigma_n^2 &= \tfrac{1}{2}[(1 \times \$50) - \$150]^2 + \tfrac{1}{2}[(5 \times 50) - \$150]^2\\ &= \tfrac{1}{2}(\$50 - \$150)^2 + \tfrac{1}{2}(\$250 - \$150)^2\\ &= \$10{,}000.\end{aligned}$$

   If he has full insurance, he makes a single fixed payment each year, so his payments do not vary with his health: His variance is $\sigma_f^2 = 0$. Finally, with partial insurance, he averages 5 visits with an average cost of \$150, so his variance is

$$\sigma_p^2 = \tfrac{1}{2}(\$90 - \$150)^2 + \tfrac{1}{2}(\$210 - \$150)^2 = \$3{,}600.$$

   Thus, $\sigma_n^2 > \sigma_p^2 > \sigma_f^2$.

3. *Discuss the trade-offs*. Because Gary is risk averse, efficiency in risk bearing requires the insurance company to bear all the risk, as with full insurance. However, full insurance results in the largest moral hazard. Removing insurance eliminates the moral hazard, but forces Gary to bear all the risk. The contingent contract is a compromise in which both the moral hazard and the degree of risk lie between the extremes.

Profit-Sharing Contracts Even if the principal cannot observe the state of nature or the agent's actions, the principal may be able to design a contingent contract that reduces the moral hazard problem by making payments contingent on an outcome, such as profit or output. One common contingent contract is a *profit-sharing contract*, in which the payoff to each party is a fraction of the observable total profit.

Suppose that Paul and Amy agree to split the earnings of the ice cream shop equally. Does making Amy's pay contingent on the firm's earnings induce Amy to work hard?

If Amy works normally, the shop earns 100 if the demand is low and Amy receives half, or 50. If demand is high, the shop earns 300, so Amy's share is 150 $\left(= \frac{1}{2} \times 300\right)$. Thus, Amy's expected value from normal effort is $100 = \left(\frac{1}{2} \times 50\right) + \left(\frac{1}{2} \times 150\right)$. The variance of her earnings is $2{,}500 = \frac{1}{2}(50 - 100)^2 + \frac{1}{2}(150 - 100)^2$.

If Amy provides extra effort, the shop earns 300 if the demand is low, and Amy receives 150, but she incurs a personal cost of 40 for providing high effort, so her net return is 110. If the demand is high, the shop's profit is 500, so that Amy nets 210 $(= 250 - 40)$. Thus, her expected return from high effort is $160 = \left(\frac{1}{2} \times 110\right) + \left(\frac{1}{2} \times 210\right)$. The variance of her earnings is $2{,}500 = \frac{1}{2}(110 - 160)^2 + \frac{1}{2}(210 - 160)^2$, which is the same as with normal effort. Because extra effort provides Amy with higher expected earnings without increasing her risk, she provides extra effort.

Given that Amy works hard, Paul makes 150 $\left(= \frac{1}{2} \times 300\right)$ if demand is low and 250 $\left(= \frac{1}{2} \times 500\right)$ if demand is high. His expected profit is $200 = \left(\frac{1}{2} \times 150\right) + \left(\frac{1}{2} \times 250\right)$, as the profit-sharing row of Table 20.1 shows. Paul prefers this profit-sharing contract to a fixed-fee contract where he pays Amy a fixed wage of 100 and makes an expected profit of 100.

However, Amy chooses to work harder only if she gets a large enough share of the profit to offset her personal cost from doing the extra work. If Amy gets less than 20% of the profit, she chooses not to work hard and earns less than she would from the wage of 100.[5] Thus, profit sharing may reduce or eliminate the moral hazard problem, especially if the agent's share of the profit is large, but may not do so if the agent's share is small.

Solved Problem 20.3

Penny, the owner of a store, makes a deal with Arthur, her manager, that at the end of the year, she receives two-thirds of the store's profit and he gets one-third. If Arthur is interested in maximizing his earnings, will Arthur act in a manner that

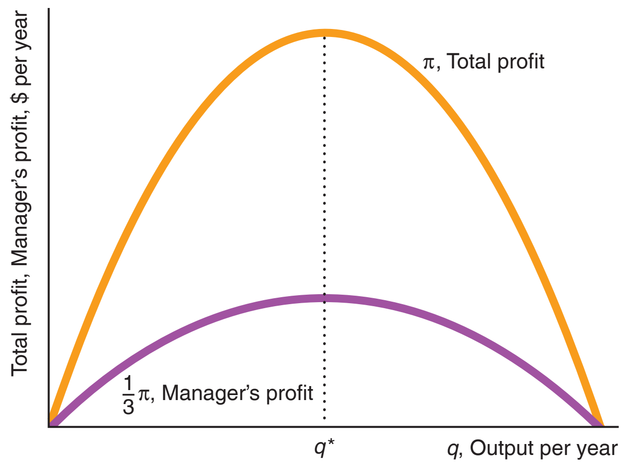
[5] If θ is Amy's share, then her expected earnings with normal effort is $\left(\frac{1}{2} \times 100\theta\right) + \left(\frac{1}{2} \times 300\theta\right) = 200\theta$, and her expected net earnings from extra effort is $\left(\frac{1}{2} \times 300\theta\right) + \left(\frac{1}{2} \times 500\theta\right) - 40 = 400\theta - 40$. She chooses not to put in extra effort if what she expects to earn from normal effort exceeds that from extra effort: $200\theta > 400\theta - 40$, or $\theta < 20\%$.

maximizes the store's total profit (which both Penny and Arthur can observe)? Answer using a graph.

Answer

1. *Draw a diagram showing the total profit curve and use it to derive Arthur's profit curve*. The figure shows that the total profit (π) curve is first increasing and then decreasing as output rises. At every output, Arthur's profit curve is one-third of the height of the profit curve, $\frac{1}{3}\pi$.

MyLab Economics Video



2. *Determine the quantity that maximizes Arthur's profit and check whether that quantity also maximizes total profit*. Because Arthur, the manager, gets a third of the total profit, $\frac{1}{3}\pi$, he sets output at q^*, which maximizes his share of the profit. Total profit and the owner's share of profit are also maximized at q^*. (Penny's share, $\frac{2}{3}\pi$—the vertical difference between the total profit curve and the agent's earnings at each output—is also maximized.)

Comment: At the output where total profit is maximized, any fraction of total profit is also maximized.[6]

Application

Sing for Your Supper

The producer of one of the world's largest annual music festivals, Outside Lands Music & Arts Festival, negotiates with dozens of food and drink vendors to sell goods at his annual event. According to his contract with them, they owe him the larger of a minimum amount (the "guarantee") and a percentage of their revenues.

[6]To determine where profit is maximized using calculus, we differentiate the profit function, $\pi(q)$, with respect to output, q, and set that derivative, which is marginal profit, equal to zero: $d\pi(q)/dq = 0$. The quantity that maximizes the manager's share of profit, $\frac{1}{3}\pi(q)$, is determined by $\frac{1}{3}d\pi(q)/dq = \frac{1}{3}d\pi(q)/dq = 0$, or $d\pi(q)/dq = 0$. That is, the condition that determines the quantity at which the manager's share of profit reaches a maximum is the same condition as the one that determines that total profit reaches a maximum.

He worried that the vendors might underreport their revenues, as he cannot easily monitor them. Using data from previous concerts, he compared reported revenues across vendors and found that the vast majority of vendors reported comparable revenues within their categories. The next year, he did not invite back the 10% of vendors who reported substantially smaller amounts. Thus, substantial cheating by vendors cost them the opportunity to participate in future events.

Then, he tried something new. Concertgoers could buy wine only by using an electronic payment system, which kept track of sales. He estimated that revenues increased by over 30% from the previous year due to more accurate reporting. Thus, by making the information symmetric, he eliminated the moral hazard problem.

Bonuses and Options To induce an agent to work hard, a principal may offer the agent a *bonus*: an extra payment if a performance target is hit. For example, Paul could offer Amy a base wage of 100 and a bonus of 200 if the shop's earnings (before paying Amy) exceed 300.

If Amy provides normal effort, the shop does not earn enough to trigger the bonus, so Amy receives 100 in both states of nature. If Amy provides extra effort but the demand is low, the shop earns 300, so Amy receives her wage of 100 and incurs a cost of 40, so her net benefit is 60. However, if she works hard and the demand is high, the shop earns 500, the bonus is triggered, and Amy gets her wage of 100 plus the bonus of 200. After subtracting her cost of extra effort, 40, she nets 260. Thus, Amy's expected return with extra effort is $160 = \left(\frac{1}{2} \times 60\right) + \left(\frac{1}{2} \times 260\right)$, which exceeds the 100 she earns with normal effort.

However, the variance in her net earnings with extra effort is $10{,}000 = \frac{1}{2}(60 - 160)^2 + \frac{1}{2}(260 - 160)^2$. Thus, whether Amy chooses to work extra hard depends on how risk averse she is. If she is nearly risk neutral, she works extra hard. However, if she is very risk averse, she puts in only normal effort, receives a modest but predictable wage, and avoids the risk of sometimes earning very little.

The next to last row (*Wage and bonus of 200; Amy is risk neutral*) in Table 20.1 shows the outcome of a bonus contract if Amy is risk neutral. If Amy is risk neutral or nearly risk neutral, she chooses to work hard. If demand is low, the bonus is not triggered, so Paul pays Amy only her base salary of 100 and keeps the residual amount of 200. With high demand, Paul pays Amy 300—the base of 100 plus the bonus of 200 —and keeps the residual of 200 (= 500 − 300). Paul expects to earn $200 = \left(\frac{1}{2} \times 200\right) + \left(\frac{1}{2} \times 200\right)$. Indeed, he earns 200 regardless of demand conditions, so he bears no risk. Thus, if Amy is risk neutral, the bonus leads to efficient payoffs and efficient risk bearing even though Amy bears all the risk. (If Amy were nearly but not quite risk neutral, she would still choose to work hard, but would dislike bearing all the risk.)

The last row (*Wage and bonus of 200; Amy is very risk averse*) in Table 20.1 shows the outcome of a bonus contract if Amy is extremely risk averse. Now, she'd rather have 100 with certainty than take a chance on sometimes netting only 60 after incurring the cost of high effort, so she works only normal hours. Consequently, the total payoffs are low and hence not efficient, but the parties share risk efficiently, with Paul bearing all the risk. Thus, this bonus may—but does not necessarily—induce Amy to work hard.

Many senior executives receive part of their salary in the form of an *option*, which is a type of bonus. An option gives the holder the right to buy up to a certain number of shares of the company at a given price (the *exercise price*) during a specified time interval. An option provides a benefit to the executive if the firm's stock price exceeds the exercise price and is therefore a bonus based on the stock price.

Piece Rates Another common type of contingent contract is a *piece-rate contract*, in which the agent receives a payment for each unit of output the agent produces. Under such a contract, Amy is paid for every serving of ice cream she sells rather than by the hour, which gives her an incentive to work hard, but she bears the risk from fluctuations in demand, which she does not control.

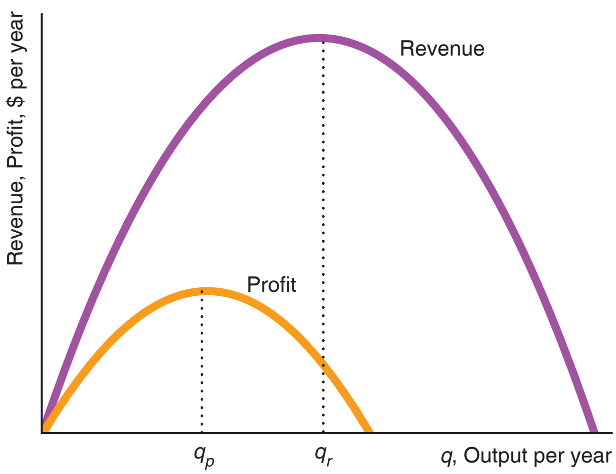
Owners often use piece rates if they can observe output but not labor. Piece rates are commonly used by agricultural, manufacturing, auto glass, and other firms where the employer wants to encourage employees to perform a repetitive job quickly.

Commissions Often, when at least one party cannot observe total profit, but both can observe revenue, they use a revenue-sharing contract, where the agent receives a share of the revenue. For example, Susan, who works in a clothing store, receives a *commission*, which is 5% of the revenue for each item she sells.

As with profit sharing, piece rates and commissions provide an incentive for agents to provide more effort than they would with a fixed-rate contract. However, as with a bonus, this incentive is not necessarily strong enough to offset the agent's cost of extra effort and the agent bears some risk.

Solved Problem 20.4

Peter, the owner of a firm, pays his salesperson, Ann, a commission on her sales (revenue). Thus, Ann has an incentive to maximize revenue. The graph shows how revenue and profit vary with output. Show that if she succeeds in maximizing revenue, she does not maximize the firm's profit.



Answer

Show the quantities of output that maximize both curves and compare them. As the figure shows, Ann maximizes revenue by selling q_r units. Profit is maximized at a smaller quantity, q_p. Thus, if Ann works to maximize sales, she does not maximize profit.[7]

Choosing the Best Contract

Which contract is best for a principal and an agent depends on their attitudes toward risk, the degree of risk, the difficulty in monitoring, and other factors. Often when the parties find that they cannot achieve both efficiency in production and efficiency in risk bearing, they choose a contract that attains neither goal. For example, they may use a contingent contract that fails to achieve efficiency in either production or risk bearing, but it strikes a compromise between the two goals.

The type of contract chosen depends on many factors. Someone who hires a lawyer and who cannot observe the lawyer's work effort or easily judge the quality of the lawyer's work worries about moral hazard. Either party may be particularly concerned about risk bearing.

Lawyers usually work for a fixed fee only if the task or case is very simple, such as writing a will or handling an uncontested divorce. The client has some idea of whether the work is done satisfactorily, so monitoring is relatively easy and little risk is involved. However, a lawyer may not accept a fixed fee if the number of hours required for a more complex project is unknown.

In riskier situations, other types of contracts are more commonly used. When the lawyer is relatively risk averse or when the principal is very concerned that the lawyer works hard, the client may pay the lawyer an hourly wage. However, unless the principal can observe the lawyer's effort, such contracts may result in the lawyer overbilling for hours.

contingent fee
a payment to a lawyer that is a share of the award in a court case (usually after legal expenses are deducted) if the client wins and nothing if the client loses

Alternatively, the parties may use a **contingent fee**: a payment to a lawyer that is a share of the award in a court case (usually after deducting legal expenses) if the client wins and nothing if the client loses. Contingent fee arrangements are particularly common for plaintiffs' lawyers who specialize in auto accidents, medical malpractice, product liability, and other *torts*: wrongful acts in which a person's body, property, or reputation is harmed and for which the injured party is entitled to compensation. Because these plaintiffs' lawyers can typically pool risks across clients, they are less concerned than their clients about risk and are willing to accept contingent fees. Moreover, accident victims often lack the resources to pay for a lawyer's time before winning at trial, so they often prefer contingent contracts.

20.3 Monitoring to Reduce Moral Hazard

A complex contract to prevent moral hazard problems may be unnecessary if the principal can monitor the agent's work. In addition, a principal who can observe the final product but not work effort may be able to avoid moral hazard by making payments contingent on the final product.

[7] Profit is $\pi(q) = R(q) - C(q)$. Profit is maximized where $d\pi(q)/dq = dR(q)/dq - dC(q)/dq = MR - MC = 0$. That is, profit reaches a maximum where marginal revenue, MR, equals marginal cost, MC. In contrast, revenue is maximized where marginal revenue equals zero: $dR(q)/dq = MR = 0$.

shirking
a *moral hazard* in which agents do not provide all the services they are paid to provide

Employees who receive a fixed salary have little incentive except self-respect to work hard if the employer cannot observe them. That is, they may engage in **shirking**: a moral hazard in which agents do not provide all the services they are paid to provide. A firm can reduce such shirking by intensively supervising or monitoring its workers. Monitoring eliminates the asymmetric information problem: Both the employee and the employer know how hard the employee works. If the cost of monitoring workers is low enough, it pays to prevent shirking by carefully monitoring and firing employees who do not work hard.

Firms have experimented with various means of lowering the cost of monitoring. Requiring employees to punch a time clock and installing video cameras to record the work effort are examples of firms' attempts to use capital to monitor job performance. Similarly, by installing assembly lines that force employees to work at a pace dictated by the firm, employers can control employees' work rate.

According to a survey by the American Management Association, nearly two-thirds of employers record employees' voice mail, email, or phone calls; review computer files; or videotape workers. A quarter of the firms that use surveillance don't tell their employees. The most common types of surveillance are tallying phone numbers called and recording the duration of the calls (37%), videotaping employees' work (16%), storing and reviewing email (15%), storing and reviewing computer files (14%), and taping and reviewing phone conversations (10%). Monitoring and surveillance are most common in the financial sector, in which 81% of firms use these techniques. Rather than watching all employees all the time, companies usually monitor selected workers using spot checks.

For some jobs, however, monitoring is counterproductive or not cost effective. Monitoring may lower employees' morale, in turn reducing productivity. Several years ago, Northwest Airlines (now Delta Airlines) took the doors off bathroom stalls to prevent workers from slacking off there. When new management eliminated this policy (and made many other changes as well), productivity increased.

It is usually impractical for firms to monitor how hard salespeople work if they spend most of their time away from the main office. As telecommuting increases, monitoring workers may become increasingly difficult.

A firm's board of directors is supposed to represent shareholders (principals) by monitoring senior executives (agents) to ensure that executives make decisions that are in the shareholders' interests. Bad executives may try to hide their actions from directors or select directors who won't "rat" on them. González et al. (2013) studied firms in which senior executives were engaging in illegal price-fixing, exposing the firm and its shareholders to significant legal liability. They found that senior executives in such firms were more inclined to recruit directors who were likely to be inattentive monitors.

When direct monitoring is very costly, firms may use various financial incentives, which we consider in the next section, to reduce the amount of monitoring that is necessary. Each of these incentives—bonding, deferred payments, and efficiency (unusually high) wages—acts as a *hostage* for good behavior (Williamson, 1983). Workers caught shirking or engaging in other undesirable acts not only lose their jobs but give up the hostage, too. The more valuable the hostage, the less monitoring the firm needs to use to deter bad behavior.

Bonding

A direct approach to ensuring good behavior by agents is to require that they deposit funds guaranteeing their good behavior, just as a landlord requires tenants to post security deposits to ensure that they will not damage an apartment. An employer may require an employee to provide a performance *bond*, an amount of money that is forfeit to the principal if the agent fails to complete certain duties or achieve certain goals. Typically, the agent *posts* (leaves) this bond with the principal or another party, such as an insurance company, before starting the job.

Many couriers who transport valuable shipments (such as jewels) or guards who watch over them have to post bonds against theft and other moral hazards. Similarly, bonds may prevent employees from quitting immediately after receiving costly training (Salop and Salop, 1976). Academics who take a sabbatical—a leave of absence that is supposed to be devoted to training or other activities that increase their future productivity—must typically sign an agreement to pay the college or university a certain sum if they quit within a year after returning from their sabbatical. Most of the other approaches we will examine as strategies for controlling shirking are equivalent to bonding.

Bonding to Prevent Shirking Some employers require a worker to post a bond that is forfeited if the employee is discovered shirking. For example, a professional athlete faces a specified fine (the equivalent of a bond) for skipping a meeting or game. The higher the bond, the less frequently the employer needs to monitor to prevent shirking.

Suppose that the value that a worker puts on the gain from taking it easy on the job is G dollars. If a worker's only potential punishment for shirking is dismissal if caught, some workers will shirk.

Suppose, however, that the worker must post a bond of B dollars that the worker forfeits if caught not working. Given the firm's level of monitoring, the probability that a worker is caught is θ. Thus, a worker who shirks expects to lose θB.[8] A risk-neutral worker chooses not to shirk if the certain gain from shirking, G, is less than or equal to the expected penalty, θB, from forfeiting the bond if caught: $G \le \theta B$. Therefore, the minimum bond that discourages shirking is

$$B = \frac{G}{\theta}. \tag{20.1}$$

Equation 20.1 shows that the bond must be larger for the higher the value that the employee places on shirking and the lower the probability that the worker is caught.

Trade-Off Between Bonds and Monitoring Thus, the larger the bond, the less monitoring is necessary to prevent shirking. Suppose that a worker places a value of $G = \$1{,}000$ a year on shirking. A bond that is large enough to discourage shirking is \$1,000 if the probability of being caught is 100%, \$2,000 at 50%, \$5,000 at 20%, \$10,000 at 10%, and \$20,000 if the probability of being caught is only 5%.

Solved Problem 20.5

MyLab Economics Solved Problem

Workers post bonds of B that are forfeited if they are caught stealing (but no other punishment is imposed). Each unit of monitoring, M, raises the probability that a firm catches a worker who steals, θ, by 5%. A unit of M costs \$10. A worker can steal a piece of equipment and resell it for its full value of G dollars. What is the optimal M that the firm uses if it believes that workers are risk neutral? In particular, if $B = \$5{,}000$ and $G = \$500$, what is the optimal M?

[8]The expected penalty is $\theta B + (1 - \theta)0 = \theta B$, where the first term on the left side is the probability of being caught times the fine of B and the second term is the probability of not being caught and facing no fine.

Answer

1. *Determine how many units of monitoring are necessary to deter stealing*. The least amount of monitoring that deters stealing is the amount at which a worker's gain from stealing equals the worker's expected loss if caught. A worker is just deterred from stealing when the gain, G, equals the expected penalty, θB. Thus, the worker is deterred when the probability of being caught is $\theta = G/B$. The number of units of monitoring effort is $M = \theta/0.05$, because each unit of monitoring raises θ by 5%.

2. *Determine whether monitoring is cost effective*. It pays for the firm to pay for M units of monitoring only if the expected benefit to the firm is greater than the cost of monitoring, $\$10 \times M$. The expected benefit if stealing is prevented is G, so monitoring pays if $G > \$10 \times M$, or $G/M > \$10$.

3. *Solve for the optimal monitoring in the special case*. The optimal level of monitoring is
$$M = \frac{\theta}{0.05} = \frac{G/B}{0.05} = \frac{500/5{,}000}{0.05} = \frac{0.1}{0.05} = 2.$$
It pays to engage in this level of monitoring because $G/M = \$500/2 = \$250 > \$10$.

Problems with Bonding Employers like the bond-posting solution because it reduces the amount of employee monitoring necessary to discourage moral hazards such as shirking and thievery. Nonetheless, firms use explicit bonding only occasionally to prevent stealing, and they rarely use it to prevent shirking.

Two major problems are inherent in posting bonds. First, to capture a bond, an unscrupulous employer might falsely accuse an employee of stealing. An employee who fears such employer opportunism might be unwilling to post a bond. One possible solution to this problem is for the firm to develop a reputation for not behaving in this manner. Another possible approach is for the firm to make the grounds for forfeiture of the bond objective and thus verifiable by others.

A second problem with bonds is that workers may not have enough wealth to post them. In Solved Problem 20.5, if the worker could steal \$10,000, and if the probability of being caught were only 5%, shirking would be deterred only if a risk-neutral worker were required to post a bond of at least \$200,000.

Principals and agents use bonds when these two problems are avoidable. Bonds are more common in contracts between firms than in those between an employer and employees. Moreover, firms have fewer problems than do typical employees in raising funds to post bonds.

Construction contractors sometimes post bonds to guarantee that they will satisfactorily finish their work by a given date. Both parties can verify whether the work is completed on time, so the principal has relatively little chance to engage in opportunistic behavior.

Application

Capping Oil and Gas Bankruptcies

Why are most onshore gas and oil producers small firms? A major reason is to avoid liability. Being small allows them to produce as much as they can and if they cause environmental damages—water pollution, toxic gas releases, of explosions—greater than their assets, they avoid liability by declaring bankruptcy. They are "judgment proof." A large firm with substantial assets would pay the damages and stay in business.

Texas has roughly 5,000 oil- and gas-producing firms. Most of these have less than two million dollars in annual revenue—much less than their liability exposure. However, as of 2001, Texas required these firms to post a *surety bond*, which is an insurance contract that obligates the insurer to compensate the state for environmental damages by the insured oil or gas producer. Insurance companies set a high premium for a firm with a bad safety record or that has little incentive to act prudently because it is financially weak. In contrast, a large, financially secure firm is less likely to act irresponsibly and hence pays a lower premium.

Boomhower (2016) showed that the bond requirement improves firms' safety incentives. As soon as the bond mandate went into effect, 6% of firms exited the market (twice the usual rate). The exiting firms were primarily small and had poor environmental records. These exiting firms transferred 88% of their oil and gas leases to larger firms. The smallest 80% of the remaining firms reduced oil production, while the large firms' production was unaffected. That is, the ability to avoid responsibility prior to bonding inflated the number of small firms and their production.

The amount of environmental damage fell after the bond mandate went into effect. Many fewer firms left their wells unplugged at the end of production, which causes a serious risk of groundwater pollution. Well blowouts and water protection violations also fell substantially.

Deferred Payments

Instead of having their employees post bonds, firms can equivalently deter bad behavior by using deferred payments. For example, a firm pays new workers a low wage for some initial period of employment. Then, over time, the firm fires any workers caught shirking, and the remaining workers receive higher wages. In another form of deferred wages, the firm provides a pension that rewards only hard workers who stay with the firm until retirement. *Deferred payments serve the same function as bonds*. They raise the cost of being fired, so less monitoring is necessary to deter shirking.

Workers care about the present value (Chapter 16) of their earnings stream over their lifetime. A firm may offer its workers one of two wage payment schemes. In the first, the firm pays w per year for each year that the employee works for the firm. In the second arrangement, the starting wage is less than w but rises over the years to a wage that exceeds w.

If employees can borrow against future earnings, those who work for one company for their entire career are indifferent between the two wage payment schemes if those plans have identical present values. The firm, however, prefers the second payment method because employees work harder to avoid losing their job and the high future earnings.

Reduced shirking leads to greater output. If the employer and employee share the extra output in the form of higher profit and lifetime earnings, both the firm and workers prefer the deferred-payment scheme that lowers incentives to shirk.

A drawback of the deferred-payment approach is that, like bond posting, it can encourage employers to engage in opportunistic behavior. For example, an employer might fire nonshirking senior workers to avoid paying their higher wages or their pensions and replace them with less expensive junior workers. However, if the firm establishes a reputation for not firing senior workers unjustifiably, the deferred-payment system helps prevent shirking.

Efficiency Wages

efficiency wage
an unusually high wage that a firm pays workers as an incentive to avoid shirking

As we've seen, the use of bonds and deferred payments discourages shirking by raising an employee's cost of losing a job. An alternative is for the firm to pay an **efficiency wage**: an unusually high wage that a firm pays workers as an incentive to avoid shirking.[9] If a worker who is fired for shirking can immediately go to another firm and earn the same wage, the worker risks nothing by shirking. However, a high wage payment raises the cost of getting fired, so it discourages shirking.[10]

How Efficiency Wages Act like Bonds Suppose that a firm pays each worker an efficiency wage w, which is more than the *going wage* $\underline{w}$ that an employee could earn elsewhere. We now show that the less frequently the firm monitors workers, the greater the wage differential must be between w and $\underline{w}$ to prevent shirking.

A worker decides whether to shirk by comparing the expected loss of earnings from getting fired to the value, G, that the worker places on shirking. A shirking worker expects to lose $\theta(w - \underline{w})$,where θ is the probability that a shirking worker is caught and fired and the term in parentheses is the lost earnings from being fired. A risk-neutral worker does not shirk if the expected loss from being fired is greater than or equal to the gain from shirking (see Appendix 20A):

$$\theta(w - \underline{w}) \geq G. \tag{20.2}$$

The smallest amount by which w can exceed $\underline{w}$ and prevent shirking is determined where this expression holds with equality, $\theta(w - \underline{w}) = G$, or

$$w - \underline{w} = \frac{G}{\theta}. \tag{20.3}$$

The extra earnings, $w - \underline{w}$, in Equation 20.3 serve the same function as the bond, B, in Equation 20.1 in discouraging bad behavior.

Suppose that the worker gets $G = \$1{,}000$ pleasure a year from not working hard and $\underline{w}$ is \$20,000 a year. If the probability that a shirking worker is caught is $\theta = 20\%$, then the efficiency wage w must be at least \$25,000 to prevent shirking. With greater monitoring, so that θ is 50%, the minimum w that prevents shirking is \$22,000. From the possible pairs of monitoring levels and efficiency wages that deter shirking, the firm picks the combination that minimizes its labor cost.

Efficiency Wages and Unemployment We've argued that it is in a firm's best interest to pay more than the "going wage" to discourage shirking. The problem with this conclusion is that if it pays for one firm to raise its wage, it pays for all firms to do so. But if all firms raise their wages and pay the same amount, no one firm can discourage shirking by paying more than the others.

Nonetheless, the overall high wages do help prevent shirking. Because all firms are paying above the competitive wage, their labor demand falls, causing unemployment. Now if a worker is fired, the worker remains unemployed while searching for a new job. Thus, the amount that the fired worker earns elsewhere, $\underline{w}$, is less than w because

---

[9] The discussion of efficiency wages is based on Yellen (1984), Stiglitz (1987), and especially Shapiro and Stiglitz (1984).

[10] Economists have varying explanations for why efficiency wages lead to higher productivity. Some claim that employers in less-developed countries pay an efficiency wage—more than they need to hire workers—to ensure that workers can afford to eat well enough that they can work hard. Other economists (such as Akerlof, 1982) and management experts contend that the higher wage acts like a gift, making workers feel beholden or loyal to the firm, so that less (or no) monitoring is needed. Efficiency wages may also attract better applicants, raising the quality of a firm's workforce.

of this period of unemployment.[11] As a result, the (high) efficiency wages discourage shirking by creating unemployment.

One implication of this theory is that unemployment benefits provided by the government actually increase the unemployment rate. Such benefits raise $\underline{w}$, decrease the markup of w over $\underline{w}$, and reduce the loss from being fired. Thus, to discourage shirking, firms have to raise their efficiency wage even higher, and even more unemployment results.

Monitoring Outcomes

So far we've concentrated on monitoring by employers looking for bad behavior as it occurs. If the employer detects shirking or other bad behavior after the fact, the employer fires or penalizes the offending employee. This punishment discourages shirking in the future.

It is often very difficult to monitor bad behavior when it occurs but relatively easy to determine it after the fact. As long as a contract holds off payment until after the principal checks for bad behavior, after-the-fact monitoring discourages bad behavior. For example, an employer can check the quality of an employee's work. If it is substandard, the employer can force the employee to make it right.

Insurance companies frequently use this approach in contracts with their customers. Insurance firms try to avoid extreme moral hazard problems by offering contracts that do not cover spectacularly reckless, stupid, or malicious behavior. If an insurance company determines after the fact that a claim resulted from reckless behavior rather than chance, the firm refuses to pay.

For example, an insurance company will not pay damages for a traffic accident if the insured driver was drunk at the time. A house insurance company disallows claims due to an explosion that resulted from an illegal activity such as making methamphetamine. It will certainly disallow claims by arsonists who torch their own homes or businesses. Life insurance companies may refuse to pay benefits to the family of someone who commits suicide (as in the play *Death of a Salesman*).

Finding out about moral hazards after they occur is too late if wrongdoers cannot be punished at that time. Indeed, it's pointless to monitor after the fact if punishment is then impossible or impractical. Although it's upsetting to find that you've been victimized, you can't do anything beyond trying to prevent the situation from happening again.

20.4 Checks on Principals

To this point, we have concentrated on situations in which the agent knows more than does the principal. Sometimes, however, the principal may have asymmetric information and engage in opportunistic behavior.

Because employers (principals) often pay employees (agents) after work is completed, employers have many opportunities to exploit workers. For example, a dishonest employer can underpay after falsely claiming that a worker took time off or that some of the worker's output was substandard. The employer can decrease piece rates over time, after employees are committed to this payment system. Employers who provide bonuses can underreport the firm's output or profit. An employer can

[11] If x is the share of time that the fired worker remains unemployed, the worker's expected earnings are $\underline{w} = (1 - x)w + x0 = (1 - x)w$.

dock earnings, claim that an employee bond is forfeit, or refuse to make deferred payments such as pensions after dishonestly claiming that a monitored worker engaged in bad behavior. Efficient contracts prevent or reduce such moral hazard problems created by employers as well as those caused by employees.

Requiring that a firm post a bond can be an effective method of deterring the firm's opportunistic behavior. For example, a firm may post bonds to ensure that it has the means of paying current wages and future pensions.

Another strategy for preventing a firm from acting opportunistically is to eliminate asymmetric information by requiring the employer to reveal relevant information to employees. For example, an employer can provide access to such information by allowing employee representatives to sit on the company board—from which vantage point they can monitor the firm's behavior. To induce workers to agree to profit sharing, a firm may provide workers with information about the company's profit by allowing them (or an independent auditor) to check its accounts. Alternatively, the firm may argue that its stock closely mirrors its profit and suggest that the known stock price be used for incentive payments.

As another means of conveying information to employees, firms may seek to establish a good reputation. For instance, a firm may publicize that it does not make a practice of firing senior employees to avoid paying pensions. The better the firm's reputation, the more likely workers are to accept a deferred payment scheme, which deters shirking.

When firms find these approaches infeasible, they may use inefficient contracts that pay employees based on easily observed revenues rather than less credible profit reports. The following Application discusses a particularly damaging but common type of inefficient contract.

Application

Layoffs Versus Pay Cuts

During recessions and depressions, demand for most firms' products fall. Many firms respond by laying off workers and reducing production rather than by lowering wages and keeping everyone employed. The average real U.S. weekly earnings fluctuated in a narrow band—$333 to $345— from 2002 through 2012, though it rose to $368 by May 2016. In contrast, the U.S. unemployment rate over this period has fluctuated substantially. It started at 5.7% in 2002, rose to 6.1% in 2003, dropped to 4.4% in 2007, rose to 10.0% in 2010, and then dropped steadily to 4.7% by May 2016.

If both sides agree to it, a wage reduction policy benefits firms and workers alike. Collectively, workers earn more than they would if they were laid off. Because the firm's costs would fall, it sells more during the downturn than it otherwise could, so its profit is higher than with layoffs. Firms that provide relatively low wages and then share profits with employees achieve this type of wage flexibility.

Why then are wage reductions less common than layoffs? A major explanation involves asymmetric information: Workers, unlike the firm, don't know whether the firm is actually facing a downturn, so they don't agree to wage cuts. In short, they don't trust the firm to tell them the truth. They fear that the firm will falsely claim that economic conditions are bad to justify a wage cut. If the firm has to lay off workers—an action that hurts the firm as well as the workers—the firm is more likely to be telling the truth about economic conditions.[12]

[12]In 2010, after several years of the Great Recession, when everyone knew that the downturn was real, increasingly pay cuts replaced layoffs, especially by state and local government employers. Sub-Zero, which makes refrigerators and other appliances, told its workers it might close one or more factories and lay off 500 employees unless they accepted a 20% cut in wages and benefits.

We illustrate this reasoning in the following matrix, which shows the payoffs if wages fall during downturns. The value of output produced by each worker is \$21 during good times and \$15 during bad times. The lower left of each cell is the amount the firm pays workers. The firm pays employees \$12 per hour if it reports that economic conditions are good and \$8 if it says that conditions are bad. The amount the firm keeps is in the upper right of each cell. If economic conditions are bad, the firm earns more by reporting these bad conditions, \$7, than it earns if it says that conditions are good, \$3. Similarly, if conditions are good, the firm earns more if it claims that conditions are bad, \$13, than if it says that they are good, \$9. Thus, regardless of the true state, the firm always claims that conditions are bad.

Wage Cut

		Firm's Claim About Conditions	
		Bad	*Good*
Actual Conditions	*Bad*	8, 7	12, 3
	Good	8, 13	12, 9

To shield themselves from such systematic lying, employees may insist that the firm lay off workers whenever it says that conditions are bad. This requirement provides the firm with an incentive to report the true conditions. In the next matrix, the firm must lay off workers for half of each period if it announces that times are bad, causing the value of output to fall by one-third. Because they now work only half the time, workers earn only half as much, \$6, as they earn during good times, \$12. If conditions are bad, the firm makes more by telling the truth, \$4, than by claiming that conditions are good, \$3. In good times, the firm makes more by announcing that conditions are good, \$9, than by claiming that they are bad, \$8. Thus, the firm reports conditions truthfully.

Worker Layoff (for half of any period the firm claims is bad)

		Firm's Claim About Conditions	
		Bad	*Good*
Actual Conditions	*Bad*	6, 4	12, 3
	Good	6, 8	12, 9

With the wage-cut contract in which the firm always says that conditions are bad, workers earn \$8 regardless of actual conditions. If economic conditions are good half the time, the firm earns an average of $\$10 = \left(\tfrac{1}{2} \times \$7\right) + \left(\tfrac{1}{2} \times \$13\right)$. Under the contract that requires layoffs, the workers earn an average of $\$9 = \left(\frac{1}{2} \times \$6\right) + \left(\frac{1}{2} \times \$12\right)$ and the firm earns an average of $\$6.50 = \left(\frac{1}{2} \times \$4\right) + \left(\frac{1}{2} \times \$9\right)$.

Therefore, the firm prefers the wage-cut contract and the workers favor the layoff contract. However, if the workers could observe actual conditions, both parties would prefer the wage-cut contract. Workers would earn an average of $\$10 = \left(\frac{1}{2} \times \$8\right) + \left(\frac{1}{2} \times \$12\right)$, and the firm would make $\$6.50 = \left(\frac{1}{2} \times \$4\right) + \left(\frac{1}{2} \times \$9\right)$. With the layoff contract, total payoffs are lower because of lost production. Thus, socially inefficient layoffs may be used because of the need to keep relatively well-informed firms honest.

20.5 Contract Choice

We have examined how to construct a single contract to prevent moral hazards. Often, however, a principal gives an agent a choice of contract. By observing the agent's choice, the principal obtains enough information to prevent agent opportunism.

Firms want to avoid hiring workers who will shirk. Employers know that not all workers shirk, even when given an opportunity to do so. So, rather than focusing on stopping lazy workers from shirking, an employer may concentrate on hiring only industrious people. With this approach, the firm seeks to avoid *moral hazard* problems by preventing *adverse selection*, whereby lazy employees falsely assert that they are hardworking.

As discussed in Chapter 19, employees may *signal* to employers that they are productive. For example, if only nonshirking employees agree to work long hours, a commitment to work long hours serves as a reliable signal. In addition, employees can signal by developing a reputation as hard workers. To the degree that employers can rely on this reputation, sorting is achieved.

When workers cannot credibly signal, firms may try to *screen out* bad workers. One way firms can determine which prospective employees will work hard and which will shirk is to give them a choice of contracts. A firm can distinguish between good and bad workers if job candidates who are hard workers select a contingent contract in which their pay depends on how hard they work and lazy job applicants choose a fixed-fee contract.

Suppose that a firm wants to hire a salesperson who will run its Cleveland office and that the potential employees are risk neutral. A hardworking salesperson sells \$100,000 worth of goods a year, but a lazy one sells only \$60,000 worth (see Table 20.2). A hard worker can earn \$30,000 if employed elsewhere, so the firm considers using a contingent contract that pays a salesperson a 30% commission on sales.

If the firm succeeds in hiring a hard worker, the salesperson makes $\$30{,}000 = \$100{,}000 \times 0.30$. The firm's share of sales is \$70,000. The firm has no costs of production (for simplicity), but maintaining this branch office costs the firm \$50,000 a year. The firm's profit is therefore \$20,000. If the firm hires a lazy salesperson under the same contract, the salesperson makes \$18,000, the firm's share of sales is \$42,000, and the firm loses \$8,000 after paying for the office.

Thus, the firm wants to hire only a hard worker. Unfortunately, the firm does not know in advance whether a potential employee is a hard worker. To acquire this information, the firm offers a potential employee a choice of contracts:

- **Contingent contract.** No salary and 30% of sales
- **Fixed-fee contract.** Annual salary of \$25,000, regardless of sales

Table 20.2 Firm's Spreadsheet

	Contingent Contract (30% of Sales), $	Fixed-Fee Contract ($25,000 Salary), $
Hard Worker		
Sales	100,000	100,000
− Salesperson's pay	−30,000	−25,000
= Firm's net revenue	70,000	75,000
− Office expenses	−50,000	−50,000
= Firm's profit	20,000	25,000
Lazy Worker		
Sales	60,000	60,000
− Salesperson's pay	−18,000	−25,000
= Firm's net revenue	42,000	35,000
− Office expenses	−50,000	−50,000
= Firm's profit	−8,000	−15,000

A prospective employee who doesn't mind hard work would earn $5,000 more by choosing the contingent contract. In contrast, a lazy candidate would make $7,000 more from a salary than from commissions. If an applicant chooses the fixed-fee contract, the firm knows that the person does not intend to work hard and decides not to hire that person.

The firm learns what it needs to know by offering this contract choice as long as the lazy applicant does not pretend to be a hard worker by choosing the contingent contract. Under the contingent contract, the lazy person makes only $18,000, but that offer may dominate others available in the market. If this pair of contracts fails to sort workers, the firm may try different pairs. If all these choices fail to sort, the firm must use other means such as monitoring to prevent shirking.

Challenge Solution

Clawing Back Bonuses

The Challenge at the beginning of the chapter asks whether evaluating a manager's performance over a longer period and using delayed compensation of clawback provisions benefit shareholders. The answer depends on whether the reward a manager receives in the short run induces the manager to sacrifice long-run profit for short-run personal gains.

Managers prefer receiving pay sooner rather than later because money today is worth more than the same amount later. Many managers receive a bonus based on a firm's annual profit. If a manager can move a major sale from January of next year to December of this year, the firm's total profit over the two years does not change, but the manager receives the resulting performance-based bonus this year rather than next year. The owners of the firm are probably not very concerned with such shifts over time, as they do not lower the present value of long-run profits substantially.

Of more concern are managers who increase this year's profit in a way that lowers profit in later years. Many firms pay a bonus on a positive profit but do not impose fines or penalties (negative bonuses) for a loss (negative profit). Suppose that a particular policy results in a large profit this year but a larger loss next year. If the manager gets a bonus based on each year's profit, the manager receives a large bonus this year and no bonus next year.

In an extreme case, a manager engages in reckless behavior that increases this year's profit but bankrupts the firm next year. The manager plans to grab this year's bonus and then disappear. Many mortgage and financial instrument managers engaged in such reckless and irresponsible behavior leading up to the 2007–2009 financial meltdown. Bad decisions at Merrill Lynch, a wealth management firm, cost shareholders billions of dollars, but senior managers kept bonuses despite the negative effects of their decisions on shareholders.

A firm can discourage bad managerial behavior if the firm can clawback a bonus given this year if the firm does badly next year. Without a clawback, the manager has a greater incentive to adopt this policy of increasing profit in one year at the cost of lower profit the next.

Equivalently, the firm may base bonuses on profits over several years. If a bonus is a percentage of the profit over two years, the manager who increases profit this year at the expense of a larger loss next year receives no bonus.

To illustrate why paying bonuses based on a longer period provides a better incentive structure, we examine the case of Jim, who is an executive at an auto loan company. We look at a two-year period. Initially, Jim receives 10% of the amount of the loans he makes in the first year. He can loan to two groups of customers. Customers in one group have excellent financial histories and repay their loans on time. Loans to this group produce revenue of $10 million this year, so that over the two-year period, the firm nets $9 million after paying Jim. Customers in the other group are much more likely to default. That group produces $30 million in revenue this year, but their defaults in the second year cost the firm $40 million. After paying Jim $3 million in the first year, the firm suffers a $13 million loss over the two years (ignoring discounting).

Because Jim prefers receiving $4 million by loaning to both groups to $1 million from loaning to only the good risks, he may expose the firm to devastating losses in the second year. He may be happy earning a gigantic amount in the first year even if he's fired in the second year. However, if his bonus is a share of the profit over two years, he no longer wants to make loans to the risky group.

Starting in 2012, Morgan Stanley paid bonuses to high-income employees over a three-year period. Similarly, Apple granted stock options and related benefits every two years rather than every year. In 2016, Wells Fargo Bank clawed back $41 million from its chairman and chief executive in response to a scandal.

Summary

1. **The Principal-Agent Problem.** A moral hazard may occur if a principal contracts with an agent to perform an action and the principal cannot observe the agent's actions so that the agent takes advantage of the principal. For example, if the owner cannot observe how hard an employee works, the employee may shirk. This moral hazard reduces their joint profit. An efficient contract leads to efficiency in production (eliminating moral hazards maximizes joint profit) and efficiency in risk bearing (the less-risk-averse party bears more of the risk). Whether production is efficient depends on the contract that the principal and the agent use and the degree to which their information is asymmetric. Ideally, the contract or agreement between the parties is efficient, so that they maximize the total profit of the parties and share risk optimally.

2. **Using Contracts to Reduce Moral Hazard.** A principal and an agent may agree to a contract that eliminates a moral hazard or at least strikes a balance between reducing the moral hazard and allocating risk optimally. Contracts that eliminate moral hazards often require the agent to bear the risk. If the agent is

more risk averse than the principal, the parties may trade off a reduction in production efficiency to lower risk for the agent.

3. **Monitoring to Reduce Moral Hazard.** Because of asymmetric information, an employer must normally monitor workers' efforts to prevent shirking. Less monitoring is necessary as the employee's interest in keeping the job increases. The employer may require the employee to post a large bond that they forfeit if the employee is caught shirking, stealing, or otherwise misbehaving. If an employee cannot afford to post a bond, the employer may use deferred payments or efficiency wages—unusually high wages—to make it worthwhile for the employee to keep the job. Employers may also be able to prevent shirking by engaging in after-the-fact monitoring. However, such monitoring works only if bad behavior can be punished after the fact.

4. **Checks on Principals.** Often both agents and principals can engage in opportunistic behavior. If a firm must reveal its actions to its employees, it is less likely to be able to take advantage of the employees. To convey information, an employer may let employees participate in decision-making meetings or audit the company's books. Alternatively, an employer may make commitments so that it is in the employer's best interest to tell employees the truth. These commitments, such as laying off workers rather than reducing wages during downturns, may reduce moral hazards but lead to nonoptimal production.

5. **Contract Choice.** A principal may be able to obtain valuable information from an agent by offering a choice of contracts. Employers avoid moral hazard problems by preventing adverse selection. For example, they may present potential employees with a choice of contracts, prompting hardworking job applicants to choose one contract and lazy candidates to choose another.

Questions

Select questions are available on MyLab Economics*; * = answer appears at the back of this book;* **A** *= algebra problem;* **C** *= calculus problem.*

1. The Principal-Agent Problem

*1.1 Under a fee-splitting arrangement, the charge for professional services is shared with the party who made the referral. Many countries consider such arrangements to represent a conflict of interest for the client since the client may not receive the services most appropriate for his or her well-being. Consequently, many countries consider fee splitting to be unethical and unacceptable, and fee splitting is often done covertly. What is the implication of this fee-sharing arrangement on the cost of the overall bill? Why?

1.2 The European Food Safety Authority provides scientific advice on questions related to food and feed safety, animal health and welfare, plant health, nutrition, and related environmental issues, and communicates its findings publicly. When public health risks are detected in food, information is shared between the European Union member states so that they can take the appropriate measures, including withdrawal or recall. However, to further enhance product quality, safety, and acceptance, companies could have their food products tested and certified to ensure compliance with national or international standards. This additional cost would be reflected in a higher price for their products. Discuss the implications of third-party testing on the moral hazard problem.

1.3 In June 2016, *The Telegraph* reported that some doctors in the United Kingdom were receiving payments from pharmaceutical companies for recommending that their drugs should be prescribed to patients in the National Health Service. Does this represent a conflict of interest? Explain.

1.4 The Government of India introduced a crop insurance scheme in February 2016 to provide financial support to farmers in the event of crop failure due to natural calamities, pests, and disease. Insurance coverage is compulsory for some farmers but voluntary for others. Part of the premium rate is paid for by the national and state governments. A government that contracts with a private insurer to deliver a program may compensate that insurer in alternative ways. Discuss which of the following compensation schemes are more likely to lead to opportunistic behavior by insurance companies: the insurer receives a percentage of each approved claim, the insurer receives an hourly rate to handle claims, the insurer receives a flat fee, or the insurer sets a premium that the government subsidizes.

*1.5 Many university bookstores will buy back used textbooks. Depending on the bookstore's present stock and which textbooks are required for upcoming courses both on campus and on campuses elsewhere, a student could receive as much as 50% of the retail purchase price after a course ends. Why would a bookstore make such a commitment in terms of dealing with moral hazard concerns?

1.6 While there are many types of mutual funds, they all pool money provided by individual investors to purchase and manage securities. Investors are charged a management fee to pay for the time and expertise of the fund manager in selecting securities and managing the portfolio in keeping with the fund's investment strategy. Management fees are typically based on a percentage of assets under management (the amount contributed by investors). Is this way of compensating mutual fund managers in the best interests of the investors? Explain.

1.7 A study published in April 2015 in the *Academic Emergency Medical* journal found that 85% of doctors surveyed admitted that they ordered tests not because they were necessary for medical reasons but to ensure that they did not overlook something that would help them diagnose their patients (and thus protect them against malpractice suits) (*Time*, 2015; Wiley Online Library, 2015). Do these findings demonstrate moral hazard or is there an alternative explanation for conducting unnecessary tests?

1.8 Sarah, the manager of a bank, can make one of two types of loans. She can loan money to local firms, and have a 75% probability of earning \$100 million and a 25% probability of earning \$80 million. Alternatively, she can loan money to oil speculators, and have a 25% probability of earning \$400 million and a 75% probability of losing \$160 million (due to loan defaults by the speculators). Sarah receives 1% of the bank's earnings. She believes that if the bank loses money, she can walk away from her job without repercussions, although she will not receive any compensation. Sarah and the bank's shareholders are risk neutral. How does Sarah invest the bank's money if all she cares about is maximizing her personal expected earnings? How would the stockholders prefer that Sarah invest the bank's money? (*Hint*: See Solved Problem 20.1.) **A**

2. Using Contracts to Reduce Moral Hazard

2.1 Padma has the rights to any treasure on the sunken ship the *Golden Calf*. Aaron is a diver who specializes in marine salvage. If Padma is risk averse and Aaron is risk neutral, does paying Aaron a fixed fee result in efficiency in risk bearing and production? Does your answer turn on how predictable the value of the sunken treasure is? Would another compensation scheme be more efficient?

2.2 Insurance companies that enter into contracts with employers to provide dental coverage limit the frequency of certain services such as regular checkups or tooth cleaning. How do such restrictions affect moral hazard and risk bearing? Show these results in a graph. (*Hint*: See Solved Problem 20.2.)

2.3 Health care coverage in Australia is provided through Medicare, which is administered by the national government. Doctors are paid mostly on a fee-for-service basis but sometimes also through incentives for enhanced services such as after-hours care, managing patients with chronic conditions, and practicing in rural areas. Proposals to address rising health care costs include patient co-payments for doctor services, an expanded role for private insurers (managed care), and an annual lump-sum payment based on the number of patients a doctor has (capitation). What would be the implications for moral hazards and for risk bearing of using lump-sum payments to compensate doctors for their services?

*2.4 Priscilla hires Arnie to manage her store. The left column of the table shows Arnie's effort. Each cell shows the net profit to Priscilla (ignoring Arnie's cost of effort).

	Low Demand	High Demand
Low Effort	20	40
Medium Effort	40	80
High Effort	80	100

Arnie's personal cost of effort is 0 at low effort, 10 at medium effort, and 30 at high effort. It is equally likely that demand will be low or high. Arnie and Priscilla are risk neutral.

They consider two possible contracts: (1) *fixed fee*: Arnie receives a fixed wage of 10; and (2) *profit sharing*: Arnie receives 50% of the firm's net income but no wage.

a. What happens if they use the fixed fee contract?
b. What happens if they use the profit-sharing contract?
c. Which contract does each prefer? **A**

2.5 In the situation described in the previous question, how do your answers change if Arnie's first contract changes so that he receives a basic fixed wage of 10 and, in addition, a bonus equal to 80% of any net income?

2.6 Patrick, the owner, makes the same offer to the manager at each of his stores: "At the end of the year, pay me a lump-sum of \$100,000, and you can keep any additional profit." Astrid, a manager at one of the stores, gladly agrees, knowing that the total profit at the store will substantially exceed \$100,000 if it is run well. If she is interested in maximizing her earnings, will Astrid act in a manner that maximizes the store's total profit? Use a figure to illustrate your answer. (*Hint*: See Solved Problem 20.3.)

*2.7 One firm (the contractee) may contract with another firm (the contractor) to provide a product or component that is unique to the contractee. As it is unique, it cannot be sold by the contractor to third parties. Suppose the contractee subsequently refuses to accept the product on the basis that it does not meet the quality standards set out in the contract. Assuming the quality standards have not been violated, why might a contractee do this, and what recourse would be available to the contractor?

*2.8 Zhihua and Pu are partners in a store in which they do all the work. They split the store's *business profit* equally (ignoring the opportunity cost of their own time in calculating this profit). Does their business profit-sharing contract give them an incentive to maximize their joint economic profit if neither can force the other to work? (*Hint*: Imagine Zhihua's thought process late one Saturday night when he is alone in the store, debating whether to keep the store open a little later or to go out on the town, and see Solved Problem 20.4.)

2.9 In Solved Problem 20.4, does joint profit increase, decrease, or remain the same as the share of revenue going to Ann increases?

*2.10 Jack and Jill live in different cities. Regardless of which one chooses to fly to visit the other, they agree to split the cost of the flight equally. What is the implication of this fee-sharing arrangement? (*Hint*: See Solved Problem 20.4.)

2.11 Diego runs a store for Mateo, the owner. Suppose that the inverse demand curve for the store's product is $p = 160 - 0.1Q$, where Q is the quantity demanded and p is the price per unit. The cost to Diego of running the store is given by the function $C(Q) = 20Q$.

a. If Mateo and Diego share the profits equally, what is the profit-maximizing quantity, price, and payoff for each?

b. If Mateo takes 50% of the earnings, what is the profit-maximizing quantity, price, and payoff for each? Is joint profit maximized?

c. If Mateo charges Diego a fixed license fee of 30,000 to run the store, what is the profit-maximizing quantity, price, and payoff for each? Is joint profit maximized?

d. If Mateo and Diego share the earnings equally, what is the profit-maximizing quantity, price, and payoff for each? Is joint profit maximized? (*Hint*: See Solved Problems 20.3 and 20.4.) **C**

2.12 Suppose that a textbook author is paid a royalty of a share of the revenue from sales where the revenue is $R = pq$, p is the competitive market price for textbooks, and q is the number of copies of this textbook (which is similar to others on the market) sold. The publisher's cost of printing and distributing the book is $C(q)$. Determine the equilibrium, and compare it to the outcome that maximizes the sum of the payment to the author plus the firm's profit. Answer using both math and a graph. Why do you think royalties are usually a share of revenue rather than profit? **C**

2.13 Suppose now that the textbook publisher in the previous question faces a downward-sloping demand curve. The revenue is $R(Q)$, and the publisher's cost of printing and distributing the book is $C(Q)$. Compare the equilibria for the following compensation methods in which the author receives the same total compensation from each method:

a. The author is paid a lump sum, $\mathcal{L}$.

b. The author is paid a share of the revenue.

c. The author receives a lump-sum payment and a share of the revenue.

Why do you think that usually authors receive a share of revenue? **C**

3. Monitoring to Reduce Moral Hazard

3.1 Many law firms consist of partners who share profits. Upon making partner, a lawyer must post a bond, a large payment to the firm that the lawyer forfeits for bad behavior. Why?

*3.2 In Solved Problem 20.5, a firm calculated the optimal level of monitoring to prevent stealing. If $G = \$500$ and $\theta = 20\%$, what is the minimum bond that deters stealing? **A**

3.3 In the previous question, suppose that for each extra \$1,000 of bonding that the firm requires a worker to post, the firm must pay that worker \$10 more per period to get the worker to work for the firm. What is the minimum bond that deters stealing? (*Hint*: See Solved Problem 20.5.) **A**

3.4 Explain why full employment may be inconsistent with no shirking.

3.5 To strengthen the incentive for hospitals to avoid making mistakes, public hospitals in Australia received no funding after July 1, 2017, for sentinel events (for example, operating on the wrong body part, incompatible blood transfusions, medication errors leading to the death of a patient, instruments or other foreign matter left inside a patient after surgery). Financial penalties for hospital-acquired complications and avoidable hospital readmissions are also being considered. Is this new policy designed to deal with adverse selection or moral hazard? Is it likely to help? Explain.

3.6 Former rental cars sell on the used car market for lower prices than cars of the same model and year that individuals formerly owned. Does this price difference reflect adverse selection or moral hazard? Could car rental companies reduce this problem by carefully inspecting rental cars for damage when renters return such cars? Why do car companies normally perform only a cursory inspection?

3.7 Many substandard condo developments have been built by small corporations that declare bankruptcy or go out of business when legal actions are started against them by condo buyers. What legal remedies might reduce this moral hazard problem? If you were considering buying a condo in a new building, what characteristics of the builder would make you more likely to buy? Explain. (*Hint*: See the Application "Capping Oil and Gas Bankruptcies.")

4. Checks on Principals

4.1 List as many ways as possible that a principal can reassure an agent that it will avoid opportunistic behavior.

4.2 Should a franchisee fail to meet the terms of a franchise agreement, the franchisor may want to terminate the agreement early to avoid damaging the reputation of the entire franchise system or to recover the trust of its customers. The default and termination provisions of the franchise agreement set out both curable and non-curable acts or omissions and the mechanism for giving notice to the franchisee to effect termination. What effect would a change in the law that restricted the ability of a franchisor to terminate a franchise agreement (for example, by extending the period in which a franchisee may cure a failure) have on production efficiency and risk bearing?

4.3 In the Application "Layoffs Versus Pay Cuts," the firm uses either a pay cut or layoffs. Can you derive a superior approach that benefits both the firm and the workers? (*Hint*: Suppose that the firm's profit or some other variable is observable.)

5. Contract Choice

5.1 List some necessary conditions for a firm to be able to sort potential employees by providing them with a choice of contracts.

5.2 In the contract choice example in the chapter, what are the implications for risk bearing of the fixed-fee and contingent contract if the sales revenue varies with market conditions? Will a worker's attitude toward risk affect which contract the worker chooses?

6. Challenge

6.1 In the Challenge Solution, show that shareholders' expected earnings are higher with the new compensation scheme than with the original one.

6.2 In 2012, Hewlett-Packard Co. announced that its new chief executive, Meg Whitman, would receive a salary of $1 and about $16.1 million in stock options, which are valuable if the stock does well (**marketwatch.com**, February 3, 2012). How would you feel about this compensation package if you were a shareholder? What are the implications for moral hazard, efficiency, and risk sharing?

6.3 Adrienne, a manager of a large firm, must decide whether to launch a new product or make a minor change to an existing product. The new product has a 30% chance of being a big success and generating profits of $20 million, a 40% chance of being fairly successful and generating profits of $5 million, and a 30% chance of being a costly failure and losing $10 million. Making minor changes in the old product would generate profits of $10 million for sure. Adrienne's contract gives her a bonus of 10% of any profits above $8 million arising from this decision. If Adrienne is risk neutral and cares only about her own income, what is her decision? Should shareholders be happy with this compensation contract? Design a contract that would be better for both Adrienne and the shareholders. **A**

6.4 Hannah manages a store for which she receives 20% of any profit earned, so she does well when the store does well and earns nothing if the store incurs a loss. Hannah and the store's owner are both risk neutral. At present, the store is making an annual profit of 500. Would Hannah bring a new product into the store if she expects there is a 60% chance that it would increase the store's profits to 1,000 but a 40% chance that the store would incur a loss of 550? Would the owner support selling the new product? If Hannah's compensation package also included a bonus of 4% of profits, would Hannah choose to sell the new product? **A**

Chapter Appendixes

Appendix 2A: Regressions

An economist's guess is as likely to be as good as anyone else's. —Will Rogers

Economists use a *regression* to estimate economic relationships such as demand curves and supply curves. A regression analysis allows us to answer three types of questions:

- How can we best fit an economic relationship to actual data?
- How confident are we in our results?
- How can we determine the effect of a change in one variable on another if many other variables are changing at the same time?

Estimating Economic Relations

We use a demand curve example to illustrate how regressions can answer these questions. The points in Figure 2A.1 show eight years of data on Nancy's annual purchases of candy bars, q, and the prices, p, she paid.[1] For example, in the year when candy bars cost 20¢, Nancy bought q_2 candy bars.

Because we assume that Nancy's tastes and income did not change during this period, we write her demand for candy bars as a function of the price of candy bars and unobservable random effects. We believe that her demand curve is linear and want to estimate the demand function:

$$q = a + bp + e,$$

where a and b are coefficients we want to determine and e is an error term. This *error term* captures random effects that are not otherwise reflected in our function. For instance, in one year, Nancy broke up with her longtime boyfriend and ate more candy bars than usual, resulting in a relatively large positive error term for that year.

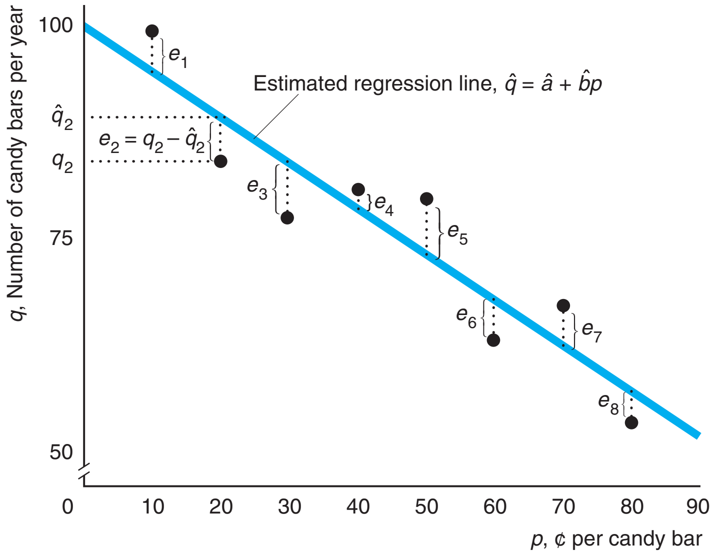
The data points in the figure exhibit a generally downward-sloping relationship between quantity and price, but the points do not lie strictly on a line because of the error terms. We could draw a line through these data points in many possible ways.

---

[1]We use a lowercase q for the quantity demanded for an individual instead of the uppercase Q that we use for a market. Notice that we violated the rule economists usually follow of putting quantity on the horizontal axis and price on the vertical axis. We are now looking at this relationship as statisticians who put the independent or explanatory variable, price, on the horizontal axis and the dependent variable, quantity, on the vertical axis.

Figure 2A.1 Regression

The circles show data on how many candy bars Nancy bought in a year at several different prices. The regression line minimizes the sum of the squared residuals, e_1 through e_8.



We fit the line in the figure by using the standard criterion that our estimates *minimize the sum of squared residuals*, where a residual, $e = q - \hat{q}$, is the difference between an actual quantity, q, and the fitted or predicted quantity on the estimated line, $\hat{q}$. That is, we choose estimated coefficients $\hat{a}$ and $\hat{b}$ so that the estimated quantities from the regression line,

$$\hat{q} = \hat{a} + \hat{b}p,$$

make the sum of the squared residuals, $e_1^2 + e_2^2 + \cdots + e_8^2$, as small as possible. By summing the square of the residuals instead of the residuals themselves, we treat the effects of a positive or negative error symmetrically and give greater weight to large errors than to small ones.[2] In the figure, the regression line is

$$\hat{q} = 99.4 - 0.49p,$$

where $\hat{a} = 99.4$ is the intercept of the estimated line and $\hat{b} = -0.49$ is the slope of the line.

Confidence in Our Estimates

Because the data reflect random errors, so do the estimated coefficients. Our estimate of Nancy's demand curve depends on the *sample* of data we use. If we were to use data from a different set of years, our estimates, $\hat{a}$ and $\hat{b}$, of the true coefficients, a and b, would differ.

[2]Using calculus, we can derive the $\hat{a}$ and $\hat{b}$ that minimize the sum of squared residuals. The estimate of the slope coefficient is a weighted average of the observed quantities, $\hat{b} = \Sigma_i w_i q_i$, where $w_i = (p_i - \overline{p})/\Sigma_i(p_i - \overline{p})^2$, $\overline{p}$ is the average of the observed prices, and Σ_i indicates the sum over each observation i. The estimate of the intercept, $\hat{a}$, is the average of the observed quantities.

If we had many estimates of the true parameter based on many samples, the estimates would be distributed around the true coefficient. These estimates are *unbiased* in the sense that the average of the estimates would equal the true coefficients.

Computer programs that calculate regression lines report a *standard error* for each coefficient, which is an estimate of the dispersion of the estimated coefficients around the true coefficient. In our example, a computer program reports

$$\underset{}{\hat{q}} = \underset{(3.99)}{99.4} - \underset{(0.08)}{0.49}p,$$

where, below each estimated coefficient, its estimated standard error appears between parentheses.

The smaller the estimated standard error, the more precise the estimate, and the more likely it is to be close to the true value. As a rough rule of thumb, there is a 95% probability that the interval that is within two standard errors of the estimated coefficient contains the true coefficient.[3] Using this rule, the *confidence interval* for the slope coefficient, $\hat{b}$, ranges from

$$-0.49 - (2 \times 0.08) = -0.65 \text{ to } -0.49 + (2 \times 0.08) = -0.33.$$

If zero were to lie within the confidence interval for $\hat{b}$, we would conclude that we cannot reject the hypothesis that the price has no effect on the quantity demanded. In our case, however, the entire confidence interval contains negative values, so we are reasonably sure that the higher the price, the less Nancy demands.

Multiple Regression

We can also estimate relationships involving more than one explanatory variable using a *multiple regression*. For example, Moschini and Meilke (1992) estimate a pork demand function, Equation 2.2, in which the quantity demanded is a function of income, Y, and the prices of pork, p, beef, p_b, and chicken, p_c:

$$Q = 171 - 20p + 20p_b + 3p_c + 2Y.$$

The multiple regression is able to separate the effects of the various explanatory variables. The coefficient 20 on the p variable says that an increase in the price of pork by \$1 per kg lowers the quantity demanded by 20 million kg per year, holding the effects of the other prices and income constant.

Appendix 3A: Effects of a Specific Tax on Equilibrium

Given that the government collects a specific or unit tax, t, from sellers, sellers receive $p - t$ when consumers pay p. We can use this information to determine the effect of a new tax on the equilibrium. In the post-tax equilibrium, the equality between the demand function and the after-tax supply function determines the price that consumers pay:

$$D(p) - S(p - t) = 0. \tag{3A.1}$$

[3]The confidence interval is the coefficient plus or minus 1.96 times its standard error for large samples (at least hundreds of observations) in which the coefficients are normally distributed. For smaller samples, the confidence interval tends to be larger.

where the supply equals demand equation is written in implicit function form (the right side of the equation is zero). That is, this equation implicitly defines the price as a function of t: $p(t)$.

We determine the effect a small tax has on the price by differentiating Equation 3A.1 with respect to t:

$$\frac{\mathrm{d}D}{\mathrm{d}p}\frac{\mathrm{d}p}{\mathrm{d}t} - \frac{\mathrm{d}S}{\mathrm{d}p}\frac{\mathrm{d}(p(t) - t)}{\mathrm{d}t} = \frac{\mathrm{d}D}{\mathrm{d}p}\frac{\mathrm{d}p}{\mathrm{d}t} - \frac{\mathrm{d}S}{\mathrm{d}p}\left(\frac{\mathrm{d}p}{\mathrm{d}t} - 1\right) = 0.$$

Rearranging the terms, it follows that the change in the price that consumers pay with respect to the change in the tax is

$$\frac{\mathrm{d}p}{\mathrm{d}t} = \frac{\dfrac{\mathrm{d}S}{\mathrm{d}p}}{\dfrac{\mathrm{d}S}{\mathrm{d}p} - \dfrac{\mathrm{d}D}{\mathrm{d}p}}. \tag{3A.2}$$

We know that $\mathrm{d}D/\mathrm{d}p < 0$ from the Law of Demand. If the supply curve slopes upward so that $\mathrm{d}S/\mathrm{d}p > 0$, then $\mathrm{d}p/\mathrm{d}t > 0$. The higher the tax, the greater the price consumers pay. If $\mathrm{d}S/\mathrm{d}p < 0$, the direction of change is ambiguous: It depends on the relative slopes of the supply and demand curves (the denominator).

By multiplying both the numerator and denominator of the right side of Equation 3A.2 by p/Q, we can express this derivative in terms of elasticities,

$$\frac{\mathrm{d}p}{\mathrm{d}t} = \frac{\dfrac{\mathrm{d}S}{\mathrm{d}p}\dfrac{p}{Q}}{\dfrac{\mathrm{d}S}{\mathrm{d}p}\dfrac{p}{Q} - \dfrac{\mathrm{d}D}{\mathrm{d}p}\dfrac{p}{Q}} = \frac{\eta}{\eta - \quad}, \tag{3A.3}$$

where the last equality follows because $\mathrm{d}S/\mathrm{d}p$ and $\mathrm{d}D/\mathrm{d}p$ are the changes in the quantities supplied and demanded as price changes, and the consumer and producer prices are identical when $t = 0$. That is, for small changes in the tax rate, Δt, the change in price is $\Delta p = [\eta/(\eta - \quad)]\Delta t$.

To determine the effect on quantity, we can combine the price result from Equation 3A.3 with information from either the demand or the supply function. Differentiating the demand function with respect to t, we know that

$$\frac{\mathrm{d}D}{\mathrm{d}p}\frac{\mathrm{d}p}{\mathrm{d}t} = \frac{\mathrm{d}D}{\mathrm{d}p}\frac{\eta}{\eta - \quad},$$

which is negative if the supply curve is upward sloping so that $\eta > 0$.

Appendix 4A: Utility and Indifference Curves

We now use calculus to examine the relationship between utility and indifference curves and some properties of indifference curves. Suppose that Lisa's utility function is $U(B, Z)$, where B is the number of burritos and Z is the number of pizzas. Lisa's marginal utility for burritos, MU_B, is the amount of extra pleasure she would get from extra burritos, holding her consumption of pizza constant. Formally, her marginal utility for burritos, B, is the partial derivative of utility, $U(B, Z)$, with respect to B holding Z constant:

$$MU_B(B, Z) = \lim_{\Delta B \to 0} \frac{U(B + \Delta B, Z) - U(B, Z)}{\Delta B} = \frac{\partial U(B, Z)}{\partial B}.$$

By assumption, marginal utility is always nonnegative: A little more of a good makes you better off or at least doesn't harm you. The marginal utility depends on the current levels of B and Z.

Which combinations of B and Z leave Lisa with a given level of pleasure, say, $\overline{U}$? We can write those combinations as

$$\overline{U} = U(B, Z). \tag{4A.1}$$

Equation 4A.1 is the equation for an indifference curve with utility level $\overline{U}$.

We can express the slope of an indifference curve—the marginal rate of substitution, *MRS*—in terms of the marginal utilities. We find the slope of the indifference curve by determining the changes in B and P that leave utility unchanged. Totally differentiating Equation 4A.1, we find that

$$\mathrm{d}\overline{U} = 0 = \frac{\partial U(B, Z)}{\partial B}\mathrm{d}B + \frac{\partial U(B, Z)}{\partial Z}\mathrm{d}Z \equiv MU_B\mathrm{d}B + MU_Z\mathrm{d}Z \tag{4A.2}$$

This equation says that a little extra utility, MU_B, times the change in B, $\mathrm{d}B$, plus the extra utility, MU_Z, times the change in Z, $\mathrm{d}Z$, must add to zero. If we increase one of the goods, we must decrease the other to hold utility constant so that we stay on the same indifference curve. In Equation 4A.2, $\mathrm{d}\overline{U} = 0$ because we are holding utility constant so that we stay on the same indifference curve. Rearranging the terms in Equation 4A.2, we find that

$$\frac{\mathrm{d}B}{\mathrm{d}Z} = -\frac{MU_Z}{MU_B}.$$

The slope of the indifference curve is the negative of the ratio of the marginal utilities.

Suppose that Lisa has the following utility function, known as a *Cobb-Douglas utility function*:

$$U(B, Z) = AB^aZ^b. \tag{4A.3}$$

Her marginal utility of burritos is

$$MU_B(B, Z) = aAB^{a-1}Z^b = a\frac{U(B, Z)}{B},$$

and her marginal utility of pizza is

$$MU_Z(B, Z) = \beta AB^aZ^{b-1} = b\frac{U(B, Z)}{Z}.$$

Suppose that $a = b = \frac{1}{2}$ and $A = 20$. If $B = Z = 4$, then $U(4, 4) = 80$ and $MU_B(4, 4) = MU_Z(4, 4) = 10$. If $B = 1$ and $Z = 4$, however, $U(1, 4) = 40$, $MU_B(1, 4) = 20$, and $MU_Z(1, 4) = 5$. The extra pleasure that Lisa gets from an extra burrito is greater, the fewer burritos she initially has, all else the same.

The slope of her indifference curve is

$$MRS = \frac{dB}{dZ} = -\frac{MU_Z}{MU_B} = -\frac{bAB^a Z^{b-1}}{aAB^{a-1}Z^b} = -\frac{bB}{aZ}.$$

The slope of the indifference curve differs with the levels of B and Z. If $a = b = \frac{1}{2}$, $B = 4$, and $Z = 1$, $MRS(4, 1) = -\left(\frac{1}{2} \times 4\right)/\left(\frac{1}{2} \times 1\right) = -4$. At $B = Z = 4$, $MRS(4, 4) = -1$.

Appendix 4B: Maximizing Utility

Lisa's objective is to maximize her *utility, U(B, Z), subject to (s.t.) a budget constraint:*

$$\max_{B,\,Z} U(B, Z)$$
$$\text{s.t.} \quad Y = p_B B + p_Z Z, \tag{4B.1}$$

where B is the number of burritos she buys at price p_B, Z is the number of pizzas she buys at price p_Z, Y is her income, and $Y = p_B B + p_Z Z$ is her budget constraint (her spending on burritos and pizza can't exceed her income). The mathematical statement of her problem shows that her *control variables* (what she chooses) are B and Z, which appear under the "max" term in the equation. We assume that Lisa has no control over the prices she faces or her budget.

To solve this type of constrained maximization problem, we use the Lagrangian method:

$$\max_{B,\,Z,\,\lambda} \mathcal{L} = U(B, Z) - \lambda(p_B B + p_Z Z - Y), \tag{4B.2}$$

where λ is called the Lagrangian multiplier. With normal-shaped utility functions, the values of B, Z, and λ that are determined by the first-order conditions of this Lagrangian problem are the same as the values that maximize the original constrained problem. The first-order conditions of Equation 4B.2 with respect to the three control variables, B, Z, and λ, are:[4]

$$\frac{\partial \mathcal{L}}{\partial B} = MU_B(B, Z) - \lambda p_B = 0, \tag{4B.3}$$

$$\frac{\partial \mathcal{L}}{\partial Z} = MU_Z(B, Z) - \lambda p_Z = 0, \tag{4B.4}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = Y - p_B B - p_Z Z = 0, \tag{4B.5}$$

where $MU_B(B, Z) \equiv \partial U(B, Z)/\partial B$ is the partial derivative of utility with respect to B (the marginal utility of B) and $MU_Z(B, Z)$ is the marginal utility of Z. Equation 4B.5 is the budget constraint. Equations 4B.3 and 4B.4 say that the marginal utility of each good equals its price times λ.

---

[4]For simplicity, we assume that we have an interior solution, B and Z are infinitely divisible, and $U(B, Z)$ is continuously differentiable at least twice (so that the second-order condition is well defined). The first-order conditions give us the necessary conditions for an interior solution in which positive quantities of both goods are consumed. We assume that the second-order (sufficient) conditions hold, which is true if the utility function is quasiconcave or if the indifference curves are convex to the origin. That is, Lisa is maximizing rather than minimizing her utility when she chooses the levels of B and Z given by the first-order conditions.

What is λ? If we equate Equations 4B.3 and 4B.4 and rearrange terms, we find that

$$\lambda = \frac{MU_B}{p_B} = \frac{MU_Z}{p_Z}. \tag{4B.6}$$

Because the Lagrangian multiplier, λ, equals the marginal utility of each good divided by its price, λ equals the extra pleasure one gets from one's last dollar of expenditures. Equivalently, λ is the value of loosening the budget constraint by one dollar.[5] Equation 4B.6 tells us that for Lisa to maximize her utility, she should pick a B and Z so that if she got one more dollar, spending that dollar on B or on Z would give her the same extra utility.

An alternative interpretation of this condition for maximizing utility. Using the ratio of Equations 4B.3 and 4B.4 (or rearranging 4B.6), we find that

$$\frac{MU_Z}{MU_B} = \frac{p_Z}{p_B}. \tag{4B.7}$$

The left side of Equation 4B.7 is the absolute value of the marginal rate of substitution, $MRS = -MU_Z/MU_B$, and the right side is the absolute value of the marginal rate of transformation, $MRT = -p_Z/p_B$. Thus, the calculus approach gives us the same condition for an optimum that we derived using graphs. The indifference curve should be tangent to the budget constraint: The slope of the indifference curve, MRS, should equal the slope of the budget constraint, MRT.

For example, suppose that the utility is Cobb-Douglas, as in Equation 4A.3: $U = AB^aZ^b$. The first-order condition, Equation 4B.5, the budget constraint, stays the same, and Equations 4B.3 and 4B.4 become

$$\frac{\partial \mathcal{L}}{\partial B} = a\frac{U(B, Z)}{B} - \lambda p_B = 0, \tag{4B.8}$$

$$\frac{\partial \mathcal{L}}{\partial Z} = b\frac{U(B, Z)}{Z} = \lambda p_Z = 0. \tag{4B.9}$$

Using Equations 4B.8 and 4B.9, we can write Equation 4B.6 as

$$\lambda = a\frac{U(B, Z)}{p_B B} = b\frac{U(B, Z)}{p_Z Z}.$$

Taking the ratio of Equations 4B.8 and 4B.9 and rearranging terms, we find that

$$bp_B B = ap_Z Z. \tag{4B.10}$$

---

[5]Differentiating utility with respect to Y, we find that

$$\frac{dU}{dY} = MU_B(B, Z)\frac{dB}{dY} + MU_Z(B, Z)\frac{dZ}{dY}.$$

Substituting from Equation 4B.6 into this expression, we obtain

$$\frac{dU}{dY} = \lambda p_B\frac{dB}{dY} + \lambda p_Z\frac{dZ}{dY} = \lambda\frac{p_B\,dB + p_Z\,dZ}{dY}.$$

Totally differentiating the budget constraint, we learn that

$$dY = p_B\,dB + p_Z dZ.$$

Substituting this expression into the previous expression gives us

$$\frac{dU}{dY} = \frac{\lambda p_B\,dB + \lambda p_Z\,dZ}{p_B\,dB + p_Z\,dZ} = \lambda.$$

Thus, λ equals the extra utility one gets from one more dollar of income.

Substituting $Y - p_B B$ for $p_Z Z$, using Equation 4B.5, into Equation 4B.10 and rearranging terms, we get

$$B = \frac{a}{a+b}\frac{Y}{p_B}. \tag{4B.11}$$

Similarly, by substituting Equation 4B.11 into Equation 4B.10, we find that

$$Z = \frac{b}{a+b}\frac{Y}{p_Z}. \tag{4B.12}$$

Thus, knowing the utility function, we can solve the expression for the B and Z that maximize utility in terms of income and prices.

Equations 4B.11 and 4B.12 are the consumer's demand curves for B and Z, respectively. (We derive demand curves using graphs in Chapter 5.)

If $a = b = \frac{1}{2}$, $A = 20$, $Y = 80$, and $p_Z = p_B = 10$, then $B = Z = 4$ and the value of loosening the budget constraint is $\lambda = MU_B/p_B = MU_Z/p_Z = 10/10 = 1$. If p_B rises to 40, then $Z = 4$, $B = 1$, and $\lambda = 20/40 = 5/10 = \frac{1}{2}$.

Appendix 5A: The Slutsky Equation

The total effect on the quantity demanded when the price of a good rises equals the sum of the substitution and income effects. The Slutsky equation (named after its discoverer, the Russian economist Eugene Slutsky) explicitly shows the relationship among the price elasticity of demand, , the pure substitution elasticity of demand, *, and the income elasticity of demand, ξ:

$$\begin{aligned} \text{Total effect} &= \text{substitution effect} + \text{income effect} \\ &= \qquad\quad {}^{*} \qquad\quad + \quad (-\theta\xi) \end{aligned}$$

where θ is the budget share of this good: the amount spent on this good divided by the total budget.

We now sketch the derivation of the Slutsky equation (for a formal derivation, see a graduate microeconomics textbook such as Varian, 1992). The total effect, $\Delta q/\Delta p$, is the change in the quantity demanded, Δq, for a given change in the good's price, Δp. The substitution effect is the change in quantity demanded for a change in price, holding utility constant, which we label $(\Delta q/\Delta p)_{U\ constant}$.

A change in the price affects how much the consumer can buy and acts like a change in income. The income effect is the change in quantity as income changes times the change in income as price changes, $(\Delta q/\Delta Y)(\Delta Y/\Delta p)$, where ΔY is the change in income. The change in income from a change in price is $\Delta Y/\Delta p = -q$. For example, if price rises by \$1, income falls by the number of units purchased. From this last result, the income effect is $-q(\Delta q/\Delta Y)$.

Using these expressions, we write the identity that the total effect equals the substitution plus the income effect as

$$\Delta q/\Delta p = (\Delta q/\Delta p)_{U\ constant} - q(\Delta q/\Delta Y).$$

Multiplying this equation through by p/q, multiplying the last term by Y/Y, and rearranging terms, we obtain

$$\frac{\Delta q}{\Delta p}\frac{p}{q} = \left(\frac{\Delta q}{\Delta p}\right)_{U\ constant}\frac{p}{q} - \frac{\Delta q}{\Delta Y}\frac{Y}{q}\frac{pq}{Y}.$$

Substituting $= (\Delta q/\Delta p)(p/q)$, $^* = (\Delta q/\Delta p)_{U\ constant}(p/q)$, $\xi = (\Delta q/\Delta Y)(Y/q)$, and $\theta = pq/Y$ into this last expression, we have the Slutsky equation:

$$= \ ^* - \theta\xi.$$

Appendix 5B: Labor-Leisure Model

Jackie's utility, U, is a function,

$$U = U(Y, N), \tag{5B.1}$$

of her leisure, N, and her income, Y, which she uses to buy all other goods and services. Jackie maximizes her utility, Equation 5B.1, subject to two constraints. The first, imposed by the clock, is that the number of hours she works, H, equals her total hours in a day minus her hours of leisure:

$$H = 24 - N. \tag{5B.2}$$

The second constraint is that her earned income (earnings), Y, equals her wage, w, times the hours she works:

$$Y = wH. \tag{5B.3}$$

For now, we assume that her unearned income is zero.

Although we can maximize Equation 5B.1 subject to Equations 5B.2 and 5B.3 using Lagrangian techniques, it is easier to do so by substitution. By substituting Equations 5B.2 and 5B.3 into Equation 5B.1, we can convert this constrained problem into an unconstrained maximization problem:

$$\max_H U = U(wH, 24 - H). \tag{5B.4}$$

By using the chain rule of differentiation, we find that the first-order condition for an interior maximum to the problem in Equation 5B.4 is

$$\frac{\mathrm{d}U}{\mathrm{d}H} = wMU_Y - MU_N = 0,$$

where MU_Y, the marginal utility of goods or income, is the partial derivative of utility with respect to income, $\partial U/\partial Y$, and MU_N, the marginal utility of leisure, is the partial derivative with respect to leisure, $\partial U/\partial N$.[6] This expression can be rewritten as $w = MU_N/MU_Y$.

To maximize her utility, Jackie must set her marginal rate of substitution of income for leisure, $MRS = -MU_N/MU_Y$, equal to her marginal rate of transformation of income for leisure, $MRT = -w$, in the market:

$$MRS = -\frac{MU_N}{MU_Y} = -w = MRT.$$

Suppose that Jackie's utility is

$$U = Y^a N^{1-a} = (wH)^a(24 - H)^{1-a},$$

which is a Cobb-Douglas utility function (Appendix 4A). Differentiating this utility function with respect to H, setting the derivative equal to zero, and rearranging

---

[6]The second-order condition for an interior maximum is

$$\frac{\mathrm{d}^2U}{\mathrm{d}H^2} = \frac{\partial^2 U}{\partial Y^2}w^2 - 2\frac{\partial^2 U}{\partial Y\,\partial N}w + \frac{\partial^2 U}{\partial N^2} < 0.$$

terms, we find that $H = 24a$. With this particular utility function, an individual's hours of leisure and work are fixed regardless of the wage. If $\alpha = \frac{1}{2}$, the individual works 12 hours a day (and has 12 hours of leisure) whether the wage is 50¢ an hour or $500 an hour.

Appendix 6A: Properties of Marginal and Average Product Curves

We can use calculus to show that the MP_L curve crosses the AP_L curve at its peak. Because capital is fixed, we can write the production function solely in terms of labor: $q = f(L)$. In Figure 6.1, $MP_L = \mathrm{d}q/\mathrm{d}L = \mathrm{d}f/\mathrm{d}L > 0$, $\mathrm{d}^2f/\mathrm{d}L^2 < 0$, and $AP_L = q/L = f(L)/L > 0$. A necessary condition to identify the amount of labor where the AP_L curve reaches a maximum is that the derivative of AP_L with respect to L equals zero:

$$\frac{\mathrm{d}AP_L}{\mathrm{d}L} = \left(\frac{\mathrm{d}q}{\mathrm{d}L} - \frac{q}{L}\right)\frac{1}{L} = 0.$$

(At the L determined by this first-order condition, AP_L is maximized if the second-order condition is negative: $\mathrm{d}^2AP_L/\mathrm{d}L^2 = \mathrm{d}^2f/\mathrm{d}L^2 < 0$.) By rearranging this first-order condition, $MP_L = \mathrm{d}q/\mathrm{d}L = q/L = AP_L$ at the peak of the AP_L curve.

Appendix 6B: The Slope of an Isoquant

We can use calculus to determine the slope at a point on an isoquant. We totally differentiate the isoquant, $\overline{q} = f(L, K)$, with respect to L and K. Along the isoquant, we can write capital as an implicit function of labor: $K(L)$. That is, for a given quantity of labor, the firm needs a certain level of capital to produce $\overline{q}$ units. Differentiating with respect to labor (and realizing that output does not change along the isoquant as we change labor), we have

$$\frac{\mathrm{d}\overline{q}}{\mathrm{d}L} = 0 = \frac{\partial f}{\partial L} + \frac{\partial f}{\partial K}\frac{\mathrm{d}K}{\mathrm{d}L} = MP_L + MP_K\frac{\mathrm{d}K}{\mathrm{d}L},$$

where $MP_K = \partial f/\partial K$ is the marginal product of capital. Rearranging this expression, we find that $-MP_L/MP_K = \mathrm{d}K/\mathrm{d}L = MRTS$.

Appendix 6C: Cobb-Douglas Production Function

The Cobb-Douglas production function is

$$q = AL^aK^b. \tag{6C.1}$$

Economists use statistical means to estimate A, α, and β, which determine the exact shape of the production function. The larger A is, the more output the firm gets from a given amount of labor and capital.

The average product of labor is determined by dividing both sides of Equation 6C.1 by q:

$$AP_L = q/L = AL^aK^b/L = AL^{a-1}K^b. \tag{6C.2}$$

The α term tells us the relationship between the average product of labor and the marginal product of labor. By differentiating the Cobb-Douglas production function with respect to L, holding K constant, we find that the marginal product of labor is

$$MP_L = \frac{\partial q}{\partial L} = aAL^{a-1}K^b = a\frac{AL^aK^b}{L} = a\frac{q}{L}.$$

The marginal product of labor equals α times the average product of labor: $MP_L = \alpha AP_L$. Consequently, $\alpha = AP_L/MP_L$. Using similar reasoning, the marginal product of capital is $MP_K = \beta q/K$. As Equation 6.7 shows, the marginal rate of technical substitution is $MRTS = -MP_L/MP_K = -(aq/L)/(bq/K) = -(a/b)K/L$.

The change in the average product of labor as labor increases is $\partial AP_L/\partial L = (a - 1)AL^{a-2}K^b = (a - 1)q/L^2$. If $\alpha - 1 < 0$ (that is, $a < 1$), then the change in the average product of labor as the number of workers increases is negative.

Appendix 7A: Minimum of the Average Cost Curve

To determine the output level q where the average cost curve, $AC(q)$, reaches its minimum, we set the derivative of average cost with respect to q equal to zero:

$$\frac{\mathrm{d}AC(q)}{\mathrm{d}q} = \frac{\mathrm{d}(C(q)/q)}{\mathrm{d}q} = \left(\frac{\mathrm{d}C(q)}{\mathrm{d}q} - \frac{C(q)}{q}\right)\frac{1}{q} = 0.$$

This condition holds at the output q where $\mathrm{d}C(q)/\mathrm{d}q = C(q)/q$, or $MC = AC$. If the second-order condition holds at that q, the average cost curve reaches its minimum at that quantity. The second-order condition requires that the average cost curve be falling to the left of this q and rising to the right.

Appendix 7B: Japanese Beer Manufacturer's Short-Run Cost Curves

We can use math to derive the various short-run cost curves for a Japanese beer manufacturer. Based on the estimates of Flath (2011), its production function is

$$q = 1.52L^{0.6}K^{0.4},$$

where labor, L, is measured in hours, K is the number of units of capital, and q is the amount of output. (*Note*: The coefficient 1.52 was chosen to produce round numbers.)

In the short run, the firm's capital is fixed at $\overline{K} = 100$. In the figure in Chapter 7's Application "A Beer Manufacturer's Short-Run Cost Curves," the rental rate of a unit of capital is \$8 and the fixed cost, F, is \$800. The figure shows that the average fixed cost, $AFC = F/q = 800/q$, falls as output increases.

We can use the production function to derive the variable cost. First, we determine how output and labor are related. Setting capital, K, at 100 units in the production function, we find that the output produced in the short run is solely a function of labor:

$$q = 1.52L^{0.6}100^{0.4} \approx 9.59L^{0.6}.$$

Rearranging this expression, we can write the number of workers per year, L, needed to produce q units of output, as a function solely of output:

$$L(q) = \left(\frac{q}{1.52 \times 100^{0.4}}\right)^{\frac{1}{0.6}} \approx 0.023q^{1.67}. \tag{7B.1}$$

Now that we know how labor and output are related, we can calculate variable cost directly. The only variable input is labor, so if the wage is \$24, the firm's variable cost is $VC(q) = wL(q) = 24L(q)$. Substituting for $L(q)$ using Equation 7B.1, we see how variable cost varies with output:

$$VC(q) = 24L(q) = 24\left(\frac{q}{1.52 \times 100^{0.4}}\right)^{\frac{1}{0.6}} \approx 0.55q^{1.67}. \tag{7B.2}$$

Using this expression for variable cost, we can construct the other cost measures.

We obtain the average variable cost as a function of output, $AVC(q)$, by dividing both sides of Equation 7B.2 by q:

$$AVC(q) = \frac{VC(q)}{q} = \frac{24L(q)}{q} \approx 24\left(\frac{0.023q^{1.67}}{q}\right) = 0.55q^{0.67}.$$

As the figure in the Application shows, the average variable cost is strictly increasing.

To obtain the equation for marginal cost as a function of output, we differentiate the variable cost, $VC(q)$, with respect to output:

$$MC(q) = \frac{dVC(q)}{dq} \approx \frac{d(0.55q^{1.67})}{dq} = 1.67 \times 0.55q^{0.67} \approx 0.92q^{0.67}.$$

Thus, to construct all the cost measures of the beer manufacturer, we need only the production function and the prices of the inputs.

Appendix 7C: Minimizing Cost

We can use calculus to derive the cost minimization conditions, Equations 7.5 and 7.8, discussed in the chapter. The problem the firm faces in the long run is to choose the level of labor, L, and capital, K, that will minimize the cost of producing a particular level of output, $\bar{q}$, given a wage of w and a rental rate of capital of r.

The relationship between inputs and output is summarized in the firm's production function: $q = f(L, K)$. The marginal product of labor, which is the extra output the firm produces from a little more labor, holding capital constant, is $MP_L(L, K) = \partial f(L, K)/\partial L$, which is positive. This has diminishing marginal returns to labor, however, so the marginal product of labor falls as labor increases: $\partial MP_L(L, K)/\partial L = \partial^2 f(L, K)/\partial L^2 < 0$. The marginal product of capital has the same properties: $\partial f(L, K)/\partial K > 0$ and $\partial MP_K(L, K)/\partial K < 0$.

The firm's problem is to minimize its cost, C, of production, through its choice of labor and capital,

$$\min_{L,\, K} C = wL + rK,$$

subject to the constraint that a given amount of output, $\overline{q}$, is to be produced:

$$f(L, K) = \overline{q}. \tag{7C.1}$$

Equation 7C.1 is the $\overline{q}$ isoquant.

We can change this constrained minimization problem into an unconstrained problem by using the Lagrangian technique. The firm's unconstrained problem is to minimize the Lagrangian, $\mathcal{L}$, through its choice of labor, capital, and the Lagrangian multiplier, λ:

$$\min_{L,\, K,\, \lambda} \mathcal{L} = wL + rK - \lambda\big(f(L, K) - \overline{q}\big).$$

We obtain the necessary conditions for a minimum by differentiating $\mathcal{L}$ with respect to L, K, and λ and setting the derivatives equal to zero:

$$\partial\mathcal{L}/\partial L = w - \lambda MP_L(L, K) = 0, \tag{7C.2}$$

$$\partial\mathcal{L}/\partial K = r - \lambda MP_K(L, K) = 0, \tag{7C.3}$$

$$\partial\mathcal{L}/\partial\lambda = f(L, K) - \overline{q} = 0. \tag{7C.4}$$

We can rewrite Equations 7C.2 and 7C.3 as $w = \lambda MP_L(L, K)$ and $r = \lambda MP_K(L, K)$. Taking the ratio of these two expressions, we obtain

$$\frac{w}{r} = \frac{MP_L(L, K)}{MP_K(L, K)} = -MRTS, \tag{7C.5}$$

which is the same as Equation 7.5. This condition states that cost is minimized when the rate at which firms can exchange capital for labor in the market, w/r, is the same as the rate at which capital can be substituted for labor along an isoquant. That is, the isocost line is tangent to the isoquant.

We can rewrite Equation 7C.5 to obtain the expression

$$\frac{MP_L(L, K)}{w} = \frac{MP_K(L, K)}{r}.$$

This equation tells us that the last dollar spent on labor should produce as much extra output as the last dollar spent on capital; otherwise, the amount of factors used should be adjusted.

We can rearrange Equations 7C.2 and 7C.3 to obtain an expression for the Lagrangian multiplier:

$$\lambda = \frac{w}{MP_L(L, K)} = \frac{r}{MP_K(L, K)}. \tag{7C.6}$$

Equation 7C.6 says that the Lagrangian multiplier, λ, equals the ratio of the factor price to the marginal product for each factor. The marginal product for a factor is the extra amount of output one gets by increasing that factor slightly, so the reciprocal of the marginal product is the extra input it takes to produce an extra unit of output. By multiplying the reciprocal of the marginal product by the factor cost, we learn the extra cost of producing an extra unit of output by using more of this factor. Thus, the Lagrangian multiplier equals the marginal cost of production: It measures how much the cost increases if we produce one more unit of output.

If a firm has a Cobb-Douglas production function, $Q = AL^aK^b$, the marginal product of capital is $MP_K = bq/K$ and the marginal product of labor is $MP_L = aq/L$ (see Appendix 6C), so the $MRTS$ is $aK/(bL)$. Thus, the tangency condition, Equation 7C.5, requires that

$$\frac{w}{r} = \frac{aK}{bL}. \tag{7C.7}$$

Using algebra, we can rewrite Equation 7C.7 as

$$K = \frac{bw}{ar}L, \tag{7C.8}$$

which is the expansion path for a Cobb-Douglas production function given w and r. According to Equation 7C.8, the expansion path of a firm with a Cobb-Douglas production function is an upward-sloping straight line through the origin with a slope of $bw/(ar)$.

Appendix 8A: The Elasticity of the Residual Demand Curve

Here we derive the expression for the elasticity of the residual demand curve given in Equation 8.2. Differentiating the residual demand (Equation 8.1),

$$D^r(p) = D(p) - S^o(p),$$

with respect to p, we obtain

$$\frac{\mathrm{d}D^r}{\mathrm{d}p} = \frac{\mathrm{d}D}{\mathrm{d}p} - \frac{\mathrm{d}S^o}{\mathrm{d}p}.$$

Because the firms are identical, the quantity produced by each is $q = Q/n$, and the total quantity produced by all the other firms is $Q_o = (n - 1)q$. Multiplying both sides of the expression by p/q and multiplying and dividing the first term on the right side by Q/Q and the second term by Q_o/Q_o, this expression may be rewritten as

$$\frac{\mathrm{d}D^r}{\mathrm{d}p}\frac{p}{q} = \frac{\mathrm{d}D}{\mathrm{d}p}\frac{p}{Q}\frac{Q}{q} - \frac{\mathrm{d}S^o}{\mathrm{d}p}\frac{p}{Q_o}\frac{Q_o}{q},$$

where $q = D^r(p)$, $Q = D(p)$, and $Q_o = S^o(p)$. This expression can in turn be rewritten as Equation 8.2,

$$_i = n \quad - (n - 1)\eta_o,$$

by noting that $Q/q = n$, $Q_o/q = (n - 1)$, $(\mathrm{d}D^r/\mathrm{d}p)(p/q) = \ _i$, $(\mathrm{d}D/\mathrm{d}p)(p/Q) = \ $, and $(\mathrm{d}S^o/\mathrm{d}p)(p/Q_o) = \eta_o$.

Appendix 8B: Profit Maximization

In general, a firm maximizes its profit, $\pi(q) = R(q) - C(q)$, by its choice of output q. We determine a *necessary condition* for a maximum at a positive level of output by differentiating profit with respect to q and setting the derivative equal to zero:

$$\frac{d\pi}{dq} = \frac{dR(q^*)}{dq} - \frac{dC(q^*)}{dq} = 0, \tag{8B.1}$$

where q^* is the profit-maximizing output. Because $dR(q)/dq$ is the marginal revenue, $MR(q)$, and $dC(q)/dq$ is the marginal cost, $MC(q)$, Equation 8B.1 says that marginal revenue equals marginal cost at q^*:

$$MR(q^*) = MC(q^*). \tag{8B.2}$$

A *sufficient condition* for profit to be maximized at $q^* > 0$ is that the second-order condition holds:

$$\frac{d^2\pi}{dq^2} = \frac{d^2R(q^*)}{dq^2} - \frac{d^2C(q^*)}{dq^2} = \frac{dMR(q^*)}{dq} - \frac{dMC(q^*)}{dq} < 0. \tag{8B.3}$$

We can rewrite Equation 8B.3 as

$$\frac{dMR(q^*)}{dq} < \frac{dMC(q^*)}{dq}. \tag{8B.4}$$

Thus, a sufficient condition for a maximum is that the slope of the marginal revenue curve is less than that of the marginal cost curve and that the MC curve cuts the MR curve from below at q^*.

For a competitive firm, $\pi(q) = pq - C(q)$, so the necessary condition for profit to be maximized, Equation 8B.1 or 8B.2, is

$$p = MC(q^*). \tag{8B.5}$$

Equation 8B.5 says that a profit-maximizing, competitive firm sets its output at q^*, where its marginal cost equals its price.

Because a competitive firm's marginal revenue, p, is a constant, $dMR/dq = dp/dq = 0$. Thus, the sufficient condition for profit to be maximized, Equation 8B.4, can be rewritten as

$$0 < \frac{dMC(q^*)}{dq} \tag{8B.6}$$

for a competitive firm. Equation 8B.6 shows that a sufficient condition for a competitive firm to be maximizing its profit at q^* is that its marginal cost curve is upward sloping at the equilibrium quantity.

Appendix 9A: Demand Elasticities and Surplus

If the demand curve is linear, as in Figure 9.3, the lost consumer surplus, area $B + C$, equals the sum of the area of a rectangle, $Q\Delta p$, with length Q and height Δp, plus the area of a triangle, $\frac{1}{2}\Delta Q\Delta p$, of length ΔQ and height Δp. We can approximate any demand curve with a straight line, so that $\Delta CS = Q\Delta p + \frac{1}{2}\Delta Q\Delta p$ is a reasonable approximation to the true change in consumer surplus. We can rewrite this expression for ΔCS as

$$\Delta p(Q + \tfrac{1}{2}\Delta Q) = Q\Delta p\left[1 + \tfrac{1}{2}\left(\frac{\Delta Q}{Q}\,\frac{p}{\Delta p}\right)\frac{\Delta p}{p}\right]$$

$$= (pQ)\frac{\Delta p}{p}\left(1 + \tfrac{1}{2}\ \frac{\Delta p}{p}\right)$$

$$= Rx\left(1 + \tfrac{1}{2}\ x\right),$$

where $x = \Delta p/p$ is the percentage increase in the price, $R(= pQ)$ is the total revenue from the sale of good Q, and   is the elasticity of demand. (We use this equation to calculate the last column in the table in the Application "Goods with a Large Consumer Surplus Loss from Price Increases.")

Appendix 11A: Relationship Between a Linear Demand Curve and Its Marginal Revenue Curve

When the demand curve is linear, its marginal revenue curve is twice as steep and hits the horizontal axis at half the quantity of the demand curve. A linear demand curve can be written generally as $p = a - bQ$. The monopoly's revenues are quadratic, $R = pQ = aQ - bQ^2$. Differentiating revenue with respect to quantity, we find that the marginal revenue, $\mathrm{d}R(Q)/\mathrm{d}Q$, is linear, $MR = a - 2bQ$. The demand and MR curves hit the price axis at a. The slope of the demand curve, $\mathrm{d}p/\mathrm{d}Q = -b$, is half (in absolute value) the slope of the marginal revenue curve, $\mathrm{d}MR/\mathrm{d}Q = -2b$. The MR curve hits the quantity axis at half the distance, $a/(2b)$, of the demand curve, a/b.

Appendix 11B: Incidence of a Specific Tax on a Monopoly

In a monopolized market, the incidence of a specific tax falling on consumers can exceed 100%: The price may rise by an amount greater than the tax. To demonstrate this possibility, we examine a market where the demand curve has a constant elasticity of   and the marginal cost is constant at $MC = m$.

Suppose that the inverse demand curve the monopoly faces is

$$p = Q^{1/\ }. \tag{11B.1}$$

The monopoly's revenue is $R = pQ = Q^{1+1/\ }$. By differentiating, we learn that the monopoly's marginal revenue is $MR = (1 + 1/\ )Q^{1/\ }$.

To maximize its profit, the monopoly operates where its marginal revenue equals its marginal cost:

$$MR = (1 + 1/\ )Q^{1/\ } = m = MC.$$

Solving this equation for the profit-maximizing output, we find that $Q = [m/(1 + 1/\ )]$. Substituting that value of Q into Equation 11B.1, we find that

$$p = m/(1 + 1/\ ).$$

A specific tax of t per unit raises the marginal cost to $m + t$, so that the monopoly price increases to

$$p_t = (m + t)/(1 + 1/\ ).$$

Consequently, the increase in price is $t/(1 + 1/\ )$. The incidence of the tax that falls on consumers is $\Delta p/\Delta t = [t/(1 + 1/\ )]/t = 1/(1 + 1/\ ) > 1$, because $\ < -1$ (a monopoly never operates in the inelastic portion of its demand curve).

Appendix 12A: Perfect Price Discrimination

A perfectly price-discriminating monopoly charges each customer the reservation price $p = D(Q)$, where $D(Q)$ is the inverse demand function and Q is total output. The discriminating monopoly's revenue, R, is the area under the demand curve up to the quantity, Q, it sells:

$$R = \int_0^Q D(z)\mathrm{d}z,$$

where z is a placeholder for quantity. The monopoly's objective is to maximize its profit through its choice of Q:

$$\max_Q \pi = \int_0^Q D(z)\mathrm{d}z - C(Q). \tag{12A.1}$$

Its first-order condition for a maximum is found by differentiating Equation 12A.1 to obtain

$$\frac{\mathrm{d}\pi}{\mathrm{d}Q} = D(Q) - \frac{\mathrm{d}C(Q)}{\mathrm{d}Q} = 0. \tag{12A.2}$$

According to Equation 12A.2, the discriminating monopoly sells units up to the quantity, Q, where the reservation price for the last unit, $D(Q)$, equals its marginal cost, $\mathrm{d}C(Q)/\mathrm{d}Q$. (This quantity is $Q_c = Q_d$ in Figure 12.2.)

For this solution to maximize profits, the second-order condition must hold: $\mathrm{d}^2\pi/\mathrm{d}Q^2 = \mathrm{d}D(Q)/\mathrm{d}Q - \mathrm{d}^2C(Q)/\mathrm{d}Q^2 < 0$. Thus, the second-order condition holds if the marginal cost curve has a nonnegative slope (because the demand curve has a negative slope). More generally, the second-order condition holds if the demand curve has a greater (absolute) slope than the marginal cost curve.

The perfectly price-discriminating monopoly's profit is

$$\pi = \int_0^Q D(z)\mathrm{d}z - C(Q).$$

For example, if $D(Q) = a - bQ$,

$$\pi = \int_0^Q (a - bz)dz - C(Q) = aQ - \frac{b}{2}Q^2 - C(Q). \tag{12A.3}$$

The monopoly finds the output that maximizes the profit by setting the derivative of the profit in Equation 12A.3 equal to zero:

$$a - bQ - \frac{\mathrm{d}C(Q)}{\mathrm{d}Q} = 0.$$

By rearranging terms, we find that $D(Q) = a - bQ = \mathrm{d}C(Q)/\mathrm{d}Q = MC$, as in Equation 12A.2. Thus, the monopoly produces the quantity at which the demand curve hits the marginal cost curve.

Appendix 12B: Group Price Discrimination

Suppose that a monopoly can divide its customers into two groups, as in Figure 12.3. It sells Q_1 to the first group and earns revenues of $R_1(Q_1)$, and it sells Q_2 units to the second group and earns $R_2(Q_2)$. Its cost of producing total output $Q = Q_1 + Q_2$ units is $C(Q)$. The monopoly can maximize its profit through its choice of prices or quantities to each group. We examine its problem when it chooses quantities:

$$\max_{Q_1, Q_2} \pi = R_1(Q_1) + R_2(Q_2) - C(Q_1 + Q_2). \tag{12C.1}$$

We obtain the first-order conditions by differentiating Equation 12C.1 with respect to Q_1 and Q_2 and setting the partial derivatives equal to zero:

$$\frac{\partial \pi}{\partial Q_1} = \frac{\mathrm{d}R_1(Q_1)}{\mathrm{d}Q_1} - \frac{\mathrm{d}C(Q)}{\mathrm{d}Q}\frac{\partial Q}{\partial Q_1} = 0, \tag{12C.2}$$

$$\frac{\partial \pi}{\partial Q_2} = \frac{\mathrm{d}R_2(Q_2)}{\mathrm{d}Q_2} - \frac{\mathrm{d}C(Q)}{\mathrm{d}Q}\frac{\partial Q}{\partial Q_2} = 0. \tag{12C.3}$$

Equation 12C.2 says that the marginal revenue from sales to the first group, $MR^1 = \mathrm{d}R_1(Q_1)/\mathrm{d}Q_1$, should equal the marginal cost of producing the last unit of total output, $MC = \mathrm{d}C(Q)/\mathrm{d}Q$, because $\partial Q/\partial Q_1 = 1$. Similarly, Equation 12C.3 says that the marginal revenue from the second group, MR^2, should also equal the marginal cost. By combining Equations 12C.2 and 12C.3, we find that the two marginal revenues are equal where the monopoly is profit maximizing:

$$MR^1 = MR^2 = MC.$$

Appendix 12C: Block Pricing

In the block-pricing utility monopoly example in the chapter, we assume that the utility monopoly faces an inverse demand curve and that its marginal and average cost is $m = 30$. Consequently, the quantity-discounting utility's profit is

$$\begin{aligned}\pi &= p(Q_1)Q_1 + p(Q_2)(Q_2 - Q_1) - mQ_2 \\ &= (90 - Q_1)Q_1 + (90 - Q_2)(Q_2 - Q_1) - 30Q_2,\end{aligned}$$

where Q_1 is the largest quantity for which the first-block rate, $p_1 = 90 - Q_1$, is charged and Q_2 is the total quantity a consumer purchases. The utility chooses Q_1 and Q_2 to maximize its profit. It sets the derivative of profit with respect to Q_1 equal to zero, $Q_2 - 2Q_1 = 0$, and the derivative of profit with respect to Q_2 equal to zero, $Q_1 - 2Q_2 + 60 = 0$. By solving these two equations, the utility determines its profit-maximizing quantities, $Q_1 = 20$ and $Q_2 = 40$. The corresponding block prices are $p_1 = 90 - 20 = 70$ and $p_2 = 50$.

Appendix 12D: Two-Part Pricing

In the example of a two-part pricing with nonidentical consumers, the demand curves for Valerie, Consumer 1, and Neal, Consumer 2, are $q_1 = 80 - p$ and $q_2 = 100 - p$. The consumer surplus for Consumer 1 is $CS_1 = \frac{1}{2}(80 - p)q_1 = \frac{1}{2}(80 - p)^2$. Similarly, $CS_2 = \frac{1}{2}(100 - p)^2$. If the monopoly charges the lower fee, $\mathcal{L} = CS_1$, it sells to both consumers and its profit is

$$\pi = 2\mathcal{L} + (p - m)(q_1 + q_2) = (80 - p)^2 + (p - 10)(180 - 2p).$$

Setting the derivative of π with respect to p equal to zero, we find that the profit-maximizing price is $p = 20$. The monopoly charges a fee of $\mathcal{L} = CS_1 = \$1{,}800$ and makes a profit of \$5,000. If the monopoly charges the higher fee, $\mathcal{L} = CS_2$, it sells only to Consumer 2, and its profit is

$$\pi = \mathcal{L} + (p - m)q_2 = \tfrac{1}{2}(100 - p)^2 + (p - 10)(100 - p).$$

The monopoly's profit-maximizing price is $p = 10$, and its profit is $\mathcal{L} = CS_2 = \$4{,}050$. Thus, the monopoly makes more by setting $\mathcal{L} = CS_1$ and selling to both customers.

Appendix 12E: Profit-Maximizing Advertising and Production

To maximize its profit, a monopoly must optimally set its advertising, A, and quantity, Q. Suppose that advertising affects only current sales, so the demand curve the monopoly faces is $p = p(Q, A)$.

As a result, the firm's revenue is $R = p(Q, A)Q = R(Q, A)$. The firm's cost of production is the function $C(Q)$. Its cost of advertising is A, because each unit of advertising costs \$1 (we chose the units of measure appropriately). Thus, its total cost is $C(Q) + A$.

The monopoly maximizes its profit through its choice of quantity and advertising:

$$\max_{Q,\,A} \pi = R(Q, A) - C(Q) - A. \tag{12E.1}$$

To obtain the necessary (first-order) conditions, we partially differentiate the profit function, Equation 12E.1 with respect to Q and A, and set the partial derivatives equal to zero:

$$\frac{\partial \pi(Q, A)}{\partial Q} = \frac{\partial R(Q, A)}{\partial Q} - \frac{dC(Q)}{dQ} = 0, \tag{12E.2}$$

$$\frac{\partial \pi(Q, A)}{\partial A} = \frac{\partial R(Q, A)}{\partial A} - 1 = 0. \tag{12E.3}$$

The profit-maximizing output and advertising levels are the Q^* and A^* that simultaneously satisfy Equations 12E.2 and 12E.3. Equation 12E.2 shows that output should be chosen so that the marginal revenue, $\partial R(Q, A)/(\partial Q)$, equals the marginal cost, $\mathrm{d}C(Q)/(\mathrm{d}Q)$. According to Equation 12E.3, the monopoly advertises to the point where its marginal revenue from the last unit of advertising, $\partial R(Q, A)/(\partial A)$, equals the marginal cost of the last unit of advertising, \$1.

Appendix 13A: Nash-Cournot Equilibrium

Here we use calculus to determine the Nash-Cournot equilibrium for n identical oligopolistic firms. We first solve for the equilibrium using general demand and cost functions, which are identical for all firms. Then we apply this general solution to a linear example. Finally, using the linear example, we determine the equilibrium when two firms have different marginal costs.

General Model

Suppose that the market demand function is $p(Q)$ and that each firm's cost function is the same $C(q_i)$. To analyze a Cournot market of identical firms, we first examine the behavior of a representative firm. Firm 1 tries to maximize its profits through its choice of q_1:

$$\max_{q_1} \pi_1(q_1, q_2, \cdots, q_n) = q_1 p(q_1 + q_2 + \cdots + q_n) - C(q_1), \quad (13A.1)$$

where $q_1 + q_2 + \cdots + q_n = Q$, the total market output. Firm 1 takes the outputs of the other firms as fixed. If Firm 1 changes its output by a small amount, the price changes by $(\mathrm{d}p(Q)/\mathrm{d}Q)(\mathrm{d}Q/\mathrm{d}q_1) = \mathrm{d}p(Q)/\mathrm{d}Q$. We determine its necessary condition to maximize profit (the first-order condition) by partially differentiating profit in Equation 13A.1 with respect to q_1 and setting the result equal to zero. After we rearrange terms, this necessary condition is

$$MR = p(Q) + q_1 \frac{\mathrm{d}p(Q)}{\mathrm{d}Q} = \frac{\mathrm{d}C(q_1)}{\mathrm{d}q_1} = MC, \quad (13A.2)$$

or marginal revenue equals marginal cost. Equation 13A.2 specifies the firm's best-response function: the optimal q_1 for any given output of other firms.

The marginal revenue expression can be rewritten as $p[1 + (q_1/p)(\mathrm{d}p/\mathrm{d}Q)]$. Multiplying and dividing the last term by n, noting that $Q = nq_1$ (given that all firms are identical), and observing that , the market elasticity of demand, is $(\mathrm{d}Q/\mathrm{d}p)(p/Q)$, we can rewrite Equation 13A.2 as

$$p\left(1 + \frac{1}{n}\right) = \frac{\mathrm{d}C(q_1)}{\mathrm{d}q_1}. \quad (13A.3)$$

The left side of Equation 13A.3 expresses Firm 1's marginal revenue in terms of the elasticity of demand of its residual demand curve, n , which is the number of firms, n, times the market demand elasticity, . Holding   constant, the more firms, the more elastic the residual demand curve, and hence the closer a firm's marginal revenue to the price.

We can rearrange Equation 13A.3 to obtain an expression for the Lerner Index, $(p - MC)/p$, in terms of the market demand elasticity and the number of firms:

$$\frac{p - MC}{p} = -\frac{1}{n}. \tag{13A.4}$$

The larger the Lerner Index, the greater the firm's market power. As Equation 13A.4 shows, if we hold the market elasticity constant and increase the number of firms, the Lerner Index falls. As n approaches ∞, the elasticity any one firm faces approaches $-\infty$, so the Lerner Index approaches 0 and the market is competitive.

Linear Example

Now suppose that the market demand is linear, $p = a - bQ$, and each firm's marginal cost is m, a constant, and it has no fixed cost. Firm 1, a typical firm, maximizes its profits through its choice of q_1:

$$\max_{q_1} \pi_1(q_1, q_2, \cdots, q_n) = q_1[a - b(q_1 + q_2 + \cdots + q_n)] - mq_1. \tag{13A.5}$$

Setting the derivative of profit with respect to q_1, holding the output levels of the other firms fixed, equal to zero, and rearranging terms, we find that the necessary condition for Firm 1 to maximize its profit is

$$MR = a - b(2q_1 + q_2 + \cdots + q_n) = m = MC. \tag{13A.6}$$

Because all firms have the same cost function, $q_2 = q_3 = \cdots = q_n \equiv q$ in equilibrium. Substituting this expression into Equation 13A.6, we find that the first firm's best-response function is

$$q_1 = R_1(q_2, \cdots, q_n) = \frac{a - m}{2b} - \frac{n - 1}{2}q. \tag{13A.7}$$

We derive the other firms' best-response functions similarly.

All these best-response functions must hold simultaneously. The intersection of the best-response functions determines the Nash-Cournot equilibrium. Setting $q_1 = q$ in Equation 13A.7 and solving for q, we find that the Nash-Cournot equilibrium output for each firm is

$$q = \frac{a - m}{(n + 1)b}. \tag{13A.8}$$

Total market output, $Q = nq$, equals $n(a - m)/[(n + 1)b]$. The corresponding price is obtained by substituting this expression for market output into the demand function:

$$p = \frac{a + nm}{n + 1}. \tag{13A.9}$$

Setting $n = 1$ in Equations 13A.8 and 13A.9 yields the monopoly quantity and price. As n becomes large, each firm's quantity approaches zero, total output approaches $(a - m)/b$, and price approaches m, which are the competitive levels. In Equation 13A.9, the Lerner Index is

$$\frac{p - MC}{p} = \frac{a - m}{a + nm}.$$

As n grows very large, the denominator goes to ∞, so the Lerner Index goes to 0, and the firm has no market power.

Different Costs

In the linear example with two firms, how does the equilibrium change if the firms have different marginal costs? The marginal cost of Firm 1 is m_1, and that of Firm 2 is m_2. Firm 1 chooses output to maximize its profit:

$$\max_{q_1} \pi_1(q_1, q_2) = q_1[a - b(q_1 + q_2)] - m_1 q_1. \tag{13A.10}$$

Setting the derivative of Firm 1's profit with respect to q_1, holding q_2 fixed, equal to zero, and rearranging terms, we find that the necessary condition for Firm 1 to maximize its profit is $MR_1 = a - b(2q_1 + q_2) = m_1 = MC$. Using algebra, we can rearrange this expression to obtain Firm 1's best-response function:

$$q_1 = \frac{a - m_1 - bq_2}{2b}. \tag{13A.11}$$

By similar reasoning, Firm 2's best-response function is

$$q_2 = \frac{a - m_2 - bq_1}{2b}. \tag{13A.12}$$

To determine the equilibrium, we solve Equations 13A.11 and 13A.12 simultaneously for q_1 and q_2:

$$q_1 = \frac{a - 2m_1 + m_2}{3b}, \tag{13A.13}$$

$$q_2 = \frac{a - 2m_2 + m_1}{3b}. \tag{13A.14}$$

By inspecting Equations 13A.13 and 13A.14, we find that the firm with the smaller marginal cost has the larger equilibrium output. Similarly, the low-cost firm has a higher profit. If m_1 is less than m_2, then

$$\pi_1 = \frac{(a + m_2 - 2m_1)^2}{9b} > \frac{(a + m_1 - 2m_2)^2}{9b} = \pi_2.$$

Appendix 13B: Nash-Stackelberg Equilibrium

We use calculus to derive the Nash-Stackelberg equilibrium for the linear example given in Appendix 13A with two firms that have the same marginal cost, m. Because Firm 1, the Stackelberg leader, chooses its output first, it knows that Firm 2, the follower, will choose its output using its best-response function, which is (see Equation 13A.7, where $n = 2$)

$$q_2 = R_2(q_1) = \frac{a - m}{2b} - \frac{1}{2}q_1. \tag{13B.1}$$

The Stackelberg leader's profit, $\pi_1(q_1 + q_2)$, can be written as $\pi_1(q_1 + R_2(q_1))$, where we've replaced the follower's output with its best-response function. The Stackelberg leader maximizes its profit by taking the best-response function as given:

$$\max_{q_1} \pi_1(q_1, R_2(q_1)) = q_1\left[a - b\left(q_1 + \frac{a - m}{2b} - \frac{1}{2}q_1\right)\right] - mq_1. \quad (13B.2)$$

Setting the derivative of Firm 1's profit (in Equation 13B.2) with respect to q_1 equal to zero and solving for q_1, we find that the profit-maximizing output of the leader is

$$q_1 = \frac{a - m}{2b}. \quad (13B.3)$$

Substituting the expression for q_1 in Equation 13B.3 into Equation 13B.1, we obtain the equilibrium output of the follower:

$$q_2 = \frac{a - m}{4b}. \quad (13B.4)$$

Appendix 13C: Nash-Bertrand Equilibrium

We can use math to determine the cola market Nash-Bertrand equilibrium discussed in the chapter. First, we determine the best-response functions each firm faces. Then we equate the best-response functions to determine the equilibrium prices for the two firms.

Coke's best-response function tells us the price Coke charges that maximizes its profit as a function of the price Pepsi charges. We use the demand curve for Coke to derive the best-response function.

The reason Coke's price depends on Pepsi's price is that the quantity of Coke demanded, q_c, depends on the price of Coke, p_c, and the price of Pepsi, p_p. Coke's demand curve is

$$q_c = 58 - 4p_c + 2p_p. \quad (13C.1)$$

Partially differentiating Equation 13C.1 with respect to p_c (that is, holding the price of Pepsi fixed), we find that the change in quantity for every dollar change in price is $\partial q_c/\partial p_c = -4$, so a \$1-per-unit increase in the price of Coke causes the quantity of Coke demanded to fall by 4 units. Similarly, the demand for Coke rises by 2 units if the price of Pepsi rises by \$1, while the price of Coke remains constant: $\partial q_c/\partial p_p = 2$.

If Coke faces a constant marginal and average cost of m per unit, its profit is

$$\pi_c = (p_c - m)q_c = (p_c - m)(58 - 4p_c + 2p_p), \quad (13C.2)$$

where $p_c - m$ is Coke's profit per unit. To determine Coke's profit-maximizing price (holding Pepsi's price fixed), we set the partial derivative of the profit function, Equation 13C.2, with respect to the price of Coke equal to zero,

$$\frac{\partial \pi_c}{\partial p_c} = q_c + (p_c - m)\frac{\partial q_c}{\partial p_c} = q_c - 4(p_c - m) = 0, \quad (13C.3)$$

and solve for p_c as a function of p_p and m to find Coke's best-response function:

$$p_c = 7.25 + 0.25p_p + 0.5m. \quad (13C.4)$$

Equation 13C.4 shows that Coke's best-response price is 25¢ higher for every extra dollar that Pepsi charges and 50¢ higher for every extra dollar of Coke's marginal cost.

If Coke's average and marginal cost of production is \$5 per unit, its best-response function is

$$p_c = 9.75 + 0.25p_p, \tag{13C.5}$$

as Figure 13.8 shows. If $p_p = \$13$, then Coke's best response is to set $p_c = \$13$.

Pepsi's demand curve is

$$q_p = 63.2 - 4p_p + 1.6p_c. \tag{13C.6}$$

Using the same approach, we find that Pepsi's best-response function (for $m = \$5$) is

$$p_p = 10.4 + 0.2p_c. \tag{13C.7}$$

The intersection of Coke's and Pepsi's best-response functions (Equations 13C.5 and 13C.7) determines the Nash equilibrium. By substituting Pepsi's best-response function, Equation 13C.7, for p_p in Coke's best-response function, Equation 13C.5, we find that $p_c = 9.75 + 0.25(10.4 + 0.2p_c)$. Solving this equation for p_c, we determine that the equilibrium price of Coke is \$13. Substituting $p_c = \$13$ into Equation 13C.6, we discover that the equilibrium price of Pepsi is also \$13.

Appendix 15A: Factor Demands

If a competitive firm hires L units of labor at a wage rate of w and K units of capital at a rental rate of r, it can produce $q = f(L, K)$ units of output. The firm sells its output at the market price of p. The firm picks L and K to maximize its profit:

$$\max_{L,\,K} \pi = pq - (wL + rK) = pf(L, K) - (wL + rK). \tag{15A.1}$$

Thus, the firm's revenue, pq, and cost both depend on L and K, so its profit depends on L and K.

Profit is maximized by setting the partial derivatives of profit (in Equation 15A.1) with respect to L and K equal to zero:

$$\frac{\partial \pi}{\partial L} = pMP_L - w = 0, \tag{15A.2}$$

$$\frac{\partial \pi}{\partial L} = pMP_K - r = 0, \tag{15A.3}$$

where $MP_L = \partial f(L, K)/\partial L$, the marginal product of labor, is the partial derivative of the production function with respect to L, and $MP_K = \partial f(L, K)/\partial K$ is the marginal product of capital. Solving Equations 15A.2 and 15A.3 simultaneously produces the factor demand equations.

Rearranging Equations 15A.2 and 15A.3, we can write these factor demand equations as

$$MRP_L \equiv pMP_L = w,$$

$$MRP_K \equiv pMP_K = r.$$

Thus, the firm maximizes its profit when it picks its inputs such that the marginal revenue product of labor equals the wage and the marginal revenue product of capital equals the rental rate of capital. For these conditions to produce a maximum, the second-order conditions must also hold. These second-order conditions say that the MRP_L and MRP_K curves slope downward.

If the production function is Cobb-Douglas, $q = AL^aK^b$, then Equations 15A.2 and 15A.3 are

$$\frac{\partial \pi}{\partial L} = paAL^{a-1}K^b - w = 0,$$
$$\frac{\partial \pi}{\partial K} = pbAL^aK^{b-1} - r = 0.$$

Solving these equations for L and K, we find that the factor demand functions are

$$L = \left(\frac{a}{w}\right)^{(1-b)/d}\left(\frac{b}{r}\right)^{b/d}(Ap)^{1/d}, \tag{15A.4}$$

$$K = \left(\frac{\alpha}{w}\right)^{\alpha/d}\left(\frac{\beta}{r}\right)^{(1-\alpha)/d}(Ap)^{1/d}, \tag{15A.5}$$

where $d = 1 - a - b$. By differentiating Equations 15A.4 and 15A.5, we can show that the demand for each factor decreases with w or r and increases with p.

If the Cobb-Douglas production function has constant returns to scale, $d = 0$, then Equations 15A.4 and 15A.5 are not helpful. The problem is that with constant returns to scale, a competitive firm with a Cobb-Douglas production function does not care how much it produces (and hence how many inputs it uses) as long as the market price and input prices are consistent with zero profit.

A competitive firm with a Cobb-Douglas production function pays labor the value of its marginal product, $w = pMP_L = paAL^{a-1}K^b = paQ/L$. As a result, the share of the firm's revenues that labor receives is $\omega_L = wL/(pQ) = a$. Similarly, $\omega_K = rK/(pQ) = b$. Thus, with a Cobb-Douglas production function, the shares of labor and of capital are fixed and independent of prices.

Appendix 15B: Monopsony

If only one firm can hire labor in a town, the firm is a monopsony. It chooses how much labor to hire to maximize its profit,

$$\pi = p(Q(L))Q(L) - w(L)L,$$

where $Q(L)$ is the production function, the amount of output produced using L hours of labor, and $w(L)$ is the labor supply curve, which shows how the wage varies with the amount of labor the firm hires. The firm maximizes its profit by setting the derivative of profit with respect to labor equal to zero (if the second-order condition holds):

$$\left(p + Q(L)\frac{\mathrm{d}p}{\mathrm{d}Q}\right)\frac{\mathrm{d}Q}{\mathrm{d}L} - w(L) - \frac{\mathrm{d}w}{\mathrm{d}L}L = 0. \tag{15B.1}$$

Rearranging terms in Equation 15B.1, we find that the maximization condition is that the marginal revenue product of labor,

$$MRP_L = pMPL = \left(p + Q(L)\frac{\mathrm{d}p}{\mathrm{d}Q}\right)\frac{\mathrm{d}Q}{\mathrm{d}L} = p\left(1 + \frac{1}{\varepsilon}\right)\frac{\mathrm{d}Q}{\mathrm{d}L},$$

equals the marginal expenditure,

$$ME = w(L) + \frac{\mathrm{d}w}{\mathrm{d}L}L = w(L)\left(1 + \frac{w}{L}\frac{\mathrm{d}w}{\mathrm{d}L}\right) = w(L)\left(1 + \frac{1}{\eta}\right), \tag{15B.2}$$

where η is the supply elasticity of labor.

If the supply curve is linear, $w(L) = g + hL$, the monopsony's expenditure is $E = w(L)L = gL + hL^2$, and the monopsony's marginal expenditure is $ME = \mathrm{d}E/\mathrm{d}L = g + 2hL$. Thus, the slope of the marginal expenditure curve, $2h$, is twice as great as that of the supply curve, h.

By rearranging the terms in Equation 15B.2, we find that

$$\frac{ME - w}{w} = \frac{1}{\eta}.$$

Thus, the markup of the marginal expenditure (and the value to the monopsony) to the wage, $(ME - w)/w$, is inversely proportional to the elasticity of supply. If the firm is a price taker, so η is infinite, the wage equals the marginal expenditure.

Appendix 16A: The Present Value of Payments over Time

We now derive the last equality in Equation 16.3, $PV = (f/i)[1 - 1/(1 + i)^t]$, which is the present value from a stream of annual payments f that lasts for t periods at an interest rate i. From the first equality in Equation 16.3, we know that the present value is

$$PV = f\left[\frac{1}{(1 + i)^1} + \frac{1}{(1 + i)^2} + \cdots + \frac{1}{(1 + i)^t}\right]. \tag{16A.1}$$

We multiply both sides of Equation 16A.1 by $(1 + i)$, we obtain

$$\begin{aligned} PV(1 + i) &= f\left[1 + \frac{1}{(1 + i)} + \frac{1}{(1 + i)^2} + \cdots + \frac{1}{(1 + i)^{t-1}}\right] \\ &= f\left[1 + \frac{PV}{f} - \frac{1}{(1 + i)^t}\right]. \end{aligned} \tag{16A.2}$$

Rearranging the terms in Equation 16A.2 using algebra, we obtain

$$PV = \frac{f}{i}\left[1 - \frac{1}{(1 + i)^t}\right]. \tag{16A.3}$$

If payment continues forever (t goes to ∞), then the last term in Equation 16A.3, $1/(1 + i)^t$, goes to zero, so that we obtain Equation 15.17,

$$PV = \frac{f}{i}. \tag{16A.4}$$

Appendix 18A: Welfare Effects of Pollution in a Competitive Market

We now show the welfare effects of a negative externality in a competitive market where demand and marginal costs are linear, as in Figure 18.1. The inverse demand curve is

$$p = a - bQ, \tag{18A.1}$$

where p is the price of the output and Q is the quantity. The private marginal cost is the competitive supply curve if pollution is an externality:

$$MC^p = c + dQ. \tag{18A.2}$$

The marginal cost to people exposed to the pollution (gunk) is

$$MC^g = eQ. \tag{18A.3}$$

Equation 18A.3 shows that no pollution harm occurs if output is zero and that the marginal harm increases linearly with output. The social marginal cost is the sum of the private marginal cost and the marginal cost of the externality:

$$MC^s = c + (d + e)Q. \tag{18A.4}$$

The intersection of the demand curve, Equation 18A.1, and the supply curve, Equation 18A.2, determines the competitive equilibrium where pollution is an externality:

$$p_c = a - bQ_c = c + dQ_c = MC^p. \tag{18A.5}$$

If we solve Equation 18A.5 for Q, the competitive equilibrium quantity is

$$Q_c = \frac{a - c}{b + d}.$$

Substituting this quantity into the demand curve, we find that the competitive price is $p_c = a - b(a - c)/(b + d)$.

If the externality is taxed at a rate equal to its marginal cost, so the externality is internalized, the market produces the social optimum. We find the social optimum by setting p in Equation 18A.1 equal to MC^s in Equation 18A.4 and solving for the resulting quantity:

$$Q_s = \frac{a - c}{b + d + e}.$$

The corresponding price is $p_s = a - b(a - c)/(b + d + e)$.

If only the monopoly sells the output, we find the monopoly's revenue by multiplying both sides of Equation 18A.1 by quantity: $R = aQ - bQ^2$. Differentiating with respect to quantity, we find that the monopoly's marginal revenue is

$$MR = a - 2bQ. \tag{18A.6}$$

If the monopoly is unregulated, its equilibrium is found by setting MR, Equation 18A.6, equal to private marginal cost, Equation 18A.2, and solving for output:

$$Q_m = \frac{a - c}{2b + d}.$$

The corresponding price is $p_m = a - b(a - c)/(2b + d)$. If the monopoly internalizes the externality due to a tax equal to MC^g, the equilibrium quantity is

$$Q_m^* = \frac{a - c}{2b + d + e}.$$

The price is $p_m^* = a - b(a - c)/(2b + d + e)$.

In Figure 18.1, $a = 450$, $b = 2$, $c = 30$, $d = 2$, and $e = 1$. Substituting these values into the equations, we solve for the following equilibrium values:

	Quantity	Price
Competition	105	240
Social optimum (competition with a tax)	84	282
Monopoly	70	310
Monopoly with a tax	60	330

Appendix 20A: Nonshirking Condition

An efficiency wage acts like a bond to prevent shirking. An employee who never shirks is not fired and earns the efficiency wage, w. A fired worker goes elsewhere and earns the lower, going wage, $\underline{w}$. The expected value to a shirking employee is

$$\theta\underline{w} + (1 - \theta)w + G,$$

where the first term is the probability of being caught shirking, θ, times earnings elsewhere if caught and fired; the second term is the probability of not being caught times the efficiency wage; and the third term, G, is the value a worker derives from shirking. The worker chooses not to shirk if the certain high wage from not shirking exceeds the expected return from shirking:

$$w \geq (1 - \theta)w + \theta\underline{w} + G,$$

which simplifies to Equation 20.2, $\theta(w - \underline{w}) \geq G$. That is, a risk-neutral worker does not shirk if the expected loss from being fired is greater than or equal to the gain from shirking.

Answers to Selected Questions and Problems

I know the answer! The answer lies within the heart of all mankind! The answer is twelve? I think I'm in the wrong building. —Charles Schultz

Chapter 2

1.1 Substituting the values for price and income into the demand function, we find that $Q = 63 - 11p + 7p_b + 3p_c + 2Y = 63 - 11p + 7(19) + 3(6) + 2(78) = 370 - 11p$.

1.2 Holding prices constant, $\Delta Q = Q_2 - Q_1 = 2(Y_2 - Y_1) = 2\Delta Y$ million kg, where Y is in thousands of Australian dollars. Thus, an AU\$200 increase in annual per capita income would cause demand to change by $\Delta Q = 2\Delta Y = 2 \times \frac{200}{1{,}000} = 0.4$ million kg per year. The demand curve would shift to the right at any price (that is, demand would increase) by 400,000 kg per year (from $Q = 370 - 11p$ to $Q = 370.4 - 11p$).

1.7 $Q = Q_1 + Q_2 = (120 - p) + \left(60 - \frac{1}{2}p\right) = 180 - 1.5p$.

2.1 Holding the price of lamb constant, $\Delta Q = Q_2 - Q_1 = -9(p_s^2 - p_s^1) = -9\Delta p_s$ million kg per year. If the price of sheep rises by AU\$0.50 per kg, then supply would change by $\Delta Q = -9\Delta p_s = -9 \times 0.5 = -4.5$. The supply curve would shift to the left at any price (that is, supply would increase) by 4.5 million kg per year (from $Q_1 = 104 + 8p$ to $Q_2 = 99.5 + 8p$).

3.1 The statement "Talk is cheap because supply exceeds demand" makes sense if we interpret it to mean that the *quantity supplied* of talk exceeds the *quantity demanded* at a price of zero. Imagine a downward-sloping demand curve that hits the horizontal, quantity axis to the left of where the upward-sloping supply curve hits the axis. (The correct aphorism is "Talk is cheap until you hire a lawyer.")

3.5 In equilibrium, the quantity demanded, $Q = a - bp$, equals the quantity supplied, $Q = c + ep$, so $a - bp = c + ep$. By solving this equation for p, we find that the equilibrium price is $p = (a - c)/(b + e)$. By substituting this expression for p into either the demand curve or the supply curve, we find that the equilibrium quantity is $Q = (ae + bc)/(b + e)$.

3.6 The supply curve for coffee beans is found by substituting the price of coffee cherries in the supply function: $Q_s = 3.15 + 0.1p - 0.5 \times 0.8 = 2.75 + 0.1p$. Set quantity demanded equal to quantity supplied, $Q_d = Q_s = Q$, to solve for the equilibrium price, p:

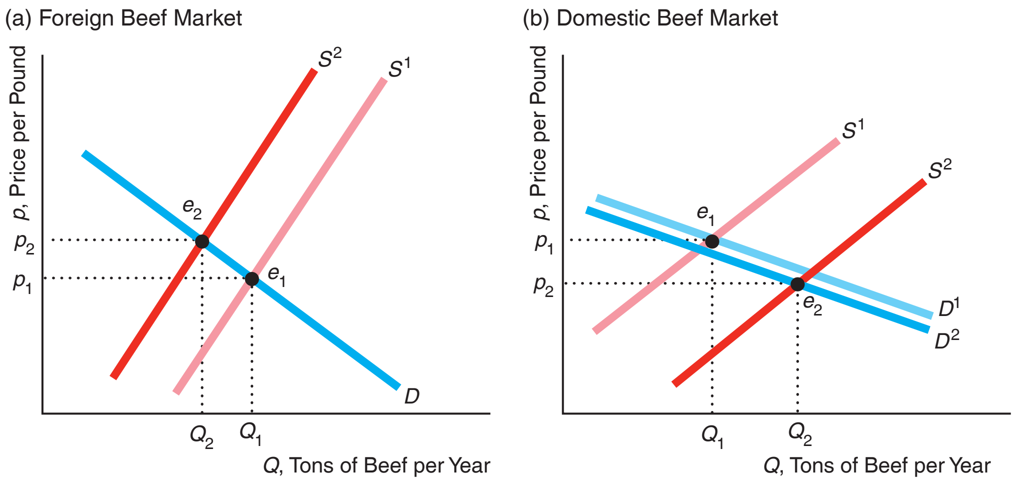
$$4.1 - 0.2p = 2.75 + 0.1p$$
$$0.3p = 1.35$$

$p = 4.5$ million pesos per thousand 60-kg bags. Substitute $p = 4.5$ into either the demand or supply curve to solve for the equilibrium quantity Q: $Q = 4.1 - 0.2p = 2.75 + 0.1p = 3.2$ thousand 60-kg bags.

4.1 The supply shock is unusually good luck or an unexpected increase in the number of lobsters in the sea (possibly due to a change in temperature). This supply shock causes the supply curve to shift to the right and does not affect the demand curve. Thus, the equilibrium moves along the demand curve. The equilibrium price falls and the equilibrium quantity increases.

5.2 A ban has no effect if foreigners supply nothing at the pre-ban, equilibrium price. Thus, if imports occur only at prices above those actually observed, a ban has no practical effect.

5.9 In the absence of price controls, the leftward shift of the supply curve as a result of a severe disruption in the supply of a product would push the market price up and reduce the equilibrium quantity. A legal requirement to keep the price below its market-clearing level—a binding price ceiling—would cause consumers to demand a larger quantity of the product than the quantity supplied. The resulting shortage would impose significant search costs on consumers and harm those unable to obtain the product. The reduced quantity and price would also reduce profits for firms selling the product. This policy would benefit the consumers who are able to acquire the product at the lower price.



7.1 As shown in figure (a), the beef supply curve would shift to the left from S^1 to S^2 in the importing countries that imposed the ban. The foreign demand curve, D, would be unaffected if the risk of foreign buyers consuming tainted meat did not increase. Thus, the leftward shift of the supply curve would cause the equilibrium to move along the demand curve from e_1 to e_2. The equilibrium price would rise from p_1 to p_2, and the equilibrium quantity would fall from Q_1 to Q_2.
As shown in figure (b), the closure of export markets for a country's beef products would be reflected in a rightward shift of the supply curve for beef for domestic sale in the short run from S^1 to S^2. In addition, the fall in domestic sales would be reflected in a leftward shift of the demand curve from D^1 to D^2. Both of these shifts work to reduce the equilibrium price from p_1 to p_2. If the rightward shift of the supply curve outweighs the leftward shift of the demand curve, then the equilibrium quantity would rise from Q_1 to Q_2, and the short-run equilibrium would shift from e_1 to e_2.

Chapter 3

2.2 According to Equation 3.1, the elasticity of demand is $\ $ = (percentage change in quantity demanded) ÷ (percentage change in price) = −3.8% ÷ 10% = −0.38, which is inelastic.

2.6 Differentiating the demand function as $Q = Ap^{\ }$ with respect to p, we find that $dQ/dp = \ Ap^{\ -1}$. To get the elasticity, we multiply dQ/dp by $p/Q = p/Ap^{\ } = 1/Ap^{\ -1}$. That is, the elasticity is $\ Ap^{\ -1} \times 1/Ap^{\ -1} = \ $. Because this result holds for any p, the elasticity is the same, $\ $, at every point along the demand curve.

2.14 The elasticity of demand is (slope) × (p/Q) = $(\Delta Q/\Delta p)(p/Q)$ = (−9.5 thousand metric tons per year per cent) × (45¢/1,275 thousand metric tons per year) ≈ −0.34. That is, for every 1% fall in the price, a third of a percent more coconut oil is demanded. The cross-price elasticity of demand for coconut oil with respect to the price of palm oil is $(\Delta Q/\Delta p_p)(p_p/Q) = 16.2 \times (31/1{,}275) \approx 0.39$.

3.1 Because the linear supply function is $Q = g + hp$, a change in price of Δp causes a $\Delta Q = h\Delta p$ change in quantity. Thus, $\Delta Q/\Delta p = h$, and the elasticity of supply is $\eta = (\Delta Q/\Delta p)(p/Q) = hp/Q$. By substituting for Q using the supply function, we find that $\eta = hp/(g + hp)$. By using the supply function to substitute for p, we learn that $\eta = (Q - g)/Q$.

4.6 By dividing both the numerator and the denominator of the right side of Equation 3.7 by η, we can rewrite that incidence equation as

$$\frac{\eta}{\eta - \ } = \frac{1}{1 - \ /\eta}.$$

As η goes to infinity, $\ /\eta$ goes to zero, so the incidence approaches 1.

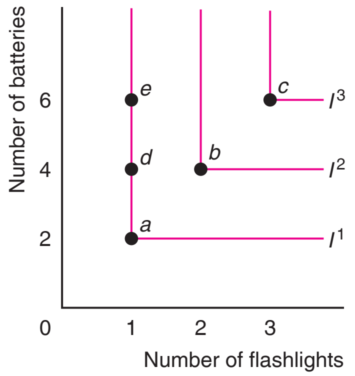
4.14 In a competitive market, the effect of a specific tax is the same whether the government collects it from suppliers or demanders. Thus, if the market for milk is competitive, consumers will pay the same price in equilibrium regardless of whether the government taxes consumers or stores.

4.18 The incidence of the tax on consumers is zero if the demand curve is perfectly elastic or the supply curve is perfectly inelastic.

Chapter 4

1.5 If the neutral product is on the vertical axis, the indifference curves are parallel vertical lines.

2.2 As flashlights and batteries are perfect complements for Aliza, her indifference curves are right angles. Her utility function is $U = \min(2F, B)$, where "min" denotes the minimum of the two arguments, F is the number of flashlights, and B is the number of batteries.



2.7 Because José Maria's utility is $U(B, Z) = AB^aZ^b$, his marginal utility of B is $MU_B = aAB^{a-1}Z^b$, his marginal utility of Z is $MU_Z = bAB^aZ^{b-1}$, and his marginal rate of substitution is $MRS = -MU_Z/MU_B = (bB)/(aZ)$.

3.4 Suppose that Dale purchases two goods at prices p_1 and p_2. If her original income is Y, the intercept of the budget line on the Good 1 axis (where the consumer buys only Good 1) is Y/p_1. Similarly, the intercept is Y/p_2 on the Good 2 axis. A 50% income tax lowers income to half its original level, $Y/2$. As a result, the budget line shifts inward toward the origin. The intercepts on the Good 1 and Good 2 axes are $Y/(2p_1)$ and $Y/(2p_2)$, respectively. The opportunity set shrinks by the area between the original budget line and the new line.

4.5 Andy's marginal utility of apples divided by its price is $\frac{3}{2} = 1.5$. The marginal utility for kumquats is $\frac{5}{4} = 1.2$. That is, a dollar spent on apples gives him more extra utils than a dollar spent on kumquats. Thus, he maximizes his utility by spending all his money on apples and buying $40/2 = 20$ pounds of apples.

4.6 If we plot B on the vertical axis and Z on the horizontal axis, the slope of David's indifference curve is $-MU_Z/MU_B = -2$. The marginal utility from one extra unit of Z is twice that from one extra unit of B. Thus, if the price of Z is less than twice as much as that of B, David buys only Z (the optimal bundle is on the Z axis at Y/p_Z, where Y is his income and p_Z is the price of Z). If the price of Z is more than twice that of B, David buys only B. If the price of Z is exactly twice as much as that of B, he is indifferent between buying any bundle along his budget line.

4.9 Using Equations 4B.11 and 4B.12, we find that the necessary conditions for a utility maximum are $B = 100\alpha/[2(\alpha + \beta)]$ and $Z = 100\beta/(\alpha + \beta)$.

4.19 If a wealthy person spends more on food than a poor person before the subsidy, then the wealthy person is more likely to be spending more than the value of the food stamps prior to receiving them and hence is less likely to have a tangency at a point like f in Figure 4.12.

5.2 Consumers do not always notice taxes that are added at the register, so including the tax in the list price may discourage sales. This effect is less likely to be important for people buying a car because they are more likely to keep the tax in mind.

6.2 Consumers who live near a border will weigh small transportation costs against sales tax rate differences in deciding which state to shop in. Higher sales tax rates increase the relative prices of the same products. If the rate differential becomes large enough to offset the transportation costs of shopping across the border, sales will fall (possibly significantly) in the area close to the border in the higher-taxed state and adversely affect businesses there. To mitigate the situation, the higher-taxed state might consider reducing the rate of sales tax in the border region, which is what Mexico did prior to 2014 in areas bordering the United States.

Chapter 5

1.2 Point E_1 corresponds to Bundle e_1 on indifference curve I^1, whereas E_2 corresponds to Bundle e_2 on indifference curve I^2, which is farther from the origin than I^1, so Mimi's utility is higher at E_2 than at E_1. Mimi is better off at E_2 than at E_1 because the price of beer is lower at E_2, so she can buy more goods with the same budget. This result generalizes: consumers are better off along the demand curve at lower prices.

2.4 An opera performance must be a normal good for Don because he views the only other good he buys as an inferior good. To show this result in a graph, draw a figure similar to Figure 5.3, but relabel the vertical "Housing" axis as "Opera performances." Don's equilibrium will be in the upper-left quadrant at a point like a in Figure 5.3.

2.5 The consumer's budget constraint is

$$p_1q_1 + p_2q_2 + \cdots + p_nq_n = Y,$$

where Y is income, and p_i is the price and q_i is the quantity of Good i. Differentiating with respect to Y, we find that

$$p_1\frac{dp_1}{dY} + p_2\frac{dq_2}{dY} + \cdots + p_n\frac{dq_n}{dY} = \frac{dY}{dY} = 1.$$

Multiplying and dividing each term by q_iY, we rewrite this last equation as

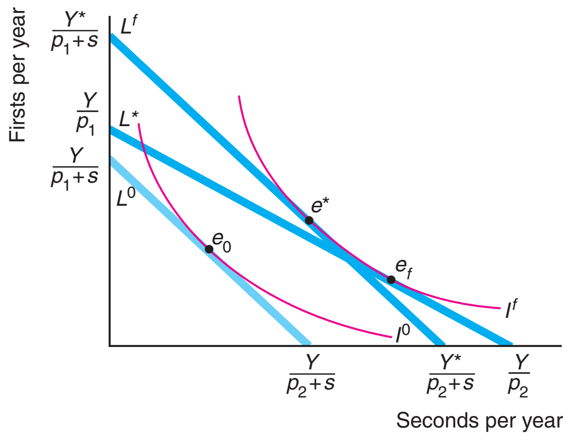
$$\frac{p_1q_1}{Y}\frac{dq_1}{dY}\frac{Y}{q_1} + \frac{p_2q_2}{Y}\frac{dq_2}{dY}\frac{Y}{q_2} + \cdots + \frac{p_nq_n}{Y}\frac{dq_n}{dY}\frac{Y}{q_n} = 1,$$

or

$$\omega_1\eta_1 + \omega_2\eta_2 + \cdots + \omega_n\eta_n = 1,$$

where η_i, the income elasticity for each Good i, equals $(dq_i/dY)(Y/q_i)$, and the budget share of Good i is $\omega_i = p_iq_i/Y$. That is, the weighted sum of the income elasticities equals 1. For this equation to hold, at least one of the goods must have a positive income elasticity; hence, not all the goods can be inferior.

3.7 In the graph, L^f is the budget line at the factory store and L^o is the constraint at the outlet store. At the factory store, the consumer maximum occurs at e_f on indifference curve I^f. Suppose that we increase the income of a consumer who shops at the outlet store to Y^*, so that the resulting budget line L^* is tangent to the indifference curve I^f. The consumer would buy Bundle e^*. That is, the pure substitution effect (the movement from e_f to e^*) causes the consumer to buy relatively more firsts. The total effect (the movement from e_f to e_o) reflects both the substitution effect (firsts are now relatively less expensive) and the income effect (the consumer is worse off after paying for shipping).



4.4 The CPI accurately reflects the true cost of living because Alix does not substitute between the goods as the relative prices change.

5.10 Afonso's average tax rate in 2016 was

$$\frac{0.145 \times 7{,}035 + 0.21 \times 13{,}065 + 0.37 \times 9{,}900}{30{,}000} = 25\%.$$

If he had one more euro, he would have kept $1 - 0.37 = €0.63$. This means that his marginal tax rate—the tax rate on that additional euro—was 37%. A comparison tells us that the marginal tax is higher than the average tax under a progressive income tax system.

5.14 A negative wage elasticity of labor supply, which represents the total effect of a wage change, means that a rise in wages will reduce hours worked. This inverse relationship indicates that the labor supply curve slopes downwards. The substitution effect of a rise in wages always causes leisure to fall (and hours worked to increase). Since the total effect shows less hours worked overall, the income effect must be positive and greater than the substitution effect. In other words, leisure is a normal good and thus more is consumed as income rises, which means that fewer hours are spent working. A fall in the tax rate from t_H to t_L would increase the effective wage from $w(1 - t_H)$ to $w(1 - t_L)$. As the income effect is greater than the substitution effect, the tax rate cut would reduce the number of hours worked and tax revenues would fall for both reasons. A downward sloping labor supply curve is typically associated with a wage that is very high.

6.2 The government could give a smaller lump-sum subsidy that shifts the L^{LS} curve down so that it is parallel to the original curve but tangent to indifference curve I^2. This tangency point is to the left of e_2, so the parents would consume fewer hours of childcare than with the original lump-sum payment.

6.3 Parents who do not receive subsidies prefer that poor parents receive lump-sum payments rather than a subsidized hourly rate for childcare. If the supply curve for day-care services is upward sloping, by shifting the demand curve farther to the right, the price subsidy raises the price of day care for these other parents.

Chapter 6

2.3 No, it is not possible for $q = 10$, $L = 3$, and $K = 6$ to be a point on this production function. Holding output and other inputs fixed, a production function shows the minimum amount needed of a given factor. As only 5 units of capital are needed to produce 10 units of output given that 3 units of labor are used, using 6 units of capital would imply excess capital.

3.1 The total product curve is $q = 6L$, the average product of labor curve is $q/L = 6$, and the marginal product of labor curve is $\Delta q/\Delta L = 6$. The graph of the total product curve, with q on the vertical axis and L on the horizontal axis, is a line with a slope of 6. The graphs of the average and marginal product of labor curves, with average and marginal product on the vertical axis and L on the horizontal axis, are horizontal lines at a height of 6 units.

3.4 a. The average product of labor, holding capital fixed at $\overline{K}$, is $AP_L = q/L = L^{-0.25}\,\overline{K}^{0.25} = (\overline{K}/L)^{0.25}$.

b. The marginal product of labor is $MP_L = \mathrm{d}q/\mathrm{d}L = \frac{3}{4}(\overline{K}/L)^{0.25}$.

4.7 Using Equation 6.3, we know that the marginal rate of technical substitution is $MRTS = -MP_L/MP_k = -\frac{2}{3}$.

4.8 The isoquant for $q = 10$ is a straight line that hits the B axis at 10 and the G axis at 20. The marginal product of B is 1 everywhere along the isoquant. The marginal rate of technical substitution is $-\frac{1}{2}$ if B is on the horizontal axis.

4.9 The isoquant looks like the "right angle" ones in panel b of Figure 6.3 because the firm cannot substitute between disks and machines but must use them in equal proportions: one disk and one hour of machine services.

5.5 Returns to scale is a property of production functions that relates the percentage change in production to the same percentage change in all inputs. For example, if both K and L are doubled (both are increased by 100%), then for the Cobb-Douglas production function, $q = K^aL^b$, output rises to

$$(2K)^a(2L)^b = 2^{a+b}K^aL^b = 2^{a+b}q.$$

Any production function exhibits constant returns to scale when a doubling of both K and L causes output to double (increase by 100%). For the Cobb-Douglas production function, this means that $a + b = 1$, so the percentage increase in output is

$$\frac{2q - q}{q} \times 100 = \frac{(2 - 1)q}{q} \times 100 = 100\%.$$

Any production function exhibits increasing returns to scale when a doubling of both K and L causes output to more than double (increase by more than 100%). For this Cobb-Douglas production function, this means that $a + b > 1$. To illustrate this, if $a + b = 1.5$, then the percentage increase in output would be

$$\frac{2^{1.5}q - q}{q} \times 100 = \frac{(2.83 - 1)q}{q} \times 100 = 183\%.$$

6.2 The marginal product of labor of Firm 1 is only 90% of the marginal product of labor of Firm 2 for a particular level of inputs. Using calculus, we find that the MP_L of Firm 1 is $\partial q_1/\partial L = 0.9\partial f(L, K)/\partial L = 0.9\partial q_2/\partial L$.

7.2 Not enough information is given to fully answer this question. If we assume that Japanese and U.S. firms have identical production functions and produce using the same ratio of factors during good times, Japanese firms will have a lower average product of labor during recessions because they are less likely to lay off workers. However, it is not clear how Japanese and U.S. firms expand output during good times (do they hire the same number of extra workers?). As a result, we cannot predict which country has the higher average product of labor.

Chapter 7

1.3 If the plane cannot be resold, its purchase price is a sunk cost, which is unaffected by the number of times the plane is flown. Consequently, the average cost per flight falls with the number of flights, but the total cost of owning and operating the plane rises because of extra consumption of gasoline and maintenance. Thus, the more frequently someone has reason to fly, the more likely that flying one's own plane costs less per flight than a ticket on a commercial airline. However, by making extra ("unnecessary") trips, Mr. Agassi raises his total cost of owning and operating the airplane.

1.4 The sunk cost is $1 per pipe. The opportunity cost of each pipe is $9.

2.10 The firm's average cost curve is:

$$AC(q) = \frac{C(q)}{q} = \frac{25}{q} + q.$$

Average cost changes with output according to the equation

$$\frac{dAC(q)}{dq} = 1 - \frac{25}{q^2}(\gtreqless)0 \text{ for } q(\gtreqless)5.$$

This means that average cost falls as q rises until $q = 5$ units, at which output level average cost is at a minimum at $AC = 10$. Average cost increases as q rises above $q = 5$ units.

3.2 You produce your output, exam points, using as inputs the time spent on Question 1, t_1, and the time spent on Question 2, t_2. If you have diminishing marginal returns to extra time on each problem, your isoquants have the usual shapes: They curve away from the origin. You face a constraint that you may spend no more than 60 minutes on the two questions: $60 = t_1 + t_2$. The slope of the 60-minute isocost curve is -1: For every extra minute you spend on Question 1, you have one less minute to spend on Question 2. To maximize your test score, given that you can spend no more than 60 minutes on the exam, you want to pick the highest isoquant that is tangent to your 60-minute isocost curve. At the tangency, the slope of your isocost curve, -1, equals the slope of your isoquant, $-MP_1/MP_2$. That is, your score on the exam is maximized when $MP_1 = MP_2$, where the last minute spent on Question 1 would increase your score by as much as spending it on Question 2 would. Therefore, you've allocated your time on the exam wisely if you are indifferent as to which question to work on during the last minute of the exam.

3.3 According to Equation 7.8, if the firm were minimizing its cost, the extra output it gets from the last dollar spent on labor, $MP_L/w = \frac{50}{200} = 0.25$, should equal the extra output it derives from the last dollar spent on capital, $MP_K/r = \frac{200}{1{,}000} = 0.2$. Thus, the firm is not minimizing its costs. It would do better if it used relatively less capital and more labor, from which it gets more extra output from the last dollar spent.

3.6 To minimize costs, the firm will produce where $MRTS = -w/r$. For the given production function,

$$MRTS = -\frac{MP_L}{MP_K} = -\frac{0.8AP_L}{0.2AP_K} = -4\frac{K}{L}.$$

In Country 1, the firm would set

$$MRTS_1 = -4\frac{K}{L} = -\frac{w_1}{r_1} = -\frac{10}{10} = -1$$

and use four times as much labor as capital in production, $L = 4K$. Substituting this result into the production function,

$$\begin{aligned} 100 &= K^{0.5}(4K)^{0.5} = 2K \\ K &= 50 \\ L &= 200. \end{aligned}$$

The firm would use 50 units of capital and 200 units of labor to produce 100 units of output, and the cost of production would be $C_1 = w_1L + r_1k = 10 \times 200 + 10 \times 50 = 2{,}500$.

In Country 2, the firm would set

$$\begin{aligned} MRTS_2 = -4\frac{K}{L} &= -\frac{w_2}{r_2} = -\frac{w_1 \times 1.2}{r_1 \times 0.8} \\ &= -\frac{12}{8} = -1.5, \end{aligned}$$

and use 2.67 times as much labor as capital in production, $L = 2.67K$. Substituting this result into the production function,

$$100 = K^{0.5}(2.67K)^{0.5} = 1.63K$$
$$K = 61$$
$$L = 163.$$

The firm would use 61 units of capital and 163 units of labor to produce 100 units of output, and the cost of production would be $C_2 = w_2L + r_2k = 12 \times 163 + 8 \times 61 = 2{,}500$. Thus, while the firm would use different combinations of factors to produce the same amount of output in the two countries, the total cost of production would be the same.

3.7 From the given information, it appears that cricket balls are produced using a constant returns to scale, fixed-proportion production function. A fixed quantity of cork, yarn, leather, and labor is used to produce each cricket ball, and these factors cannot be substituted. If a firm wants to increase its production from one to two cricket balls per hour, it must use exactly twice as much of each input.

The corresponding cost function is $C(q) = (w + c + y + l)q$, where w is the wage for the period it takes to make one cricket ball, c is the cost of cork used for one cricket ball, y is the cost of the yarn used to make one cricket ball, and l is the cost of the leather used to produce one cricket ball.

3.12 Let w be the cost of a unit of L and r be the cost of a unit of K. Because the two inputs are perfect substitutes in the production process, the firm uses only the less expensive of the two inputs. Therefore, the long-run cost function is $C(q) = wq$ if $w \le r$; otherwise, it is $C(q) = rq$.

4.2 As q increases, its long-run average cost falls. For example, if output increases from $q = 1$ to $q = 2$ units, average cost falls from $AC(q = 1) = 2$ to $AC(q = 2) = 1$. This relationship is called economies of scale: average cost decreases as returns to scale increase.

$$\frac{dAC(q)}{dq} = -\frac{2}{q^2} < 0$$

If the firm also learns by doing, then average cost would fall even more as q increases. This would be reflected in the exponent of the output variable increasing as well. For example, if the exponent rises from $q(= q^1)$ to q^2, then an increase in output from $q = 1$ to $q = 2$ units would cause average cost to fall from $AC(q = 1) = 2$ to

$$AC(q) = \frac{2}{q^2} = 0.5.$$

4.4 If $a = 0$, then $AC = 12$. Since average cost is constant, this means that the firm does not learn by doing (average cost does not fall as cumulative output increases).

On the other hand, if $a < 0$, then average cost falls as cumulative output increases, so the firm does learn by doing. For example, if $a = 0.5$ and cumulative output increases from 1 to 2,000 units, then average cost falls from

$$AC = 10 + 2(1)^{-0.5} = 12 \text{ to}$$
$$AC = 10 + 2(2{,}000)^{-0.5} = 10.04.$$

5.2 This firm has significant economies of scope, as producing gasoline and heating oil separately would cost approximately twice as much as producing them together. In this case, the measure of economies of scope, SC, is a positive number.

6.1 If $-w/r$ is the same as the slope of the line segment connecting the wafer-handling stepper and stepper technologies, then the isocost will lie on that line segment, and the firm will be indifferent between using either of the two technologies (or any combination of the two). In all the isocost lines in the figure, the cost of capital is the same, and the wage varies. The wage such that the firm is indifferent lies between the relatively high wage on the C^2 isocost line and the lower wage on the C^3 isocost line.

Chapter 8

2.4 The first-order condition to maximize profit is the derivative of the profit function with respect to q set equal to zero: $120 - 40 - 20q = 0$. Thus, profit is maximized where $q = 4$, so that $R(4) = 120 \times 4 = 480$, $VC(4) = (40 \times 4) + (10 \times 16) = 320$, $\pi(4) = R(4) - VC(4) - F = 480 - 320 - 200 = -40$. The firm should operate in the short run because its revenue exceeds its variable cost: $480 > 320$.

3.1 Suppose that a U-shaped marginal cost curve cuts a competitive firm's demand curve (price line) from above at q_1 and from below at q_2. By increasing output to $q_1 + 1$, the firm earns extra profit because the last unit sells for price p, which is greater than the marginal cost of that last unit. Indeed, the price exceeds the marginal cost of all units between q_1 and q_2, so it is more profitable to produce q_2 than q_1. Thus, the firm should either produce q_2 or shut down (if it is making a loss at q_2). We can also derive this result using calculus. The second-order condition, Equation 8B.3, for a competitive firm requires that marginal cost cut the demand line from below at q^*, the profit-maximizing quantity: $dMC(q^*)/dq > 0$.

3.3 The competitive firm's marginal cost function is found by differentiating its cost function with respect to quantity: $dC(q)/dq = b + 2cq + 3dq^2$. The firm's necessary profit-maximizing condition is $p = MC = b + 2cq + 3dq^2$. The firm solves this equation for q for a specific price to determine its profit-maximizing output.

3.13 Some fishermen would stay in port (or harvest other species) when the price of the haddock is less than their average variable costs. Others with lower average variable costs would choose to continue harvesting haddock.

4.3 Laws of this type reduce the flexibility firms have to react to changes in market conditions. If conditions change suddenly for the worse, a firm may have to operate at a loss for some time before it can adjust or shut down. This potential extra expense may initially discourage some firms from entering the market.

4.9 To derive the expression for the elasticity of the residual or excess supply curve in Equation 8.7, we differentiate the residual supply curve (Equation 8.6), $S^r(p) = S(p) - D^o(p)$, with respect to p to obtain

$$\frac{dS^r}{dp} = \frac{dS}{dp} - \frac{dD^0}{dp}.$$

Let $Q_r = S^r(p)$, $Q = S(p)$, and $Q_o = D(p)$. We multiply both sides of the differentiated expression by p/Q_r, and for convenience, we also multiply the second term by $Q/Q = 1$ and the last term by $Q_o/Q_o = 1$:

$$\frac{dS^r}{dp}\frac{p}{Q_r} = \frac{dS}{dp}\frac{p}{Q_r}\frac{Q}{Q} - \frac{dD^o}{dp}\frac{p}{Q_r}\frac{Q_o}{Q_o}.$$

We can rewrite this expression as Equation 8.7,

$$\eta_r = \frac{\eta}{\theta} - \frac{1-\theta}{\theta}\ {}_o,$$

where $\eta_r = (dS^t/dp)(p/Q_r)$ is the residual supply elasticity, $\eta = (dS/dp)(p/Q)$ is the market supply elasticity, ${}_o = (dD^o/dp)(p/Q_o)$ is the demand elasticity of the other countries, and $\theta = Q_r/Q$ is the residual country's share of the world's output (hence, $1 - \theta = Q_o/Q$ is the share of the rest of the world). *Note*: If n countries have equal outputs, then $1/\theta = n$, so this equation can be rewritten as $\eta_r = n\eta - (n-1)\ {}_o$.

4.10 The incidence of a tax on consumers is the share of the tax that falls on consumers. As the retail price rises by 50% of the national excise tax increase, the incidence of the tax is shared equally by consumers and producers. In contrast, since the retail price rises by 100% of the regional excise tax, its incidence falls fully on consumers; the price received by producers is unaffected. From Chapter 3, the change in the price that consumers pay with respect to the change in the tax is

$$\frac{\Delta p}{\Delta t} = \frac{\eta}{\eta - \ },$$

where η is the elasticity of supply and is the elasticity of demand for the good. As smaller values of η cause the consumer price to change less, the incidence on consumers also falls. For example, if $\eta = 0$ (the supply curve is vertical or perfectly inelastic), then $\Delta p/\Delta t = 0$, and the incidence of the tax falls entirely on producers.

From Chapter 8, we would expect the residual supply curve for a region to be more elastic than for the country as a whole:

$$\eta_r = \frac{\eta}{\theta} - \frac{1-\theta}{\theta}\ {}_0 > \eta,$$

where η_r is the residual elasticity, $\eta > 0$ is the national elasticity of supply, $0 < \theta \le 1$ is the region's share of the good, and ${}_0 \approx \ < 0$ is the elasticity of demand for the other regions. As $\eta < \eta_r$, we would expect the incidence of a national excise tax increase on consumers to be less than the incidence of a regional tax.

If all 20 regions are identical, then $\theta = 0.05$ and $\eta_r = 20\eta - 19\ {}_0$. As ${}_0 < 0$, the supply elasticity for any one region is at least 20 times larger than for the country as a whole.

5.4 The entry of Uber shifted the market supply curve for ride services to the right, reducing the market price and increasing the market quantity. In the short run, existing taxi companies would reduce output and incur losses since the market price is below their minimum average cost of production. In the long run, some taxi companies would leave the market.

Reducing the annual licence fee for taxis, a type of lump-sum tax, would lower their fixed costs, but not affect their marginal costs, so that the average cost curve for a representative taxi company would shift downwards along its marginal cost curve. Each would produce less, but losses could be significantly reduced or eliminated.

Chapter 9

2.1 The consumer surplus at a price of 30 is $450 = \frac{1}{2}(30 \times 30)$.

6.6 The Challenge Solution in Chapter 8 shows the long-run equilibrium effect of a lump-sum tax in a competitive market. Consumer surplus falls by more than tax revenue increases, and producer surplus remains zero, so welfare falls.

6.7 If the tax is based on *economic* profit, the tax has no long-run effect because the firms make zero economic profit. If the tax is based on *business* profit and business profit is greater than economic profit, the profit tax raises firms' after-tax costs and results in fewer firms in the market. The exact effect of the tax depends on why business profit is less than economic profit. For example, if the government ignores opportunity labor cost but includes all capital cost in computing profit, firms will substitute toward labor and away from capital.

6.11 a. The initial equilibrium is determined by equating the quantity demanded to the quantity supplied: $100 - 10p = 10p$. That is, the equilibrium is $p = 5$ and $Q = 50$. At the support price, the quantity supplied is $Q_s = 60$. The market-clearing price is $p = 4$. The deficiency payment was $D = (\underline{p} - p)Q_s = (6 - 4)60 = 120$.

b. Consumer surplus rises from

$$CS_1 = \tfrac{1}{2}(10 - 5)50 = 125 \text{ to}$$
$$CS_2 = \tfrac{1}{2}(10 - 4)60 = 180.$$

Producer surplus rises from

$$PS_1 = \tfrac{1}{2}(5 - 0)50 = 125 \text{ to}$$
$$PS_2 = \tfrac{1}{2}(6 - 0)60 = 180.$$

Welfare falls from

$$CS_1 + PS_1 = 125 + 125 = 250 \text{ to}$$
$$CS_2 + PS_2 - D = 180 + 180 - 120 = 240.$$

Thus, the deadweight loss is 10.

Chapter 10

1.4 A subsidy is a negative tax. Thus, we can use the same analysis as in Solved Problem 10.2 to answer this question (reversing the signs of the effects).

2.4 Amos' marginal rate of substitution is $MRS_a = [\alpha/(1 - \alpha)]H_a/G_a$, and Elise's is $MRS_e = [\beta/(1 - \beta)] H_e/G_e$. Along the contract curve, the two marginal rates of substitution are equal: $MRS_a = MRS_e$. Thus, to find the contract curve, we equate the right sides of the expressions for MRS_a and MRS_e. Using the information about the endowments and some algebra, we can write the (quadratic) formula for the contract curve as

$$(\beta - \alpha)G_aH_a + \beta(\alpha - 1)50G_a + \alpha(1 - \beta)100H_a = 0.$$

4.1 If you draw the convex production possibility frontier on panel c of Figure 10.6, you will see that it lies strictly inside the concave production possibility frontier. Thus, more output can be obtained if Jane and Denise use the concave frontier. That is, each should specialize in producing the good for which she has a comparative advantage.

4.2 As Chapter 5 shows, the slope of the budget constraint facing an individual equals the negative of that person's wage. Panel a of the figure below illustrates that Pat's budget constraint is steeper than Chris' because Pat's wage is larger than Chris'. Panel b shows their combined budget constraint after they marry. Before they marry, each spends some time in the market earning money and other time at home cooking, cleaning, and consuming leisure. After they marry, one can specialize in earning money and the other at working at home. If they are both equally skilled at household work (or if Chris is better), then Pat has a comparative advantage (see Figure 10.6) in working in the market and Chris has a comparative advantage in working at home. Of course, if both enjoy consuming leisure, they may not fully specialize. As an example, suppose that before they lived together Chris and Pat each spent 10 hours a day in sleep and leisure activities, 5 hours working in the marketplace, and 9 hours working at home. Because Chris earns \$10 an hour and Pat earns \$20, they collectively earned \$150 a day and worked 18 hours a day at home. After they marry, they can benefit from specialization. If Chris works entirely at home and Pat works 10 hours in the market and the rest at home, they collectively earn \$200 a day (a one-third increase) and still have 18 hours of work at home. If they do not need to spend as much time working at home because of economies of scale, one or both could work more hours in the marketplace, and they will have even greater disposable income.

Chapter 11

1.3 At $Q = 10$, $p = 500 - 10(10) = 400$. The elasticity is $= -0.1(400)/10 = -4$. Revenue $R = pQ = 10(400) = 4{,}000$.

2.10 The firm's Lerner index is

$$\frac{p - MC}{p} = \frac{1{,}400 - 840}{1{,}400} = 0.4.$$

The firm faces a demand elasticity of

$$= -\frac{p}{p - MC} = -\frac{1}{0.4} = -2.5.$$

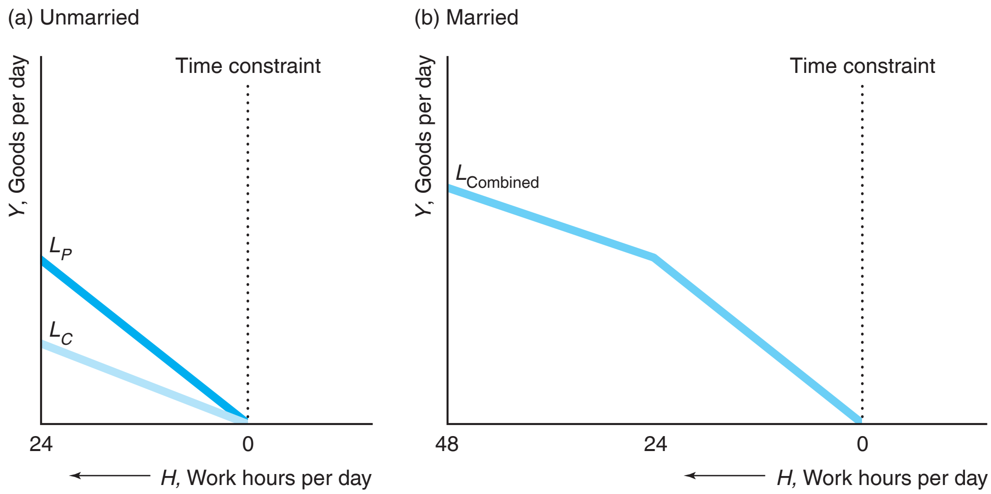
3.3 Before the specific tax is imposed, the profit-maximizing monopolist chooses to produce where marginal cost, $MC = 4Q$, equals marginal revenue, $MR = 500 - Q$, at $Q = 100$ units. It sets price equal to 450 and makes a profit of $\pi = pQ - C = 45{,}000 - 21{,}000 = 24{,}000$.
The specific tax shifts the monopolist's marginal cost curve up by 100 per unit. The monopolist now chooses to produce where the new marginal cost, $MC = 4Q + 100$, equals the same marginal revenue, $MR = 500 - Q$, at $Q = 80$ units. It sets price equal to 460 and makes an after-tax profit of

$$\pi = pQ - C_t = pQ - (1{,}000 + 2Q^2 + 100Q)$$
$$= 45{,}000 - 21{,}000 = 15{,}000.$$

The imposition of a profits tax, τ, instead of a specific will not alter the monopolist's MC curve and therefore not affect the initial monopoly optimum. Output and price will remain unchanged. However, the monopoly profit will be reduced by the amount of the profits tax. Since the specific tax raised tax revenue of 8,000, the rate of profits tax that raises the same amount of revenue for government would be

$$(1 + \tau)\pi = 8{,}000$$
$$\tau = \frac{8000}{24{,}000} - 1 = 33.3\%.$$

3.5 Suppose that the monopoly faces a constant-elasticity demand curve, with elasticity , has a constant marginal cost, m, and that the government imposes a specific tax of t. The monopoly sets



its price such that $p = (m + t)/(1 + 1/\ \ )$. Thus, $dp/dt = 1/(1 + 1/\ \ ) > 1$.

4.1 Yes. For example, if the demand curve cuts the average cost curve only in its downward-sloping section, the average cost curve is strictly downward sloping in the relevant region.

6.2 a. The marginal revenue curve is $MR = 20 - 8Q$. To maximize profits, the monopolist would set $MR = MC$ and produce $Q = 2$ units at the price, $p = 12$, and earn a profit of 16 in each of the two periods for a total profit of $\pi = 32$.

b. Setting $p_1 = MC = 4$, the firm would produce $Q_1 = 4$ units in the first period and earn no profit. This means that profit would have to be greater than 32 in the second period for the firm to benefit from the positive network externality. If the inverse demand curve in the second period is $p_2 = 20 - bQ_2$, where $b > 0$ is its flatter slope, then the marginal revenue curve is $MR_2 = 20 - 2bQ_2$. To maximize profit, the monopolist would set $MR = MC$ and produce $Q_2 = 8/b$ units at the price $p_2 = 12$. Since

$$\pi_2 = (p_2 - 4)Q_2 = \frac{64}{b} > 32,$$
$$b < 2.$$

Thus, the slope of the demand curve would have to fall below 2 for the firm's total profits over the two periods to be larger. For example, if $b = 1$, then $Q_2 = 8$ and $\pi_2 = 64 > 32 = \pi$.

Chapter 12

1.2 Governments may provide social housing as a form of charity, to meet the distinct housing needs of groups such as persons with disabilities, to give families the stability they need to succeed at school, in work and in the community, or to price discriminate by lowering the final price for less wealthy families (who presumably have higher elasticities of demand).

2.6 Under perfect price discrimination, the firm's profit is the area below the demand curve and above marginal cost. This area is $0.5 \times 60 \times 60 = 1{,}800$. The consumer surplus is zero, as all consumer surplus is extracted by the monopoly. The total surplus is therefore 1,800. The deadweight loss is zero because the monopoly produces up to the point where marginal cost cuts the demand curve. For a single-price monopoly, the profit is 900, the consumer surplus is 450, the total surplus is 1,350, and the deadweight loss is 450.

3.5 See MyLab Economics, Chapter 12, Supplemental Material, "Aibo," for more details about this robot. The two marginal revenue curves are $MR_J = 3{,}500 - Q_J$ and $MR_A = 4{,}500 - 2Q_A$. Equating the marginal revenues with the marginal cost of \$500, we find that $Q_J = 3{,}000$ and $Q_A = 2{,}000$. Substituting these quantities into the inverse demand curves, we learn that $p_J = \$2{,}000$ and $p_A = \$2{,}500$. Rearranging Equation 11.9, we know that the elasticities of demand are $\ _J = p/(MC - p) = 2{,}000/(500 - 2{,}000) = -4/3$ and $\ _A = 2{,}500/(500 - 2{,}500) = -5/4$. Thus, using Equation 12.3, we find that

$$\frac{p_J}{p_A} = \frac{2{,}000}{2{,}500} = 0.8 = \frac{1 + 1/(-5/4)}{1 + 1/(-4/3)} = \frac{1 + 1/\ _A}{1 + 1/\ _J}.$$

The profit in Japan is $(p_J - m)Q_J = (\$2{,}000 - \$500) \times 3{,}000 = \$4.5$ million, and the U.S. profit is \$4 million. The deadweight loss is greater in Japan, \$2.25 million $(= \frac{1}{2} \times \$1{,}500 \times 3{,}000)$, than in the United States, \$2 million $(= \frac{1}{2} \times \$2{,}000 \times 2{,}000)$.

3.6 The marginal revenue function corresponding to a linear inverse demand function has the same intercept and twice as steep a slope (see Chapter 11). Thus, the U.S. marginal revenue function is $MR_A = 100 - 2Q_A$, and the Japanese one is $MR_J = 80 - 4Q_J$. To determine how many units to sell in the United States, the monopoly sets its U.S. marginal revenue equal to its marginal cost, $MR_A = 100 - 2Q_A = 20$, and solves for the optimal quantity, $Q_A = 40$ units. Similarly, because $MR_J = 80 - 4Q_J = 20$, the optimal quantity is $Q_J = 15$ units in Japan. Substituting $Q_A = 40$ into the U.S. demand function, we find that $p_A = 100 - 40 = \$60$. Similarly, substituting $Q_J = 15$ units into the Japanese demand function, we learn that $p_J = 80 - (2 \times 15) = \$50$. Thus, the price-discriminating monopoly charges 20% more in the United States than in Japan. We can also show this result using elasticities. From Equation 3.3, we know that the elasticity of demand is $_A = -p_A/Q_A$ in the United States and $_J = -\frac{1}{2}P_J/Q_J$ in Japan. In the equilibrium, $_A = -60/40 = -3/2$ and $_J = -50/(2 \times 15) = -5/3$. As Equation 12.3 shows, the ratio of the prices depends on the relative elasticities of demand:

$$p_A/p_J = 60/50 = (1 + 1/\ _J)/(1 + 1/\ _A)$$
$$= (1 - 3/5)/(1 - 2/3) = 6/5.$$

3.7 Given that $p_A = 39 - 3Q_A$ and $p_B = 71 - 7Q_B$, the corresponding marginal revenue functions are $MR_A = 39 - 6Q_A$ and $MR_B = 71 - 14Q_B$. Setting MR_A equal to the marginal cost, 1, and solving, we find that $39 - 6Q_A = 1$, or $38 = 6Q_A$, or $Q_A = 38/6 \approx 6.33$. Substituting this quantity into the inverse demand function, we find that $p_A = 39 - 3(38/6) = 20$. Similarly, $Q_B = 5$ and $p_B = 36$.

3.9 From the problem, we know that the profit-maximizing Chinese price is $p = 3$ and the quantity is $Q = 0.1$ (million). The marginal cost is $m = 1$. The Lerner markup equation is $(p - MC)/p = -1/$ (Equation 11.9). Thus, $(p_C - m)/p_C = (3 - 1)/3 = -1/\ _C$, so $_C = -3/2$. If the Chinese inverse demand curve is $p = a - bQ$, then the corresponding marginal revenue curve is $MR = a - 2bQ$. Warner maximizes its profit where $MR = a - 2bQ = m = 1$, so its optimal $Q = (a - 1)/(2b)$. Substituting this expression into the inverse demand curve, we find that its optimal $p = (a + 1)/2 = 3$, or $a = 5$. Substituting that result into the output equation, we have $Q = (5 - 1)/(2b) = 0.1$ (million). Thus, $b = 20$, the inverse demand function is $p = 5 - 20Q$, and the marginal revenue function is $MR = 5 - 40Q$. Using this information, you can draw a figure similar to Figure 12.3.

3.12 With the ban, the monopoly's profit in each country is the output times the difference between the price and its constant average cost, 1. The monopoly's profit in Country 1 is $\pi_1 = (3.50 - 1)5 = 12.50$. Its profit in Country 2 is $\pi_2 = (5 - 1)4 = 16$. Thus, its total profit is $\pi = \pi_1 + \pi_2 = 12.50 + 16 = 28.50$. Without the ban, the monopoly's profits are $\pi_1 = (4 - 1)4 = 12$, $\pi_2 = (4 - 1)5 = 15$, and $\pi = 12 + 15 = 27$. Thus, the monopoly's profit falls from 28.50 to 27 if it loses the ability to price discriminate.

3.14 A limited-time offer allows a firm to maximize its profit by price discriminating. They induce people with a relatively higher price elasticity of demand to act. Consumers must incur some cost, such as their time, to obtain the discount being offered. By spending that extra time, price-sensitive consumers differentiate themselves.

4.3 If the monopoly sets the first-block price equal to \$120, the cost to consumers of buying 60 units is \$3,600, which equals their consumer surplus. Thus, consumers are willing to make this purchase, and the monopoly captures the entire potential surplus. This outcome is the same as perfect price discrimination.

5.3 Under the uniform price, Elena would pay €60 and take 10 lessons per month, and the club would earn a profit of €400.

$$MR = 100 - 8q = 20 = MC$$
$$q = 10$$
$$p = 60$$
$$\pi = (p - AC)q = 400$$

Under two-part pricing, Elena would pay €20 and take 20 lessons per month, and the club would maximize its profit from a membership fee that equals Elena's consumer surplus of €800. The club's profit from Elena would double.

$$p = 100 - 4q = 20 = MC$$
$$q = 20$$
$$p = 20$$
$$\pi = CS = 0.5(100 - 20)20 = 800$$

6.4 a. When the firm prices each course separately, it maximizes profit by charging £48 for the main course (and selling to two potential customers) and £7 for dessert (and selling to 3 potential customers). Its total revenue is £117.

b. Yes, the firm's revenue is higher at £128 if it bundles the main course and the dessert together, charges £32 for the combination and sells to all four potential customers. In addition, the firm's revenue would be only slightly lower if it charged £42 for the bundle and sold to three of the potential customers.

7.3 These types of events would reduce tourism to the area and shift the marginal benefit curve for advertising to the left. To maximize profit, the quantity of advertising should be chosen where the marginal benefit (the extra gross profit from one more unit of advertising or the marginal revenue from one more unit of output) equals its marginal cost. Accordingly, a profit-maximizing business would reduce the amount it spends on advertising for that region.

7.4 The monopoly's profit is $\pi = (800 - 4Q + 0.2A^{0.5})Q - 2Q - A = 798Q - 4Q^2 + 0.2A^{0.5}Q - A$. Setting the partial derivatives of the profit function with respect to Q and A equal to zero, we obtain the first-order conditions $\partial\pi/\partial Q = 798 - 8Q + 0.2A^{0.5} = 0$ and $\partial\pi/\partial A = 0.1A^{-0.5}Q - 1 = 0$. Rearranging the first of these conditions, we learn that $A^{0.5} = 0.1Q$. Substituting this expression into the second condition, we find that $798 - 8Q + 0.02Q = 0$, or $Q = 100$. Thus, $A^{0.5} = 0.1Q = 10$, so $A = 100$.

Chapter 13

2.4 While cartels are illegal in most developed countries, the lure of collusive profits entices many firms to enter into cartel agreements. At the same time, cartel members have an incentive to cheat by producing more than the agreed-upon amount (or by lowering their price) and increasing their own profits at the expense of other cartel members. In the absence of legal means to enforce the cartel agreement, cartels tacitly rather than explicitly coordinate their activities to avoid running afoul of competition laws. Cooperating to put informal operating rules in place to regulate activities was a means used by the Dutch cartels to build trust among members, reduce instability, and prolong the duration and profitability of the cartel.

2.7 The profit-maximizing cartel output is the monopoly output. Setting $MR = MC$ yields $100 - 4Q = 20$, so $Q = 20$. Each of the four firms produces $q = 20/4 = 5$.

3.2 The inverse demand curve is $p = 1 - 0.001Q$. Thus, $\pi_1 = [1 - 0.001(q_1 + q_2)]q_1 - 0.28q_1$. Its first-order condition is $d\pi_1/dq_1 = 1 - 0.001(2q_1 + q_2) - 0.28 = 0$. If we rearrange the terms, the first firm's best-response function is $q_1 = 360 - \frac{1}{2}q_2$. Similarly, the second firm's best-response function is $q_2 = 360 - \frac{1}{2}q_1$. By substituting one of these best-response functions into the other, we learn that the Nash-Cournot equilibrium occurs at $q_1 = q_2 = 240$, so the equilibrium price is 52¢.

3.4 The market elasticity of demand is

$$= \frac{\Delta Q}{\Delta p}\frac{p}{Q} = -1.$$

From Equation 13.7 or 13.8,

$$p = \frac{MC}{\left(1 + \dfrac{1}{n}\right)} = \frac{MC}{\left(1 - \dfrac{1}{n}\right)} = \beta MC,$$

where β is the markup of price over marginal cost. As MC rises holding the number of firms constant, p increases by the same proportion, $\Delta p = \beta \Delta MC$, so market power does not change. As the number of firms falls holding MC constant, β rises and market power increases.

3.9 Firm 1's profit is $\pi_1 = q_1[a - b(q_1 + q_2)] - mq_1$. Consequently, its best-response function is $q_1 = (a - m - bq_2)/(2b)$, where we replace m_1 with m in Equation 13A.11. (Alternatively, you can derive this result using calculus.) Firm 2's profit is $\pi_2 = q_2[a - b(q_1 + q_2)] - (m + x)q_2$. Simultaneously solving these best-response functions for q_1 and q_2, we get the equilibrium quantities in Equations 13A.13 and 13A.14, where we've substituted for the appropriate marginal costs:

$$q_1 = \frac{a - 2m + (m + x)}{3b} = \frac{a - m + x}{3b},$$

$$q_2 = \frac{a - 2(m + x) + m}{3b} = \frac{a - m - 2x}{3b},$$

By inspection,

$$q_1 = [a - m + x]/[3b] > q_2$$
$$= [a - m - 2x]/[3b].$$

The low-cost firm, Firm 1, has the higher profit. The profits are $\pi_1 = (a + [m + x] - 2m)^2/[9b]$ and $\pi_2 = (a + m - 2[m + x])^2/[9b]$. Thus,

$$\pi_1 = \frac{(a - m + x)^2}{9b} > \frac{(a - m - 2x)^2}{9b} = \pi_2.$$

3.10 By differentiating its product, a firm makes the residual demand curve it faces less elastic everywhere. For example, no consumer will buy from that firm if its rival charges less and the goods are homogeneous. In contrast, some consumers who prefer this firm's product to that of its rival will still buy from this firm even if its rival charges less. As the chapter shows, a firm sets a higher price, the lower the elasticity of demand at the equilibrium.

4.1 a. Since $p = 100 - 5q_1 - 5q_2$, the best-response function for Firm 1 is:

$$MR_1 = 100 - 10q_1 - 5q_2 = 10 = MC$$
$$q_1 = 9 - 0.5q_2.$$

Similarly, $q_2 = 9 - 0.5q_1$.

Alternatively, starting with Firm 1's profit function and using calculus, the best-response function for Firm 1 is:

$$\pi_1 = (100 - 5q_1 - 5q_2)q_1 - 10q_1$$
$$= 90q_1 - 5q_1^2 - 5q_2q_1$$
$$\frac{\partial \pi_1}{\partial q_1} = 0 = 90 - 10q_1 - 5q_2$$
$$q_1 = 9 - 0.5q_2.$$

Similarly, $q_2 = 9 - 0.5q_1$.

Solve the system of best-response functions to find the Nash-Cournot equilibrium quantities and price.

$$q_1 = 9 - 0.5(9 - 0.5q_1)$$
$$q_1 = 6 = q_2$$
$$Q = q_1 + q_2 = 12$$
$$p = 40.$$

b. When Firm 1 moves first, substitute Firm 2's best-response function into the inverse demand curve to find Firm 1's optimal output and then use Firm 2's best-response function to obtain its optimal output.

$$p = 100 - 5q_1 - 5(9 - 0.5q_1) = 55 - 2.5q_1$$
$$MR = 55 - 5q_1 = 10 = MC$$
$$q_1 = 9$$
$$q_2 = 9 - 0.5 \times 9 = 4.5.$$

Alternatively, substitute Firm 2's best-response function into Firm 1's profit function to find each firm's optimal output.

$$\pi_1 = 90q_1 - 5q_1^2 - 5q_1(9 - 0.5q_1)$$
$$= 45q_1 - 2.5q_1^2$$
$$\frac{\partial \pi_1}{\partial q_1} = 0 = 45 - 5q_1$$
$$q_1 = 9$$
$$q_2 = 4.5.$$

Both approaches yield the total Stackelberg output, $Q = q_1 + q_2 = 13.5$, and price, $p = 32.5$.

4.4 Since it is an oligopoly, the retail outlet of a wireless communications company would likely be willing to pay a higher rent for the space than a convenience store, which operates in a more competitive market environment with lower profit.

4.5 The inverse demand curve is $p = a - bQ = 1 - 0.001Q$, and the marginal cost is $m = 0.28$. We can use the formulas in Appendix 13B to determine the Stackelberg quantities. Thus, the first firm's quantity is $q_1 = (1 - 0.28)/0.002 = 360$. The second firm's quantity is $q_2 = (1 - 0.28)/0.004 = 180$. In contrast, the Nash-Cournot quantities are both 240. Thus, the Stackelberg leader produces more and the follower producers less than the Nash-Cournot quantities.

4.6 Duopoly: Since $p = 30 - 2q_1 - 2q_2$, the best-response function for Firm 1 is:

$$MR_1 = 30 - 4q_1 - 2q_2 = 6 = MC$$
$$q_1 = 6 - 0.5q_2.$$

Similarly, $q_2 = 6 - 0.5q_1$.

Solve the system of best-response functions to find the Nash-Cournot solution.

$$q_1 = 6 - 0.5(6 - 0.5q_1)$$
$$q_1 = 4 = q_2$$
$$Q = q_1 + q_2 = 8$$
$$p = 14$$
$$\pi = (p - AC)Q = (14 - 6)8 = 64$$
$$CS = 0.5(30 - p)Q = 0.5(30 - 14)8 = 64$$
$$DWL = 0.5(p - MC)(Q_C - Q) = 16$$

where $Q_C = 12$ is the level of output with marginal cost pricing.

Monopoly: The combined firm would set marginal revenue equal to marginal cost to maximize profit.

$$p = 30 - 2Q$$
$$MR = 30 - 4Q = 6 = MC$$
$$Q = 6$$
$$p = 18$$
$$\pi = (p - AC)Q = 72$$
$$CS = 0.5(30 - p)Q = 36$$
$$DWL = 0.5(p - MC)(Q_C - Q) = 36$$

A comparison reveals that the merged firm produces less, sells it product at a higher price, and earns larger profit. Consumer surplus falls and deadweight loss rises.

5.2 Given that the duopoly firms produce identical goods, the equilibrium price is lower if the duopolies set price rather than quantity. If the goods are heterogeneous, we cannot answer this question definitively.

5.5 Firm 1 wants to maximize its profit: $\pi_1 = (p_1 - 10)q_1 = (p_1 - 10)(100 - 2p_1 + p_2)$. Its first-order condition is $d\pi_1/dp_1 = 100 - 4p_1 + p_2 + 20 = 0$, so its best-response function is $p_1 = 30 + \frac{1}{4}p_2$. Similarly, Firm 2's best-response function is $p_2 = 30 + \frac{1}{4}p_1$. Solving, the Nash-Bertrand equilibrium prices are $p_1 = p_2 = 40$. Each firm produces 60 units.

6.6 a. The Nash-Cournot equilibrium in the absence of a government intervention is $q_1 = 30$, $q_2 = 40$, $p = 50$, $\pi_1 = 900$, and $\pi_2 = 1{,}600$.
b. The Nash-Cournot equilibrium is now $q_1 = 33.3$, $q_2 = 33.3$, $p = 53.3$, $\pi_1 = 1{,}108.9$, and $\pi_2 = 1{,}108.9$.
c. As Firm 2's profit was 1,600 in part a, a fixed cost slightly greater than 1,600 will prevent entry.

7.1 Before any subsidies, the best-response functions are:

$$q_1 = 9 - 0.5q_2$$
$$q_2 = 9 - 0.5q_1.$$

And the Nash-Cournot equilibrium quantities, price, and profits are:

$$q_1 = 9 - 0.5(9 - 0.5q_1)$$
$$q_1 = 6 = q_2$$
$$Q = q_1 + q_2 = 12$$
$$p = 40$$
$$\pi = \pi_1 + \pi_2 = (p - AC)(q_1 + q_2)$$
$$= (40 - 10)(6 + 6)$$
$$= 180 + 180 = 360.$$

If only Firm 1 receives a subsidy, the best-response functions are:

$$MR_1 = 100 - 10q_1 - 5q_2 = 7 = MC_1$$
$$q_1 = 9.3 - 0.5q_2,$$
$$MR_2 = 100 - 5q_1 - 10q_2 = 10 = MC_2$$
$$q_2 = 9 - 0.5q_1.$$

And the Nash-Cournot equilibrium quantities, price, and profits are:

$$q_1 = 9.3 - 0.5(9 - 0.5q_1)$$
$$q_1 = 6.4$$
$$q_2 = 5.8$$
$$Q = q_1 + q_2 = 12.2$$
$$p = 39$$
$$\pi = \pi_1 + \pi_2 = (p - AC)(q_1 + q_2)$$
$$= (39 - 7)(6.4 + 5.8)$$
$$= 204.8 + 185.6 = 390.4.$$

Firm 1's output rises while Firm 2's output falls, price falls, and profit rises for both firms, but much more for Firm 1.

If both firms receive the subsidy, the best-response functions are:

$$q_1 = 9.3 - 0.5q_2$$
$$q_2 = 9.3 - 0.5q_1,$$

And the Nash-Cournot equilibrium quantities, price, and profits are:

$$\begin{aligned}
q_1 &= 9.3 - 0.5(9.3 - 0.5q_1) \\
q_1 &= 6.2 = q_2 \\
Q &= q_1 + q_2 = 12.4 \\
p &= 38 \\
\pi &= \pi_1 + \pi_2 = (p - AC)(q_1 + q_2) \\
&= (38 - 7)(6.2 + 6.2) \\
&= 192.2 + 192.2 = 384.4.
\end{aligned}$$

Compared to case where only Firm 1 receives the subsidy, Firm 1's output falls slightly while Firm 2's output rises until the two are equal (and higher than in the no subsidy equilibrium), price falls even more, and profit falls slightly for Firm 1 but rises significantly for Firm 2.

Chapter 14

1.1 The payoff matrix in this prisoners' dilemma game is

		Duncan	
		Squeal	*Stay Silent*
Larry	*Squeal*	–2 –2	–5 0
	Stay Silent	0 –5	–1 –1

If Duncan stays silent, Larry gets 0 if he squeals and −1 (a year in jail) if he stays silent. If Duncan confesses, Larry gets −2 if he squeals and −5 if he does not. Thus, Larry is better off squealing in either case, so squealing is his dominant strategy. By the same reasoning, squealing is also Duncan's dominant strategy. As a result, the Nash equilibrium is for both to confess.

1.4 A dominant strategy is one that produces a higher payoff to all possible strategies (best responses) that the other player might use.

a. If Divit does not enter, Sara's payoff is higher if she does enter. However, if Divit enters, Sara's payoff is higher if she does not enter. Thus, Sara does not have a dominant strategy.

b. If Sara does not enter, Divit's payoff is higher if he does enter. If Sara enters, Divit's payoff is still higher if he enters. Thus, regardless of what Sara does, Divit is always better off to enter and this is his dominant strategy.

A set of strategies is a Nash equilibrium if, when all other players use best-response strategies, no player can obtain a higher payoff by choosing a different strategy. This game does not have a dominant strategy solution, but it still has a Nash equilibrium: Because Sara knows that Divit will enter, Sara will stay out of the market, and neither wants to change this strategy.

The subsidy does not change Divit's dominant strategy: He enters the market. With the subsidy, Sara's payoff rises by 250 if she enters the market: She earns 50 if Divit also enters the market and 750 if he does not. Thus, entering the market becomes Sara's dominant strategy. Since this game has a dominant strategy solution with a subsidy to Sara, this solution is also a Nash equilibrium. Both Divit and Sara will start new businesses in the region.

1.12 A dominant strategy is one that produces a higher payoff to all possible strategies (best responses) that the other player might use.

a. If Firm 2 charges the low price, Firm 1's payoff is higher if it charges a low price. If Firm 2 charges a high price, Firm 1's payoff is higher if it charges a high price. Thus, Firm 1 does not have a dominant strategy.

b. If Firm 1 charges the low price, Firm 2's payoff is higher if it charges a high price. If Firm 1 charges a high price, Firm 2's payoff is higher if it charges a low price. Thus, Firm 2 does not have a dominant strategy.

A set of strategies is a Nash equilibrium if, when all other players use best response strategies, no player can obtain a higher payoff by choosing a different strategy. This game does not have a dominant strategy solution, nor a pure-strategy Nash equilibrium. None of the outcomes are Nash equilibria because at least one player could benefit by choosing a different strategy.

However, there is a mixed-strategy Nash equilibrium.

- If Firm 1 charges the low price, its expected utility is $EU_1^L = \delta_2(49) + (1 - \delta_2)(48)$ where δ_2 is the probability that Firm 2 also charges a low price. If Firm 1 charges a high price, its expected utility is $EU_1^H = \delta_2(47) + (1 - \delta_2)(52)$. Firm 1 is indifferent between the two if $\delta_2 = 0.67$.
- If Firm 2 charges the low price, its expected utility is $EU_2^L = \delta_1(47) + (1 - \delta_1)(57)$ where δ_1 is the probability that Firm 1 also charges a low price. If Firm 2 charges a high price, its expected

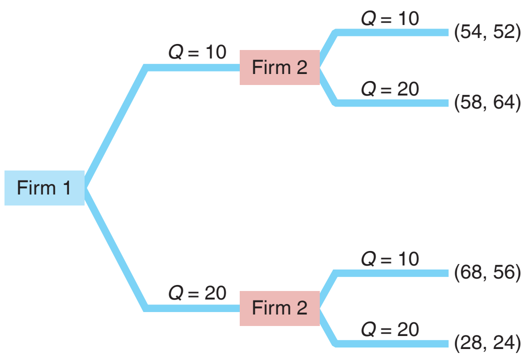
utility is $EU_2^H = \delta_1(52) + (1 - \delta_1)(55)$. Firm 2 is indifferent between the two if $\delta_1 = 0.29$.

Thus, the mixed-strategy Nash equilibrium is for Firm 1 to charge the low price 29% of the time and for Firm 2 to charge the low price 67% of the time.

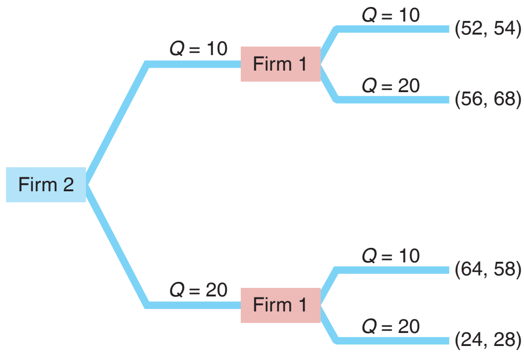
2.2 If the game is repeated a finite number of times, the outcome will yield the noncooperative solution if the players are fully rational. That is, $q_u = 64$ and $q_a = 64$ in each period. The same outcome occurs if both firms know that one firm cares only about current period profits.

3.1 a. There are two Nash equilibria (the off diagonals). If either firm produces 20 while the other produces 10, neither has an incentive to change strategies, given the strategy of the other.

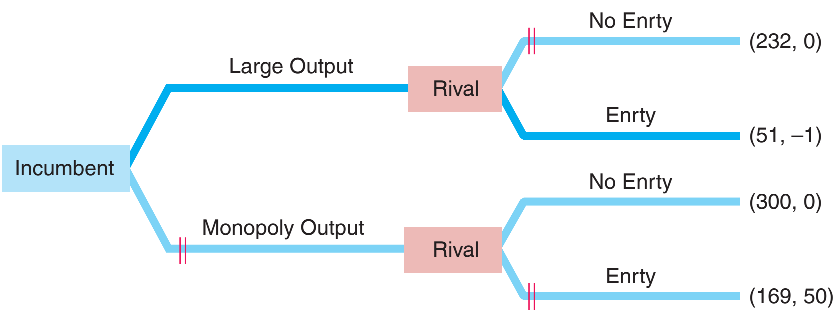
b. If Firm 1 moves first and picks $Q = 10$, then Firm 2 will pick $Q = 20$, and Firm 1's profit will be \$58. If Firm 1 moves first and picks $Q = 20$, then Firm 2 will pick $Q = 10$, and Firm 1's profit will be \$68. Profit of \$68 is greater than \$58, so Firm 1 will pick $Q = 20$, and Firm 2 will pick $Q = 10$.



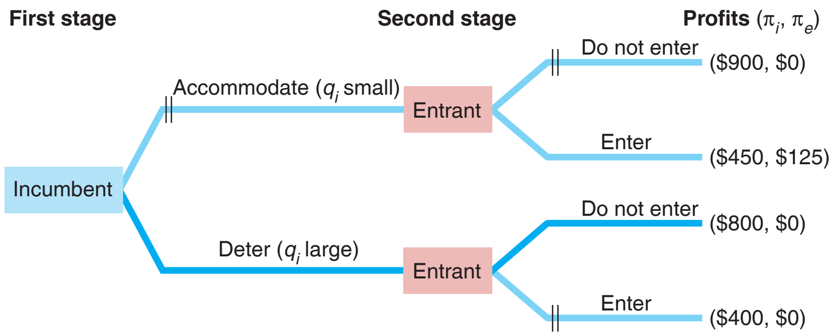
c. If Firm 2 moves first and picks $Q = 10$, then Firm 1 will pick $Q = 20$, and Firm 2's profit will be \$56. If Firm 2 moves first and picks $Q = 20$, then Firm 1 will pick $Q = 10$, and Firm 2's profit will be \$64. Profit of \$64 is greater than \$56, so Firm 2 will pick $Q = 20$, and Firm 1 will pick $Q = 10$.



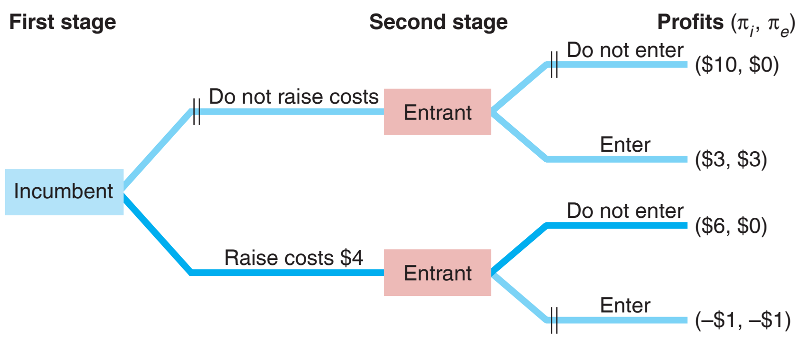
3.11 The game tree illustrates why the incumbent may install the robotic arms to discourage entry even though its total cost rises. If the incumbent fears that a rival is poised to enter, it invests to discourage entry. The incumbent can invest in equipment that lowers its marginal cost. With the lowered marginal cost, it is credible that the incumbent will produce larger quantities of output, which discourages entry. The incumbent's monopoly (no-entry) profit drops from \$900 to \$500 if it makes the investment because the investment raises its total cost. If the incumbent doesn't buy the robotic arms, the rival enters because it makes \$300 by entering and nothing if it stays out of the market. With entry, the incumbent's profit is \$400. With the investment, the rival loses \$36 if it enters, so it stays out of the market, losing nothing. Because of the investment, the incumbent earns \$500. Nonetheless, earning \$500 is better than earning only \$400, so the incumbent invests.



3.12 The incumbent firm has a *first-mover advantage*, as the game tree illustrates. Moving first allows the incumbent or leader firm to *commit* to producing a relatively large quantity. If the incumbent does not make a commitment before its rival enters, entry occurs and the incumbent earns a relatively low profit. By committing to produce such a large output level that the potential entrant decides not to enter because it cannot make a positive profit, the incumbent's commitment discourages entry. Moving backward in time (moving to the left in the diagram), we examine the incumbent's choice. If the incumbent commits to the small quantity, its rival enters and the incumbent earns \$450. If the incumbent commits to the larger quantity, its rival does not enter and the incumbent earns \$800. Clearly, the incumbent should commit to the larger quantity because it earns a larger profit and the potential entrant chooses to stay out of the market. Their chosen paths are identified by the darker blue in the figure.



3.13 It is worth more to the monopoly to keep the potential entrant out than it is worth to the potential entrant to enter, as the figure shows. Before the pollution-control device requirement, the entrant would pay up to \$3 to enter, whereas the incumbent would pay up to $\pi_m - \pi_d = \$7$ to exclude the potential entrant. The incumbent's profit is \$6 if entry does not occur, and it loses \$1 if entry occurs. Because the new firm would lose \$1 if it enters, it does not enter. Thus, the incumbent has an incentive to raise costs by \$4 to both firms. The incumbent's profit is \$6 if it raises costs rather than \$3 if it does not.



Chapter 15

1.2 The answer is given in Appendix 15A.

1.4 The competitive firm's marginal revenue of labor is $MRP_L = 2pK$.

1.5 A competitive firm hires labor, L, up to the point at which the marginal revenue product of labor equals the marginal cost of employing that worker, which is the wage, w.

$$pMP_L = w$$
$$10(9 - L) = 10$$
$$L = 8$$

With the specific tax, the price the firm receives falls to $p - 5$ and the competitive firm hires

$$(p - 5)MP_L = w$$
$$5(9 - L) = 10$$
$$L = 7$$

Thus, the specific tax reduces the firm's demand for labor from 8 to 7 workers.

Chapter 16

1.5 An individual with a zero discount rate views current and future consumption as equally attractive. An individual with an infinite discount rate cares only about current consumption and puts no value on future consumption.

1.13 If the interest rate is set in real terms, putting \$2,000 in the bank today results in an annual flow of \$200 in real terms. If the interest rate is set in nominal terms, the real payment will shrink over time, so you cannot receive a real payment of \$200 annually. (If the nominal rate were set at 15.5%, an initial \$2,000 investment would ensure an annual flow of \$200 in real terms.)

1.14 The present value of your payment today is €500. With inflation, €1 saved today can purchase $(1 + i)/(1 + g)$ euros next year, where i is the nominal interest rate and g is the rate of inflation. Thus, the present value of your €500 payment next year is €495.15 today:

$$500\left(\frac{1+g}{1+i}\right) = 500\left(\frac{1.02}{1.03}\right) = 495.15.$$

The present value of your new TV is therefore €995.15.

2.1 The present value of the contract with Retailer 1 is:

$$PV_1 = 20{,}000 + \frac{80{,}000}{1+\delta},$$

where δ is the discount rate. When $\delta = 0.03$, $PV_1 = 97{,}670$, and when $\delta = 0.07$, $PV_1 = 94{,}766$. The present value of the contract with Retailer 2 is:

$$PV_2 = 20{,}000 + \frac{40{,}000}{1+\delta} + \frac{42{,}000}{(1+\delta)^2}.$$

When $\delta = 0.03$, $PV_2 = 98{,}424$, and when $\delta = 0.07$, $PV_2 = 94{,}068$.

Comparing the two, the parts supplier should accept the contract with Retailer 2 if its discount rate is 3%, but the contract with Retailer 1 if its discount rate is 7%.

2.8 Currently, you are buying 600 gallons of gas at a cost of \$1,200 per year. With a more gas-efficient car, you would spend only \$600 per year, saving \$600 per year in gas payments. If we assume that these payments are made at the end of each year, the present value of this savings for five years is \$2,580 at a 5% annual interest rate and \$2,280 at 10% (using Table 16.4). The present value of the amount you must spend to buy the car in five years is \$6,240 at 5% and \$4,960 at 10% (using Table 16.3). Thus, the present value of the additional cost of buying now rather than later is \$1,760 (= \$8,000 − \$6,240) at 5% and \$3,040 at 10%. The benefit from buying now is the present value of the reduced gas payments. The cost is the present value of the additional cost of buying the car sooner rather than later. At 5%, the benefit is \$2,580 and the cost is \$1,760, so you should buy now. However, at 10%, the benefit, \$2,280, is less than the cost, \$3,040, so you should buy later.

2.18 Solving for *irr*, we find that *irr* equals 1 or 9. This approach fails to give us a unique solution, so we should use the *NPV* approach instead.

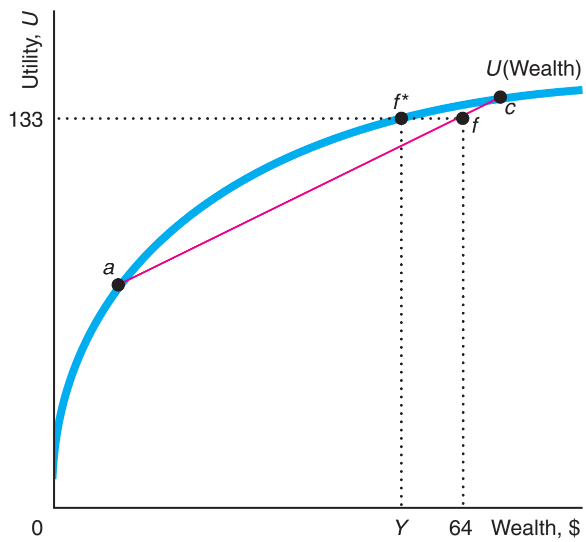
Chapter 17

1.2 Assuming that the painting is not insured against fire, its expected value is $\$550 = (0.2 \times \$1{,}000) + (0.1 \times \$0) + (0.7 \times \$500)$.

1.3 The expected value is $EV = 0.2 \times 15 \times 50 + 0.8 \times 25 \times 50 = €1{,}150$.
The variance is $V = 0.2(750 - 1{,}150)^2 + 0.8(1{,}250 - 1{,}150)^2 = 40{,}000$.

1.6 From a monetary viewpoint, the expected value of having to pay a parking ticket is $EV = \theta F$, where θ is the probability of being caught and fined, and F is the amount of the penalty. Increasing the probability of being caught through greater enforcement or the amount of the fine charged would both make the expected punishment more severe and work to deter illegal parking. However, since it is essentially costless, a city government would likely prefer to increase the fine. Raising the probability of being caught would entail substantial court costs as well as the costs of hiring and paying for additional parking enforcement officers.

2.5 As the graph shows, Irma's expected utility of 133 at point *f* (where her expected wealth is \$64) is the same as her utility from a certain wealth of *Y*.



2.9 Hugo's expected wealth is $EW = \left(\frac{2}{3} \times 144\right) + \left(\frac{1}{3} \times 225\right) = 96 + 75 = 171.$

His expected utility is

$$\begin{aligned} EU &= \left[\tfrac{2}{3} \times U(144)\right] + \left[\tfrac{1}{3} \times U(225)\right] \\ &= \left[\tfrac{2}{3} \times \sqrt{144}\right] + \left[\tfrac{1}{3} \times \sqrt{225}\right] \\ &= \left[\tfrac{2}{3} \times 12\right] + \left[\tfrac{1}{3} \times 15\right] = 13. \end{aligned}$$

He would pay up to an amount P to avoid bearing the risk, where $U(EW - P)$ equals his expected utility from the risky stock, EU. That is, $U(EW - P) = U(171 - P) = \sqrt{171 - P} = 13 = EU$. Squaring both sides, we find that $171 - P = 169$, or $P = 2$. That is, Hugo would accept an offer for his stock today of \$169 (or more), which reflects a risk premium of \$2.

3.3 a. $EU(\text{No Insurance}) = 0.2U(90{,}000) + 0.8U(160{,}000) = 240 + 1{,}280 = 1{,}520$. $U(\text{Insurance}) = U(160{,}000 - 15{,}000) = 1{,}523.15$. Because $U(\text{Insurance}) > EU(\text{No Insurance})$, this household should buy the insurance.

b. The fair price is the expected value of the payout, which is $0.2 \times 70{,}000 = 14{,}000$.

c. The maximum the household would pay for this insurance is the price that would give it the same utility as not buying the insurance. If this price is p, then $U(160{,}000 - p) = EU(\text{No Insurance}) = 1{,}520$. Therefore, $4(160{,}000 - p)^{0.5} = 1{,}520$ or $160{,}000 - p = (1{,}520/4)^2 = 144{,}400$. It follows that $p = 15{,}600$.

4.2 If they were married, Andy would receive half the potential earnings whether they stayed married or not. As a result, Andy will receive \$12,000 in present-value terms from Kim's additional earnings. Because the returns to the investment exceed the cost, Andy will make this investment (unless a better investment is available). However, if they stay unmarried and split, Andy's expected return on the investment is the probability of staying together, $\frac{1}{2}$, times Kim's half of the returns if they stay together, \$12,000. Thus, Andy's expected return on the investment, \$6,000, is less than the cost of the education, so Andy is unwilling to make that investment (regardless of other investment opportunities).

5.4 Either Joe or Sue might be overconfident, but the question provides no information to indicate that is the reason for their different behaviors. Prospect theory might provide a good explanation. Under prospect theory, people are risk seeking in the domain of losses, so Joe is willing to take a substantial risk on the last race. Sue is in the domain of gains and is therefore fundamentally risk averse. A small gamble is still fun for her (see the Application "Gambling"), but she does not want to take a large risk.

Chapter 18

3.2 As Figure 18.2 shows, a specific tax of \$84 per ton of output or per unit of emissions (gunk) leads to the social optimum.

3.6 Use the model in Appendix 18A to determine the equilibrium if the marginal harm of gunk is $MC^g = \$84$ (instead of Equation 18A.3). We care only about the marginal harm of gunk at the social optimum, which we know is $MC^g = \$84$ (because it is the same at every level of output). That is the same marginal cost as in the table at the end of Appendix 18A. Thus, the social optimum is the same as in that example (and no algebra is necessary). Using algebra, we set the demand curve equal to the new social marginal cost, $MC^2 = c + dQ + 84$, and we find that the socially optimal quantity is

$$\begin{aligned} Q_s &= (a - c - 84)/(b + d) \\ &= (450 - 30 - 84)/(2 + 2) = 84. \end{aligned}$$

5.2 The government charges Alice's auto body shop a tax of \$200 if it works on one car and \$400 if it works on two cars per hour, which equals the marginal harm to the tea house. Alice makes \$100 after tax if she works on one car and \$0 if she works on two cars, so the social optimum is achieved. If they can bargain, the tea house offers Alice an amount between \$100 and \$200 to shut down, which maximizes their after-tax joint profits but not social welfare. Thus, the "optimal" tax may not maximize social welfare if the parties can bargain.

6.11 If the industry were advertising optimally, the marginal revenue from advertising would equal the marginal cost. Since $MC = 2\% < 3\% = MR$, the agreed-upon level of advertising is less than optimal, and advertising is underfunded. This may be because some franchisees who benefit from collective advertising do not want to pay for it and would like to free ride. This is why the fee is usually mandatory.

7.1 You can use several ways to demonstrate that welfare can go up despite the pollution. For example, you could redraw panel b with flatter supply curves so that area C became smaller than A (area A remains unchanged). Similarly, if the marginal pollution harm is very small, then we are very close to the no-distortion case, so that welfare will increase.

7.2 Going from no trade to free trade, consumers gain areas B and C, while domestic firms lose B. Thus, if consumers give firms an amount between B and $B + C$, both groups will be better off than with no trade.

Chapter 19

1.7 The most that risk neutral buyers would be willing to pay for a used dress of unknown quality is the average price: $p = \delta p_L + (1 - \delta)p_G = 30\delta + 80(1 - \delta) = 80 - 50\delta$. All dresses will sell if the market price exceeds C\$70: in this case, $\delta \leq 0.2$. Lemons will sell if the market price is between C\$30 and C\$70: in this case, $0.2 < \delta \leq 1$. No dresses will sell if the market price is less than C\$20, but this cannot happen because δ cannot be larger than one ($\delta \not> 1$).

2.2 If an insurance company omits any key factor in setting the premium it charges, buying insurance would be unattractive for people who are below average risk due to that factor, but attractive for people who have a higher risk. In this example, because insurance costs would not vary with soil type, buying insurance would be less attractive for houses on good soil and relatively more attractive for houses on bad soil. So relatively more homeowners with houses on poor soil would buy insurance at too low a price, resulting in a disproportionate number of bad outcomes for the insurer following a large earthquake. This omission (hidden inaction) is a moral hazard.

2.3 Brand names allow consumers to identify a particular company's product in the future. If a mushroom company expects to remain in business over time, it would be foolish to brand its product if its mushrooms are of inferior quality. (Just ask Babar's grandfather.) Thus, all else the same, we would expect branded mushrooms to be of higher quality than unbranded ones.

4.1 If almost all consumers know the true prices, and all but one firm charges the full-information competitive price, then it does not pay for a firm to set a high price. It gains a little from charging ignorant consumers the high price, but it sells to no informed customers. Thus, the full-information competitive price is charged in this market.

Chapter 20

1.1 This arrangement would likely increase the cost of the overall bill due to the moral hazard of an informed party (the professional) taking a hidden action (referral fee arrangement) that harms a less-informed party (the client).

1.5 By making this commitment, the bookstore may be trying to assure students who cannot judge how quickly the textbook will deteriorate that it is durable enough to maintain at least a certain value in the future. The bookstore is trying to eliminate asymmetric information to increase the demand for its product relative to alternatives such as e-textbooks.

2.4 a. If Arnie is paid a fixed wage of 10, then Arnie provides low effort. Any additional effort would be costly to him and would not increase his wage and therefore be a net loss.

b. Given that they share profit equally, Arnie chooses medium effort. With medium effort the expected profit is $(0.5 \times 40) + (0.5 \times 80) = 60$. Arnie gets 50% or 30 and subtracts the cost of effort, 10, yielding a net gain of 20. With either low effort or high effort, Arnie's net gain is only 15.

c. Arnie receives 20 under the profit-sharing contract and 10 with a fixed wage, so he prefers the fixed wage. Priscilla has an expected value of $30 - 10 = 20$ with a fixed wage and $0.5(60) = 30$ under profit sharing, so she also prefers profit sharing.

2.7 The contractee might refuse to accept a perfectly good product to negotiate a lower price for it.

Since contracts are enforceable in court, one option would be for the contractor to seek an objective determination of whether or not the terms of the contract have been satisfied. However, this may be costly and difficult to verify. A second option would be for the contractor to accept the lower price in the short run (something is better than nothing) and cease dealing with the contractee in the long run. However, this would prevent transactions that would otherwise be value-generating for the two firms. A third option would focus on eliminating the potential for opportunistic behavior (moral hazard) on the part of the contractee by eliminating asymmetric information (and preventing adverse selection). This might be accomplished through contractual provisions that require information sharing and contractee approval at specified stages of development and production process.

2.8 A partner who works an extra hour bears the full opportunity cost of this extra hour but gets only half the marginal benefit from the extra business profit. The opportunity cost of extra time spent at the store is the partner's best alternative use of time. A partner could earn money working for someone else or use the time to have fun. Because a partner bears the full marginal cost but gets only half the marginal benefit (the extra business profit) from an extra hour of work, each partner works only up to the point at which the marginal cost equals half the marginal benefit. Thus, each has an incentive to put in less effort than the level that maximizes their joint profit, where the marginal cost equals the marginal benefit.

2.10 This agreement may lead to more trips than if each paid the entire amount when flying. Under the sharing plan, whichever of them most wants to visit the other more receives the full marginal benefit of one more visit while having to pay only half the marginal cost.

3.2 The minimum bond that deters stealing is \$2,500 (= \$500/0.2).

Sources for Challenges and Applications

Chapter 1

Applications

Twinkie Tax: Jacobson and Brownell (2000). Bruce Bartlett, "The Big Food Tax," *National Review Online*, April 3, 2002. Chouinard et al. (2007). **http://www.nytimes.com/2014/07/30/opinion/mark-bittman-introducing-the-national-soda-tax.html.** Emily Dugdale, "Soda Tax Brings Needed Cash to Gardening Program," *Berkeleyside*, June 8, 2015. Kate Kilpatrick, "Taxing Soda, Saving Lives," August 19, 2015, **http://www.projects.aljazeera.com/2015/08/mexico-obesity/soda-tax.html.**

Income Threshold Model and China: "Next in Line: Chinese Consumers," *Economist*, 326 (7795), January 23, 1993:66–67. Jeff Pelline, "U.S. Businesses Pour into China," *San Francisco Chronicle*, May 17, 1994:B1–B2. *China Statistical Yearbook* (Beijing: China Statistical Publishing House, 2000). **http://www.uschina.org/info/forecast/2007/foreign-investment.html.** Jiang Wei, "FDI Doubles Despite Tax Concerns," *China Daily*, February 19, 2008. Keith Bradsher, "With First Car, a New Life in China," *New York Times*, April 24, 2008. "China January FDI Grows at Strongest Pace in Four Years," Reuters, February 15, 2015. **http://www.worldometers.info/cars** (viewed July 30, 2016). **http://www.bridgingthegapresearch.org/research/sodasnack_taxes/** (viewed July 30, 2016). **http://www.tradingeconomics.com/china/foreign-direct-investment** (viewed July 30, 2016).

Chapter 2

Challenge

Quantities and Prices of Genetically Modified Foods: **http://www.redgreenandblue.org/2011/03/23/is-europes-ban-on-monsantos-gmo-crops-illegal. http://www.guardian.co.uk/environment/2012/feb/08/industry-claims-rise-gm-crops.** GMO-Compass, **http://www.gmo-compass.org/eng/agri_biotechnology/gmo_planting/257.global_gm_planting_2013.html.** *ISAAA Brief 49-2014: Executive Summary*, March 25, 2014, **http://www.isaaa.org/resources/publications/briefs/49/executivesummary/default.asp.**

Cary Funk and Lee Rainie, "Public and Scientists' Views on Science and Society," **http://www.pewinternet.org/2015/01/29/public-and-scientists-views-on-science-and-society. http://www.justlabelit.org/right-to-know-center/labeling-around-the-world** (viewed on May 26, 2016). **http://www.genewatch.org/sub-532326** (viewed May 26, 2016). National Academies of Science, *Genetically Engineered Crops: Experiences and Prospects*, **http://www.nap.edu/23395**, 2016.

Applications

Calorie Counting: Bollinger, Leslie, and Sorensen (2011). Dingman et al. (2015). Stewart and Mentzer Morrison (2015).

Aggregating Corn Demand Curves: McPhail and Babcock (2012). **http://www.ncga.com/worldofcorn** (viewed May 11, 2013).

Occupational Licensing: **http://www.ncbex.org/publications/statistics/**, 2016; **http://www.admissions.calbar.ca.gov/Examinations/Statistics.aspx**, 2014. Kleiner and Krueger (2013). Kleiner and Park (2014). Kleiner (2015). *Occupational Licensing: A Framework for Policymakers*, The White House, July 2015. **http://blogs.bls.gov/blog/2016/04/15/why-this-counts-new-timely-data-on-professional-certifications-and-licenses/**, April 15, 2016. Zumbrun, Josh, "Occupational Licenses May Be Bad for the Economy, But Good for Workers Who Have Them," **http://blogs.wsj.com/economics/2016/04/18/occupational-licenses-may-be-bad-for-the-economy-but-good-for-workers-who-have-them/**, April 18,

2016. **http://www.bls.gov/cps/certifications-and-licenses.htm** (viewed May 26, 2016). Patricia Cohen, "Moving to Arizona Soon? You Might Need a License," June 17, 2016, **http://www.nytimes.com/2016/06/18/business/economy/job-licenses.html**.

Venezuelan Price Ceilings and Shortages: William Neuman, "With Venezuelan Food Shortages, Some Blame Price Controls," *New York Times*, April 20, 2012. William Neuman, "Protests Swell in Venezuela as Places to Rally Disappear," *New York Times*, February 20, 2014. Lopez, Linette, "This Is a Food Line in Venezuela," *Business Inside*, February 28, 2014. Francisco Toro, "How Venezuela Turns Butter Vendors Into Currency Manipulators," *New Republic*, March 4, 2014. S.B., "The Dividing Line, *Economist*, March 10, 2014. Rafael Romo, "Fistfights Amid Long Bread Lines in Venezuela," **http://www.cnn.com**, May 2014. "Smugglers Turn Venezuelan Crisis Into Colombian Cash Out," March 31, 2015, **http://www.customstoday.com.pk/smugglers-turn-venezuelan-crisis-into-colombian-cash-out-2**. Rafael Romo, "$755 for a Box of Condoms? No Protection from Shortages in Venezuela," February 6, 2015, **http://www.cnn.com**. Joshua Keating, "Venezuela Is Running Out of Toilet Paper," **http://www.slate.com**, February 25, 2015. Ian Bremmer, "Venezuela Goes from Bad to Catastrophe,**http://time.com/4348972/venezuela-goes-from-bad-to-catastrophe/**, June 6, 2016. Nicholas Casey, "Venezuelans Ransack Stores as Hunger Grips the Nation," *New York Times*, June 19, 2016. Anatoly Kurmanaev, "Amid Shortages, Venezuela Frees Prices," *Wall Street Journal*, October 15–16, 2016.

Chapter 3

Challenge

Who Pays the Gasoline Tax?: Chouinard and Perloff (2007). Scott Horsley, "Obama Says Illinois' Gas Tax Holiday Didn't Work," NPR, May 9, 2008. Josh Voorhees, "Oberstar Proposes Future Gas Tax Hike," *E&ENews*, March 26, 2010. "Shuler Wants Gas Tax Break as Asheville Prices Rise," **citizen-times.com**, May 4, 2011. **http://www.api.org/oil-and-natural-gas/consumer-information/motor-fuel-taxes/gasoline-tax** (viewed May 26, 2016). **http://www.eia.gov/dnav/pet/pet_pnp_cap1_dcu_nus_a.htm** (viewed May 29, 2016). **http://www.iea.org/statistics/monthlystatistics/monthlyoilprices/**, April 2016.

Applications

The Demand Elasticities for Google Play and Apple Apps: Ghose and Han (2014). **http://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/** (viewed August 30, 2016).

Amazon Prime: Bensinger, Greg, "Amazon Raises Prime Subscription Price to $99 a Year," *Wall Street Journal*, March 13, 2014.

Anti-Smoking Policies May Reduce Drunk Driving: Krauss et al. (2014). Alan Mozes, "Anti-Smoking Policies May Also Curtail Drinking," *Healthday*, Tuesday, Sept. 23, 2014.

Oil Drilling in the Arctic National Wildlife Refuge: United States Geological Survey, "Arctic National Wildlife Refuge, 1002 Area, Petroleum Assessment, 1998, Including Economic Analysis," **http://pubs.usgs.gov/fs/fs-0028-01/fs-0028-01.pdf**, April 2001. Energy Information Administration, "The Effects of Alaska Oil and Natural Gas Provisions of H.R. 4 and S. 1776 on U.S. Energy Markets," February 2002. Borenstein (2005). Phil Taylor, "GOP Governors Urge Expanded Offshore Drilling, Opening ANWR," *E&E News PM*, August 22, 2012. **http://www.tonto.eia.doe.gov** (viewed May 11, 2013). **http://www.eia.gov** (viewed May 11, 2013).

Taxes to Prevent Obesity: Jacobson and Brownell (2000). Pierre Lemieux, "It's the Fat Police," *National Post,* April 6, 2002. Helen Tobler, "Call for Tax War on Obesity," *Australian IT,* August 16, 2002. "Soda Pop to Be Banned in L.A. Schools," **http://www.CBSNEWS.com**, August 28, 2002. Chouinard et al. (2007). "'Fat Tax' Proposed for Obese in Britain," **http://www.sify.com/news/fat-tax-proposed-for-obese-in-britain-news-national-jegmLFabedd.html**, February 25, 2008. Suzanne Daley, "Hungary Tries a Dash of Taxes to Promote Healthier Eating Habits," *New York Times*, March 2, 2013.

Subsidizing Ethanol: "Ethanol Subsidies Are $82 per Barrel of Replaced Gasoline, Says Incendiary Baker Institute Report," **http://www.biofuelsdigest.com/blog2/2010/01/07/ethanol-subsidies-are-82-per-barrel-of-replaced-gasoline-says-incendiary-baker-institute-report**, July 1, 2010. Robert Pear, "After Three Decades, Tax Credit for Ethanol Expires," *New York Times*, January 1, 2012. McPhail and Babcock (2012). **http://www.ers.usda.gov/topics/crops/corn/policy.aspx** (viewed May 27, 2013).

Chapter 4

Challenge

Why Americans Buy More Ebooks Than Do Germans: Biba, "Publishing Business Conference: International Trends in Ebook Consumption," March 20, 2012, **http://www.teleread.com**. Aaron Wiener, "We Read Best on Paper, Cultural Resistance Hobbles German E-Book Market," April 13, 2012, **http://www.spiegel.de**. Caroline Winter, "The Story Behind Germany's Scant E-Book Sales," *Bloomberg Businessweek*, April 19, 2012. Rüdiger Wischenbart et al., *The Global ebook Market*, O'Reilly:2014. Philip Jones, "Global E-book Report highlights unequal markets," *The Bookseller*, May 6, 2015, **http://www.thebookseller.com/news/global-e-book-report-highlights-unequal-markets**. **http://www.vatlive.com/vat-rates/european-vat-rates/eu-vat-rates/** (viewed June 4, 2016). **http://www.statista.com/statistics/437152/book-market-print-and-e-book-revenue-in-germany/** (viewed September 1, 2016). **http://www.publishersweekly.com/pw/by-topic/industry-news/bea/article/70350-bea-2016-e-book-sales-fell-13-in-2015-nielsen-reports.html** (viewed September 1, 2016).

Applications

You Can't Have Too Much Money: *Businessweek*, February 28, 2005, p. 13. "Money Cannot Buy Happiness but It Helps," *Gulf Times*, May 30, 2010. "Annual Happiness Index Again Finds One-Third of Americans Very Happy," Harris Interactive, June 22, 2011, **http://www.digitaljournal.com**. Emily Alpert, "Happiness Tops in Denmark, Lowest in Togo, Study Says," *Los Angeles Times*, April 2, 2012. **http://www.businessinsider.com/23-of-us-adults-would-jump-at-the-opportunity-to-go-under-the-knife-if-it-was-cheaper-2012-1**, January 18, 2012. Helliwell et al. (2013). Stevenson and Wolfers (2013). Easterlin (2015). Easterlin (2016).

Indifference Curves Between Food and Clothing: Eastwood and Craven (1981).

Substituting Alcohol for Marijuana: Crost and Guerrero (2012).

Benefiting from Food Stamps: Whitmore (2002). "U.S. Converting Food Stamps into Debit-Card Benefits," *Chattanooga Times Free Press,* June 23, 2004. U.S. Department of Agriculture, *Characteristics of Food Stamp Households: Fiscal Year 2007*, September 2008. **http://www.fns.usda.gov/snap/faqs.htm** (viewed August 31, 2009). Hoynes and Schanzenbach (2009). "State Supplemental Nutrition Program Participation Rates in 2007," **http://www.usda.gov/documents/2007SNAPparticipationreport.pdf**, 2009. "New FRAC Report on 24 Cities Finds Only 67 Percent of Eligible Urban Households Participate in SNAP/Food Stamp Program," **http://www.frac.org/urban-snapfood-stamp-report**, September 9, 2009. Center on Budget and Policy Priorities, "Policy Basics: Introduction to the Supplemental Nutrition Assistance Program (SNAP)," March 28, 2013. Bruich (2014). **http://www.fns.usda.gov/pd/supplemental-nutrition-assistance-program-snap** (viewed June 4, 2016).

Opt In Versus Opt Out: Madrian and Shea (2001, 2002). **http://www.aon.mediaroom.com/index.php?s=43&item=2285**, May 24, 2011. **http://www.arielinvestments.com/401k-Study-2012**, May 13, 2012. **http://www.bls.gov/ncs/ebs/benefits/2015/ebbl0057.pdf**.

Unaware of Taxes: Chetty et al. (2009). Goldin and Homonoff (2013). Zheng et al. (2013).

Chapter 5

Challenge

Per-Hour Versus Lump-Sum Childcare Subsidies: Blau and Tekin (2001). Elena Cherney, "Giving Day-Care Cash to Stay-at-Home Parents Sounds Like Politics to Some Canadians," *Wall Street Journal*, July 3, 2006:A2. Tekin (2007). Currie and Gahvari (2008). **http://www.hss.gov.yk.ca/pdf/childcaresubsidy_en.pdf**, October, 2015. **http://www.acf.hhs.gov/programs/ccb/ccdf/factsheet.htm** (viewed June 7, 2016). **https://www.oecd.org/els/soc/PF3_1_Public_spending_on_childcare_and_early_education.pdf** (viewed June 7, 2016). **http://www.oecd.org/els/soc/PF1_1_Public_spending_on_family_benefits_Oct2013.xls** (viewed June 7, 2016).

Applications

Smoking Versus Eating and Phoning: Busch et al. (2004). Labonne and Chase (2011). Nikaj and Chaloupka (2013). Pesko et al. (2016). **http://www.who.int/mediacentre/factsheets/fs339/en/** (viewed June 6, 2016). **http://ash.org/resources/**

tobacco-statistics-facts/ (viewed June 6, 2016). **http://www.cdc.gov/tobacco/data_statistics/fact_sheets/adult_data/cig_smoking/** (viewed June 6, 2016).

Fast-Food Engel Curve: Kim and Leigh (2011).

What's Your Smart Phone Worth to You?: David Dean et al., *The Internet Economy in the G-20*, Boston Consulting Group, 2012, **http://www.bcgperspectives.com/Images/Internet_Economy_G20_tcm80-100409.pdf**. Sarah Kidner, "What Would You Give Up for the Net?", **http://conversation.which.co.uk/technology/what-would-you-give-up-for-internet-chocolate-coffee-sex**, March 22, 2012. Wolfgang Bock et al., *The Growth of the Global Mobile Internet Economy*, Boston Consulting Group, 2015, **http://www.transactives.com/Portals/24/docs/GPSS%202015/Growth_of_the_Global_Mobile_Internet_Economy_Feb_2015_tcm80-181599.pdf.**

Paying Employees to Relocate: Herod (2008). **http://www.mthink.mercer.com/commitment-to-global-mobility-remains-strong**, April 8, 2013. Mercer, "Global Mobility: The New Business Imperative," **http://www.imercer.com/content/IGM-v3n1-intl-assignments-tackling-compensation.aspx**, 2013. Expatistan, 2015, **http://www.expatistan.com/cost-of-living/comparison/seattle/london** (viewed on April 22, 2015). **http://www.workforce.com/articles/20593-special-report-relocation-sector-keeps-moving-right-along** (viewed June 6, 2016). **http://globalmobilitytrends.brookfieldgrs.com/** (viewed June 8, 2016).

Working After Winning the Lottery: Imbens et al. (2001). Cesarini et al. (2015). Picchio et al. (2015).

Chapter 6

Challenge

More Productive Workers During Downturns: Flath (2011). Data files underlying Susan Fleck et al., "The Compensation-Productivity Gap: A Visual Essay," *Monthly Labor Review*, January 2011. Lazear et al. (2016).

Applications

Chinese State-Owned Enterprises: Gao Xu, "State-Owned Enterprises in China: How Big Are They?", January 19, 2010, **http://www.blogs.worldbank.org/eastasiapacific/state-owned-enterprises-in-china-how-big-are-they**. National Bureau of Statistics of China, *China Statistical Yearbook*, 2014. Chang-Tai Hsieh and Zheng (Michael) Song, "Grasp the Large, Let Go of the Small: The Transformation of the State Sector in China," NBER Working Paper, 2015.

Malthus and the Green Revolution: Norman Borlaug, "Nobel Lecture," **http://www.nobelprize.org**, December 11, 1970. Gregg Easterbrook, "Forgotten Benefactor of Humanity," *Atlantic Monthly*, February 1997. Brander and Taylor (1998). Alan Barkema, "Ag Biotech," *The Main Street Economist*, October 2000. "Biotechnology and the Green Revolution: Interview with Norman Borlaug," **http://www.actionBioscience.org**, November 2002. United Nations, *Millennium Development Goals Report*, New York, 2008. Food and Agriculture Organization of the United Nations, *The State of Food Insecurity in the World*, 2015. **http://www.fao.org/NEWS/2000/000704-e.htm** (viewed on June 11, 2016). **http://www.ers.usda.gov/data-products/agricultural-productivity-in-the-us.aspx#28247** (viewed June 11, 2016).

A Semiconductor Integrated Circuit Isoquant: Nile Hatch, personal communications. Roy Mallory, personal communications.

Returns to Scale in Various Industries: Hsieh (1995). Flath (2011). Fox and Smeets (2011). Devine et al. (2012). Hossain et al. (2012). Zhang et al. (2012).

Robots and the Food You Eat: Ken Teh, "Robot Waiters in China Never Lose Patience," *Associated Press*, December 22, 2010, **http://www.washingtontimes.com**. Ilan Brat, "Robots Step Into New Planting, Harvesting Roles" *Wall Street Journal*, April 23, 2015, **http://www.wsj.com**. Aviva Rutkin, "Harvey, the Robot Farmer Fixing the US Labour Shortage," *New Scientist*, June 18, 2014, **http://www.newscientist.com**. Jesse McKinley, "With Farm Robotics, the Cows Decide When It's Milking Time," *New York Times*, April 22, 2014. Dyllan Furness, "Hate Ordering Fried Chicken from Human Beings? KFC's New Restaurant Has You Covered," **http://www.digitaltrends.com/cool-tech/kfc-ai-robot-restaurant/**, May 16, 2016. Chuck Thompson and Elaine Yu, "New Order? China Restaurant Debuts Robot Waiters," **http://www.cnn.com/2016/04/19/travel/china-robot-waiters/**, April 20, 2016.

A Good Boss Raises Productivity: Lazear et al. (2015).

Chapter 7

Challenge

Technology Choice at Home Versus Abroad: **http://www.semiconductors.org/industry_statistics/historical_billing_reports** (viewed June13, 2016).

Applications

The Opportunity Cost of an MBA: Edmonston, Peter, "In Tough Times, M.B.A. Applications May Be an Economic Indicator," *New York Times*, October 7, 2008. "Profile of GMAT Candidates, 2005–06 to 2009–10," **gmac.com**. *2015 Application Trend Survey*, **gmac.com**.

The Sharing Economy and the Short Run: James R. Hagerty, "Startup Matches Heavy Equipment Owners and Renters," *Wall Street Journal*, May 28, 2015.

A Beer Manufacturer's Short-Run Cost Curves: Flath (2011).

3D Printing: "Print Me a Stradivarius," **http://www.economist.com/node/18114327**, February 10, 2011. Lucas Mearian, "3D Printing Moves from Prototypes to Production," *Computerworld*, April 30, 2014. Lucas Mearian, "Is Apple Planning a 3D Printer?" *Computerworld*, May 21, 2015. **http://phys.org/news/2016-06-airbus-3d-printed-mini-aircraft.html** (viewed June 14, 2016).

A Beer Manufacturer's Long-Run Cost Curves: Flath (2011).

Should You Buy an Inkjet or a Laser Printer?: **http://www.dell.com**. **http://www.amazon.com** (viewed June 13, 2016).

Medical Economies of Scope: Cohen and Morrison Paul (2011). Gonçalves and Barros (2013). Carey et al. (2015).

Chapter 8

Challenge

The Rising Cost of Keeping On Truckin': "About Interstate Trucking Authority," **http://www.ehow.com/about_4739672_interstate-trucking-authority.html** (viewed on June 2012). **http://www.fmcsa.dot.gov/rules-regulations/rules-regulations.htm** (viewed June 2012). **http://www.dotauthority.com/ucr.htm?gclid=CKma5tebnqQCFSFZiAodN2bunA** (viewed June 2012). James L Gattuso, "Truckers Don't Need Mandated Recorders," *Freemont Tribune*, May 29, 2012. James Jaillet, "Logging Device Mandate Could Come in 2016, Outlines Hardware Spec's, Harassment Provisions," *Overdrive*, March 13, 2014.

Applications

Fracking and Shutdowns: Agis Salpukas, "Low Prices Have Sapped Little Oil Producers," *New York Times*, April 3, 1999:B1, B4. Robert Collier, "Oil's Dirty Future," *San Francisco Chronicle*, May 22, 2005:A1, A14, A15. Jon Birger, "Oil Shale May Finally Have Its Moment," *Fortune*, November 1, 2007. Judith Kohler, "Energy Firms Cautious on Oil Shale," *OCRegister*, November 3, 2007. Steve Austin, "Falling Oil Prices Slows US Fracking," **http:www.Oil-price.net**, December 8, 2014. Shaw Tully, "The Shale Oil Revolution Is in Danger," **http:www.fortune.com**, January 9, 2015. **http://www.macrotrends.net/1369/crude-oil-price-history-chart** (viewed June 16, 2016).

The Size of Ethanol Processing Plants: "Statistics," Renewable Fuels Association, **http://www.ethanolrfa.org/resources/industry/statistics/** (viewed June 20, 2016). **http://www.eia.gov/petroleum/ethanolcapacity/** (viewed June 20, 2016).

http://www.ethanolproducer.com/plants/listplants/US/Existing/Sugar-Starch/ (viewed June 20, 2016). **http://www.tradingeconomics.com/commodity/ethanol** (viewed June 20, 2016).

Entry and Exit of Solar Power Firms: Borenstein (2015). Lucas Davis and Alex Chun (2015). I am grateful to Davis and Chun for providing the data for the figure.

Upward-Sloping Long-Run Supply Curve for Cotton: International Cotton Advisory Committee, *Survey of the Cost of Production of Raw Cotton*, September 1992:5. *Cotton: World Statistics*, April 1993:4–5. The figure shows the supply of the major producing countries for which we have cost information. The only large producers for whom cost data are missing are India and China.

Reformulated Gasoline Supply Curves: Borenstein et al. (2004). David R. Baker, "Rules Fuel Patchwork Quilt of Gas Blends Nationwide," *San Francisco Chronicle*, June 19, 2005:B1, B3. Brown et al. (2008). Ronald D. White, "Refinery Work Boosts Gasoline Prices in California," *Los Angeles Times*, September 9, 2009. David R. Baker, "Why Californians Are Paying Even More for Gas," *San Francisco Chronicle*, September 23, 2009, p. A1. Auffhammer and Kellogg (2011). **https://www.eia.gov/dnav/pet/pet_pri_gnd_dcus_nus_w.htm** (viewed June 16, 2016).

Chapter 9

Challenge

Liquor Licenses: Jim Saksa, "Rum Deal, Counting Up All the Ways America's Booze Laws Are Terrible," **http://www.slate.com/articles/business/moneybox/2014/06/america_s_booze_laws_worse_than_you_thought.html**, June 12, 2014. Various state licensing websites, 2016.

Applications

What's a Name Worth? **http://www.forbes.com/** (viewed June 18, 2016). **https://www.cia.gov** (viewed June 18, 2016).

Willingness to Pay on eBay: Bapna et al. (2008). **http://www.eBay.com**. Hasker et al. (2014).

Goods with a Large Consumer Surplus Loss from Price Increases: Revenues in 2012 from the U.S. National Income and Product Accounts, **http://www.bea.gov** and **http://www.econstats.com** (viewed June 28, 2013). Elasticities are based on Blanciforti (1982). Appendix 9A shows how ΔCS was calculated.

Deadweight Loss of Christmas Presents: Waldfogel (1993, 2005, 2009). Bauer and Schmidt (2008). Eve Mitchell, "Cashing In," *Oakland Tribune*, June 12, 2011, C1. Barbara Farfan, "2011 U.S. Christmas Holiday Retail Data, Statistics, Results, Numbers Roundup Complete U.S. Retail Industry Christmas Holiday Shopping Year-Over-Year Results," December 29, 2011. Miguel Helft, "Meet the Anti-Groupon," **http://www.CNNMoney.com**, April 30, 2012. **http://www.tech.fortune.cnn.com/2012/04/30/wrapp/**. Matt Brownell, "Gift Cards," *Forbes*, December 12, 2012. Jason Russell, "Christmas Gift Fraud Costs Billions," *Washington Examiner*, December 26, 2014. **https://www.cardcash.com/gift-card-statistics/** (viewed June 17, 2016).

Welfare Effects of Allowing Fracking: Hausman and Kellogg (2015).

The Deadweight Loss from Gas Taxes: Blundell et al. (2012).

How Big Are Farm Subsidies and Who Gets Them?: *Agricultural Policies Monitoring and Evaluations 2016*, Organization for Economic Cooperation and Development, 2016, **http://www.oecd.org/tad/agricultural-policies/monitoring-and-evaluation.htm** (viewed on June 27, 2016). **http://www.farm.ewg.org** (viewed June 27, 2016).

The Social Cost of a Natural Gas Price Ceiling: MacAvoy (2000). Davis and Kilian (2011).

Russian Food Ban: "Russia to Ban All Agriculture Imports from US: Report," **http://www.nypost.com**, August 6, 2014. "Inflation Soars Above 8% as Food Bans and Ruble Bite," *The Moscow Times*, November 5, 2014. **http://www.nbcnews.com/storyline/ukraine-crisis/russias-sanctions-war-against-west-explained-n412416**, August 22, 2015. **http://www.themoscowtimes.com/arts_n_ideas/business/article/finnish-dairy-company-valio-to-close-plant-due-to-russian-food-embargo/571162.html**, June 1, 2016. **https://rbth.com/business/2016/06/01/russia-eases-its-food-embargo_599381**, June 1, 2016. **http://sputniknews.com/russia/20160617/1041529091/russia-food-embargo-membership.html**, June 17, 2016.

Chapter 10

Challenge

Anti-Price Gouging Laws: Davis (2008). Carden (2008). Michael Giberson, "Predictable Consequences of Anti-Price Gouging Laws," **http://www.knowledgeproblem.com/2009/06/01/predictable-consequences-of-anti-price-gouging-laws**, June 1, 2009. "Bill Allows Gas Price Increase in Emergency," **http://www.ajc.com**, March 22, 2010; "Price Gouging Law Goes into Effect Following State of Emergency Declaration," **http://www.neworleans.com**, April 29, 2010. Keith Goble, "Rhode Island Imposes Price-Gouging Protections," **http://lwww.andlinemag.com**, June 25, 2012. **http://www.ag.ny.gov/press-release/**

ag-schneiderman-cracks-down-gas-stations-engaged-hurricane-sandy-price-gouging, May 2, 2013. "Virginia's Anti-Price Gouging Law Now In Effect," *Fairfax News*, February 12, 2014. "State Anti-gouging Law in Effect Amid Winter Storm," **http://www.wvgazette.com**, March 5, 2015. Jack Brammer, "Beshear Issues Storm-Related Orders to Protect Against Price Gouging," **http://www.Kentucky.com**, February 17, 2015. **http://www.kxii.com/content/news/Emergency-declaration-keys-Oklahoma-anti-price-gouging-law-379050721.html**, May 11, 2016.

Applications

Urban Flight: Econsult (2003). Crawford et al. (2004). **http://www.phila.gov/Revenue/businesses/taxes/Pages/WageTax.aspx** (viewed June 20, 2016).

The 1% Grow Wealthier: Sylvia Nasar, "The Rich Get Richer, but Never the Same Way Twice," *New York Times*, August 16, 1992:3. Davies et al. (2007). Bucks et al. (2009). Kennickell (2009). **http://www.joshuakennon.com/how-much-money-does-it-take-to-be-in-the-top-1-of-wealth-and-net-worth-in-the-united-states**, September 16, 2011. Sylvia Allegreto, "The State of Working America's Wealth," EPI Briefing Paper, March 23, 2011. Kennickell (2009, 2011). Lawrence Mishel and Natalie Sabadish, "CEO Pay and the Top 1%," EPI Issue Brief, May 2, 2012. Linda Levine, "An Analysis of the Distribution of Wealth Across Households, 1989–2010," Congressional Research Service, 2012. Alvaredo et al. (2013). **http://www.aflcio.org/Corporate-Watch/CEO-Pay-and-the-99/Trends-in-CEO-Pay** (viewed on July 10, 2015). **http://www.eml.berkeley.edu/~saez/TabFig2014prel.xls** (viewed on July 10, 2015). Peterson-Withorn, Chase, "Forbes Billionaires: Full List of the 500 Richest People in the World 2015," **http://www.forbes.com**, March 3, 2015. **http://www.huffingtonpost.com/waltons-net-worth_n_1680642.html** (viewed on July 10, 2015). *Wealth: Having It All and Wanting More*, Oxfam Issue Briefing, January 2015. Credit Suisse, *Global Wealth Databook 2015*. Saez (2016).

How You Vote Matters: "The Mathematics of Voting: Democratic Symmetry," *The Economist*, March 4, 2000:83. Nancy Vogel, "'Instant Runoff' Voting Touted," *Los Angeles Times*, December 25, 2006, **http://articles.latimes.com/2006/dec/25/local/me-voting25**. **http://www.sfelections.org/demo/**, December 2007. Rob Richie, "Of Alternative Voting Methods, Instant Runoff Is Best," **http://www.mprnews.org/story/2009/09/29/richie**. "Our View: Charter Commission Should Split Decisions," *Portland Press Herald*, June 30, 2010. Rosemary Westwood, "Explainer: What Comes Next in the Liberal Vote," *Maclean's*, April 5, 2013. **http://time.com/4352797/ranked-choice-voting-maine-donald-trump/**, June 2, 2016.

Chapter 11

Challenge

Brand-Name and Generic Drugs: **http://www.fiercepharma.com/special-reports/top-10-patent-expirations-2015**, December 17, 2014. Trefis Team, "Why Are Generic Drug Prices Shooting Up?", *Forbes*, February 27, 2015. "Pricey Hep C Successor Overtakes Solvadi," *Health News Florida*, July 13, 2015. **http://www.drugs.com/article/patent-expirations.html** (July 18, 2015). **http://www.imshealth.com** (July 18, 2015). **http://www.pharmalive.com/loss-of-exclusivity/**, June 8, 2016. **https://www.statnews.com/pharmalot/2016/06/09/pfizer-drug-prices-turing-valeant/**, June 9, 2016. **http://www.consumerreports.org/drugs/cure-for-high-drug-prices/**, June 21, 2016. **http://www.statista.com/statistics/205042/proportion-of-brand-to-generic-prescriptions-dispensed/** (viewed June 25, 2016).

Applications

Apple's iPad: Chloe Albanesius, "iSuppli: iPad Could Produce Profits for Apple," **http://www.pcmag.com**, February 10, 2010. Don Reisinger, "IDC: Apple iPad Secures 87 Percent Market Share," **http://www.cnet.com**, January 18, 2011. Don Reisinger, "Study: iPad Tallies 89 Percent of Table Traffic," **http://www.cnet.com**, January 24, 2011. Jenna Wortham, "So Far Rivals Can't Beat iPad's Price," *New York Times*, March 6, 2011. **http://www.idc.com/getdoc.jsp?containerId=prUS41218816** (viewed June 24, 2016). The marginal cost estimate (slightly rounded) is from iSuppli. I assumed that the company's gross profit margin for 2010 of about 40% (**http://www.forbes.com**) held for the iPad and used that to calculate the fixed cost. I derived the linear demand curve by assuming Apple maximizes short-run profit using the information on price, marginal cost, and quantity.

Cable Cars and Profit Maximization: Rachel Gordon, “A Fare Too Steep?” *San Francisco Chronicle*, September 12, 2006:B-1. Rachel Gordon, “Cable Car Fares Rising a Buck to $6 on July 1,” *San Francisco Chronicle*, June 7, 2011. “Cable Car Fares,” *SFMTA*, **http://www.sfmta.com/cms/mfares/fareinfo.htm#cable** (viewed on July 18, 2016).

Botox Patent Monopoly: Mike Weiss, “For S.F. Doctor, Drug Botox Becomes a Real Eye-Opener,” *San Francisco Chronicle*, April 14, 2002:A1, A19. Reed Abelson, “F.D.A. Approves Allergan Drug for Fighting Wrinkles,” *New York Times*, April 16, 2002. Joseph Walker, “Botox Itself Aims Not to Age,” *Wall Street Journal*, May 18, 2014. Ryan Sachetta and Cynthia Koons, “Worried About Wrinkles, Guys? Allergan Bets You’ll Want ‘Brotox”’ **http://www.Bloomberg.com**, June 17, 2015. **http://www.evaluategroup.com/View/15548—1002-modData/product/botox** (viewed June 24, 2016).

Natural Gas Regulation: Davis and Muehlegger (2010): We used their most conservative estimate (the one that produced the smallest deadweight loss). We approximated their demand curve with a linear one that has the same price elasticity of demand at point *b*. This figure represents the aggregation of state-level monopolies to the national level.

Movie Studios Attacked by 3D Printers!: Erich Schwartzel, “Hollywood’s Other Piracy Problem: 3-D Printers,” *Wall Street Journal*, July 20, 2015.

eBay’s Critical Mass: Ina Steiner, “Yahoo Closes Australian Auction Site,” **http://www.auctionbytes.com**, August 7, 2003. Brown and Morgan (2006, 2009). John Barrett, “MySpace Is a Natural Monopoly,” **http://www.ecommercetimes.com**, January 17, 2007. Carol Xiaojuan Ou and Robert M. Davison, “Why eBay Lost to TaoBao in China: The Global Advantage,” *Communications of the ACM*, 52(1), January 2009:145–148.

Chapter 12

Challenge

Sale Prices: Perloff and Wu (2007). Karen Datko, “My Ketchup Taste Test: It’s an Upset!” **http://www.msn.com**, January 14, 2011. Rebecca Smithers, “Heinz Left Playing Tomato Catch-Up after Ketchup Tasting Trouncing,” *The Guardian*, May 25, 2011. Andrew Adam Newman, “Ketchup Moves Upmarket, with a Balsamic Tinge,” *New York Times*, October 25, 2011. Whitney Filloon, “Heinz and French’s Are Embroiled in a Ketchup and Mustard War,” **http://www.eater.com**, April 23, 2015. **http://business.financialpost.com/news/retail-marketing/condiment-wars-how-canadian-nationalism-helped-a-u-k-owned-corporate-giants-quest-to-be-ketchup-king**, March 17, 2016. **http://www.wikinvest.com/stock/H.J._Heinz_Company_(HNZ)** (viewed July 1, 2016). **http://www.heinz.com/our-company/press-room/trivia.aspx** (viewed July 1, 2016).

Applications

Disneyland Pricing: “Couple Tries for Year of Daily Disneyland Visits,” *San Francisco Chronicle*, July 3, 2012. **http://www.disneyland.com** (viewed on August 2, 2015). Joseph Pimentel, “Disney Announces Discounted Tickets for SoCal Residents,” **http://www.ocregister.com/articles/park-698721-disneyland-ticket.html**, January 6, 2016.

Preventing Resale of Designer Bags: Eric Wilson, “Retailers Limit Purchases of Designer Handbags,” *New York Times*, January 10, 2008. **http://www.prada.com**, **http://www.gucci.com**, **http://www.saksfifthavenue.com** (viewed on August 2, 2015).

Botox and Price Discrimination: See Chapter 11, “Botox Patent Monopoly.” Mike Weiss, “For S.F. Doctor, Drug Botox Becomes a Real Eye-Opener,” *San Francisco Chronicle*, April 14, 2002:A1, A19. Reed Abelson, “F.D.A. Approves Allergan Drug for Fighting Wrinkles,” *New York Times*, April 16:2002. Vita Reed, “Botox, Medical Cosmetics Boost Allergan’s Profits,” **http://www.ocbj.com/accounts/login/?next=/news/2010/may/09/botox-medical-cosmetics-boost-allergans-profits/** (July 2010). “Nice News for Allergan—Analyst Blog,” **http://www.nasdaq.com/article/nice-news-for-allergan-analyst-blog-cm152408**, June 29, 2012. Deborah Knapp, “Botox Approved as Migraine Medicine, Spurs Headache Club,” **http://www.kens5.com/**, July 8, 2013.

Google Uses Bidding for Ads to Price Discriminate: Goldfarb and Tucker (2011).

Warner Brothers Sets Prices for a Harry Potter **DVD:** Sources of information and data for this Application and the following Solved Problem include Amazon websites for each country, **http://www.warnerbros.com**, **http://www.ukfilmcouncil.org.uk**, **http://www.the-numbers.com/dvd/charts/annual/2011.php**, and **http://www.bbc.co.uk/**

newsbeat/16444062, January 6, 2012. We assume that the demand curves in each country are linear.

Reselling Textbooks: Adam Liptak, "Justices Weigh Case on Imported Textbooks," *New York Times*, October 29, 2012. Robert Barnes, "In Copyright Case, Supreme Court Rules That Goods Made Overseas Can Be Resold Here," *Washington Post*, March 19, 2013. Daniel Fisher, "Supreme Court Upholds Right to Sell Foreign-Published Books," *Forbes*, March 19, 2013. Ry Rivard, "Libraries Can Lend Foreign Books," *Inside Higher Ed*, March 20, 2013. **http://www.amazon.com** and **http://www.amazon.co.uk** (viewed July 2, 2016).

Buying Discounts: Borenstein and Rose (1994). "Up Front," *Consumer Reports*, September 2009:7. **http://www.statista.com/topics/1156/coupon-market-trends-in-the-united-states**, 2015. **http://www.businesswire.com/news/home/20150212006009/en/Consumers-Save-3.6-Billion-Redeeming-Coupons-2014**, February 2015. **http://www.statista.com/statistics/502230/value-fsi-coupons-usa/** (viewed July 1, 2016). **http://www.statista.com/statistics/502266/number-fsi-coupons-usa/** (viewed July 1, 2016).

http://blog.accessdevelopment.com/ultimate-collection-coupon-statistics (viewed July 1, 2016).

iTunes for a Song: Shiller and Waldfogel (2011).

Ties That Bind: **http://www.hp.com** (viewed July 15, 2013). **http://www.consumerchoice.info/warranty.htm** (viewed July 15, 2013). **http://www.mlmlaw.com/library/guides/ftc/warranties/undermag.htm** (viewed July 15, 2013).

Super Bowl Commercials: Ho et al. (2009). Kim (2011). **http://www.fortune.com/2016/02/06/super-bowl-ads/**, February 6, 2016.

Chapter 13

Challenge

Government Aircraft Subsidies: Irwin and Pavcnik (2004). John Heilpin, "WTO: Boeing Got $5B In Illegal Subsidies," March 12, 2012. **http://www.manufacturing.net/news/2012/03/wto-boeing-got-5b-in-illegal-subsidies** (viewed on June 10, 2012). **http://www.opensecrets.org/lobby/clientlbs.php?id=D000000100&year=2015**. **http://www.wsj.com/articles/eu-files-complaint-with-wto-about-boeing-1419008774**. **http:www.247wallst.com/aerospace-defense/2015/02/24/wto-to-examine-boeing-777x-subsidies**. **http://www.reuters.com/article/2015/06/16/us-airshow-france-boeing-airbus-idUSKBN0OW0VM20150616**. **https://www.opensecrets.org/lobby/clientsum.php?id=D000000100** (viewed June 27, 2016).

Applications

The Apple-Google-Intel-Adobe-Intuit-Lucasfilms-Pixar Wage Cartel: Michael Liedtke, "Apple, Google, Other Tech Firms to Pay $415M in Wage Case," **http://www.seattletimes.com**, January 15, 2015. Mark Ames, "The Techtopus: How Silicon Valley's Most Celebrated CEOs Conspired to Drive Down 100,000 Tech Engineers' Wages," **http://www.pando.com**, January 23, 2014. Mark Ames, "Revealed: Apple and Google's Wage-Fixing Cartel Involved Dozens More Companies, Over One Million Employees," **http://www.pando.com**, March 22, 2014.

Cheating on the Maple Syrup Cartel: Ian, Austen "The Maple Syrup Mavericks," *New York Times*, August 23, 2015. **http:/www.siropderable.ca/home.aspx** (viewed on August 24, 2015).

Mergers to Monopolize: Collard-Wexler (2014).

Mobile Number Portability: Cho et al. (2015). **http://mcclist.com/number-portability.asp** (viewed July 5, 2016).

Bottled Water: David Lazarus, "How Water Bottlers Tap Into All Sorts of Sources," *San Francisco Chronicle*, January 19, 2007. Vinnee Tong, "What's in That Bottle?" **http://www.suntimes.com**, July 28, 2007. Kevin Cowherd, "Bottled-Water Labeling: A Source of Irritation," **http://www.baltimoresun.com**, August 1, 2007. **http://www.gallup.com/poll/153875/worry-water-air-pollution-historical-lows.aspx** (viewed on July 16, 2012); **http://www.forbes.com/sites/greatspeculations/2015/06/26/bottled-water-is-a-potential-growth-category-that-cant-be-ignored**. Venessa Wong, "Coca-Cola Wants to Buy the World a Milk," **http://www.businessweek.com**, December 1, 2014. Dan Nosowitz, "Coca-Cola to Sell Sexy Lactose-Free Milk Product of Some Kind," **http://www.modernfarmer.com**, December 2, 2014. Khushbu Shah, "Coca-Cola's New 'Super Milk' Fairlife Is Super Weird," *Eater*, February 16, 2015. **http://www.bevindustry.com/articles/89123-us-bottled-water-market-grows-64-percent-in-2015**, February 26, 2016.

Monopolistically Competitive Food Truck Market: Andrew S. Ross, "San Francisco Food Truck Empire Expanding," *San Francisco Chronicle*, February 18, 2011. **www.mobilefoodnews.com** (viewed June 17, 2012). **http://offthegrid.com/about/** (viewed July 6, 2016).

Zoning Laws as a Barrier to Entry by Hotel Chains: Suzuki (2013).

Chapter 14

Challenge

Intel and AMD's Advertising Strategies: Salgado (2008). **https://www.cpubenchmark.net/market_share.html** (viewed July 9, 2016). **http://jonpeddie.com/publications/market_watch** (viewed July 9, 2016).

Applications

Strategic Advertising: "50 Years Ago . . . ," *Consumer Reports*, January 1986. Roberts and Samuelson (1988). Gasmi et al. (1992). Elliott, Stuart, "Advertising," *New York Times*, April 28, 1994:C7. Slade (1995). Salgado (2008). Richards and Padilla (2009). David and Markowitz (2011). Chandra and Weinberg (2015). **http://adage.com/article/advertising/top-200-u-s-advertisers-spend-smarter/304625/**, July 8, 2016. **http://www.statista.com/topics/1625/procter-and-gamble** (viewed on July 8, 2016). Anderson et al. (2016).

Tough Love: Marta Choroszewicz and Pascal Wolff, "Population and Social Conditions," *Eurostat*, 50/2010. Parker, Kim, "The Boomerang Generation," Pew Research Center, March 15, 2012. **http://www.britishinsurance.com/insurance-news/2nd-july—financial-support.aspx**, July 2, 2012. Hilary Osborne, "Record Levels of Young Adults Living at Home, Says ONS," *The Guardian*, January 21, 2014. **http://www.pewsocialtrends.org/2014/07/17/in-post-recession-era-young-adults-drive-continuing-rise-in-multi-generational-living.** Shiv Malik, "The Dependent Generation: Half Young European Adults Live with Their Parents," March 24, 2014. **http://www.bloomberg.com/news/articles/2015-03-05/parents-risk-retirement-to-support-millennial-kids. http://www.pewsocialtrends.org/2015/05/21/5-helping-adult-children/**, May 21, 2015. **http://www.pewsocialtrends.org/2015/07/29/more-millennials-living-with-family-despite-improved-job-market/**, July 29, 2015. **http://ec.europa.eu/eurostat/web/youth/statistics-illustrated** (viewed on July 8, 2016). **http://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/families/bulletins/familiesandhouseholds/2015-11-05** (viewed on July 8, 2016).

Keeping Out Casinos: Cookson (2015).

Bidders' Curse: Lee and Malmendier (2011). Garratt et al. (2012). Feng et al. (2016).

GM's Ultimatum: Robert Schoenberger, "GM Sends Ultimatums to All Its 6000 US dealers," *Cleveland Plain Dealer*, June 2, 2009. "GM Dealers Sue to Keep Doors Open," *Toronto Star*, November 27, 2009. Janet Kurnovich, "GM Canada Sued for $750 Million by Former Dealers," **http://www.insurance-car.co**, May 17, 2011. Greg Keenan, "Judge Dismisses Class Action by GM Canada Dealers, Upholds Claim Against Law Firm," *The Globe and Mail*, July 9, 2015.

Chapter 15

Challenge

Athletes' Salaries and Ticket Prices: Bill Shaikin, "Experts Say Demand, Not Higher Salaries, Drives Up Baseball Ticket Prices," *Los Angeles Times*, April 1, 1999. Anthony Crupi, "RSNs Posted Across the Board Revenue Gains in '09," *MediaWeek*, June 16, 2010. Rick Snider, "Steinbrenner: Trailblazer to the Lavish, Diva Owner," *Washington Examiner*, **http://www.washingtonexaminer.com**, July 14, 2010. "Joe Strauss Live: Cardinals Chat Q-&-A," **http://www.stltoday.com/sports/baseball/professional/article_e8736348-9509-11df-bf07-00127992bc8b.html**, July 21, 2010. **http://www.usatoday.com/story/sports/mlb/2016/04/03/2016-mlb-salaries-payrolls/82592542/. https://www.baseballprospectus.com/compensation/cots/league-info/highest-paid-players/** (viewed July 13, 2016).

Applications

Unions and Profits: Lee and Mas (2012). Stanfield and Tumarkin (2016).

Monopsony and the Gender Wage Gap: Madden (1973). Webber (2016).

Chapter 16

Challenge

Does Going to College Pay?: Richard Pérez-Peña, "U.S. Bachelor Degree Rate Passes Milestone," *New York Times*, February 23, 2012. Andrew Kelly, "The Neglected Majority: What Americans Without a College Degree Think about Higher Education," *Forbes*, April 28, 2015. **http://www.usnews.com/news/articles/2015/02/10/college-grads-question-the-return-on-investment-of-todays-degrees. http://www.luminafoundation.org/news-and-events/gallup-2015** (viewed on July 14, 2016). **http://nces.ed.gov/programs/coe/indicator_caa.asp** (viewed July 14, 2016). **http;//www.nces.ed.gov/programs/digest/d15/tables/dt15_330.10.asp** (viewed on July 14, 2016). **http://www.gallup.com/poll/185819/recent-grads-less-likely-agree-college-worth-cost.aspx** (viewed July 14, 2016). **http://www.bls.gov/news.release/hsgec.nr0.htm** (viewed July 14, 2016). The statistical analysis is based on annual earnings (wages or self-employment) data from the 2009 U.S. *Current Population Survey* (Miriam King, Steven Ruggles, J. Trent Alexander, Sarah Flood, Katie Genadek, Matthew B. Schroeder, Brandon Trampe, and Rebecca Vick. *Integrated Public Use Microdata Series, Current Population Survey: Version 3.0*, Minneapolis: University of Minnesota, 2010).

Applications

Power of Compounding: Helen Huntley, Kim Norris, and Robert Trigaux, "An Early Lesson in Investing for the Long Term," *St. Petersburg Times*, March 14, 1994. **http://www.wsj.com/articles/new-york-city-property-values-surge-1452907433** (viewed July 14, 2016). Author's calculations.

Saving for Retirement: Author's calculations.

Winning the Lottery: Geetika Rudra, "Powerball Record Winner Is 84-Year-Old Woman Gloria Mackenzie," **http://www.abcnews.go.com**, June 5, 2013.

Falling Discount Rates and Self-Control: Shapiro (2004). Kan (2007). Gruber and Mullainathan (2005). Jeff Zeleny, "Occasional Smoker, 47, Signs Tobacco Bill," *New York Times*, June 23, 2009. **http://www.gallup.com/poll/1717/Tobacco-Smoking.aspx** (viewed on July 22, 2012). **http://www.aspire2025.org.nz/2012/05/22/article-support-for-a-tobacco-endgame**, May 22, 2012. **http://www.gallup.com/poll/163763/smokers-quit-tried-multiple-times.aspx**, July 31, 2013. **http://www.gallup.com/poll/173990/smokers-say-higher-cigarette-taxes-unjust.aspx**, July 18, 2014. **http://www.globaltimes.cn/content/931484.shtml** (viewed on September 3, 2015).

Redwood Trees: Berck and Bentley (1997). Peter Berck, personal communications.

Chapter 17

Challenge

BP and Limited Liability: Mark, Long, and Angel Gonzalez, "Transocean Seeks Limit on Liability," *Wall Street Journal*, May 13, 2010. David Leonhardt, "Spillonomics: Underestimating Risk," *New York Times*, May 21, 2010. Jef Feeley and Allen Johnson Jr., "BP Wins Final Approval of Guilty Plea Over Gulf Oil Spill," Bloomberg News, January 29, 2013, **http://www.bloomberg.com/news/2013-01-29/bp-wins-final-approval-of-guilty-plea-over-gulf-oil-spill.html.** Daniel Gilbert and Sarah Kent "BP Agrees to Pay \$18.7 Billion to Settle Deepwater Horizon Oil Spill Claims," *Wall Street Journal*, July 2, 2015.

Applications

Stocks' Risk Premium: "The Cost of Looking," *Economist*, 328(7828), September 11, 1993:74. Leslie Eaton, "Assessing a Fund's Risk Is Part Math, Part Art," *New York Times*, April 2, 1995. Jagannathan et al. (2000). **http://www.standardandpoors.com** (viewed on July 17, 2016). **http://pages.stern.nyu.edu/~adamodar/New_Home_Page/datafile/histretSP.html** (viewed on July 17, 2016).

Gambling: Friedman and Savage (1948). Brunk (1981). Meghan Cox Gurdon, "British Accuse Their Lottery of Robbing the Poor to Give to the Rich," *San Francisco Chronicle,* November 25, 1995. Mark Maremont and Alexandra Berzon, "How Often Do Gamblers Really Win?" *Wall Street Journal*, October 11, 2013. **http://www.gamesindustry.biz/articles/2015-04-22-gaming-will-hit-usd91-5-billion-this-year-newzoo; www.cnbc.com/2015/05/28/gambling-revenue-grew-rapidly-in-2014-but-difficulties-lie-ahead.html**.

Failure to Diversify: Paul J. Lim, "Don't Paint Nest Eggs in Company Colors," *New York Times*, March 30, 2008. Duan et al. (2015). Ron Lieber, "A Scary Movie: Filling Your 401(k) with Company Stock," *New York Times*, March 21, 2015.

Flight Insurance: **http://www.buy.travelguard.com** (viewed on September 5, 2015). **http://www.ntsb.gov/investigations/data/Pages/Accident-data-review.aspx** (viewed on July 17, 2016). **http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/subject_areas/airline_information/index.html** (viewed on July 17, 2016).

Limited Insurance for Natural Disasters: Joseph B. Treaster, "Insurer Plans to Curb Sales Along Coasts," *New York Times*, October 10, 1996. Joseph B. Treaster, "Why Insurers Shrink from Earthquake Risk," *New York Times*, November 21, 1999. Paul Vitello, "Home Insurers Canceling in East," *New York Times*, October 16, 2007. Pielke et al. (2008). "Tropical Storm Bonnie Dissipates; Still No Plan from Congress on Flood Insurance," **http://www.insurancenewsnet.com/article.aspx?id=210913&type=newswires**, July 28, 2010. **http://www.propertycasualty360.com/2015/07/27/first-half-of-2015-recorded-below-average-global-n**. **http://www.swissre.com/media/news_releases/Global_insured_losses_from_disasters_reach_USD_37_billion_in_2015.html**, March 30, 2016. **http://ir.aon.com/about-aon/investor-relations/investor-news/news-release-details/2016/Global-economy-suffers-worst-April-natural-disaster-losses-for-five-years-according-to-Aon-catastrophe-report/default.aspx**, May 9, 2016.

Biased Estimates: Benjamin et al. (2001). Arthur Hu, "Death Spectrum," **http://www.arthurhu.com/index/health/death.htm#deathrank** (viewed on July 16, 2016). **http://www.flmnh.ufl.edu/fish/sharks/statistics/statsus.htm** (viewed on July 16, 2016). **http://www.cdc.gov/nchs/fastats/homicide.htm** (viewed on July 16, 2016). **http://www.medicinenet.com/are_poinsettia_plants_poisonous_fact_or_fiction/views.htm** (viewed on July 16, 2016). **http://www.cnn.com/2016/02/09/world/shark-attacks/** (viewed on July 16, 2016).

Loss Aversion Contracts: Fryer et al. (2012). Hossain and List (2012). Imas et al. (forthcoming).

Chapter 18

Challenge

Trade and Pollution: Tideman and Tullock (1976). Lelieveld et al. (2015). **http://www.wto.org** (viewed July 24, 2016). **http://www.trade.gov** (viewed July 24, 2016). **http://www.healthdata.org/news-release/poor-air-quality-kills-55-million-worldwide-annually**, February 12, 2016.

Applications

Negative Externalities from Spam: Caliendo et al. (2012). Rao and Reiley (2012). **http://www.bloomberg.com/news/articles/2016-01-19/e-mail-spam-goes-artisanal**, January 19, 2016. **http://www.kaspersky.com/about/news/spam/2016/Kaspersky-Lab-Reports-Significant-Increase-in-Malicious-Spam-Emails-in-Q1-2016** (viewed July 20, 2016).

Global Warming: **http://www.ipcc.ch/report/ar5/syr/**, 2014. Pal and Eltahir (2015). **http://www.pewresearch.org/fact-tank/2015/06/16/ideological-divide-over-global-warming-as-wide-as-ever/**. Mark Carney, "Breaking the Tragedy of the Horizon," Speech, September 29, 2015, **http://www.bankofengland.co.uk/publications/Documents/speeches/2015/speech844.pdf**. Stephane Hallegatte et al., *Shock Waves*, World Bank, 2016, **https://openknowledge.worldbank.org/bitstream/handle/10986/22787/9781464806735.pdf**. **http://climate.nasa.gov/scientific-consensus/** (viewed July 23, 2016). **https://www.skepticalscience.com/global-warming-scientific-consensus-intermediate.htm** (viewed July 23, 2016).

Why Tax Drivers: Levitt and Porter (2001). Grabowski and Morrisey (2006), Edlin and Karaca-Mandic (2006), Parry et al. (2007), Anderson (2008). Hill et al. (2009). "Dutch Road Tax Inflames Debate in Neighboring Germany," *Deutsche Welle*, November 16, 2009, **http://www.dw.com/en/dutch-road-tax-inflames-debate-in-neighboring-germany/a-4896686**; **http://www.cesifo-group.de/portal/pls/portal/docs/1/1192818.PDF** (viewed on August 1, 2012). Anderson and Auffhammer (2014). Sheehan-Connor (2015). Lucas Davis, "Raise the Gas Tax," **http://www.energyathaas.wordpress.com/2015/01/05/raise-the-gas-tax**.

Protecting Babies: Currie and Walker (2011). Anderson (2015). Knittel et al. (2016).

Buying a Town: Katharine Q. Seelye, "Utility Buys Town It Choked, Lock, Stock and Blue Plume," *New York Times*, May 13, 2002. "Cheshire Ohio," Abandoned, **http://www.abandonedonline.net/neighborhoods/cheshire-ohio** (viewed August 7, 2013).

Acid Rain Cap-and-Trade Program: Dallas Burtraw, "Trading Emissions to Clean the Air: Exchanges Few but Savings Many," *Resources*, 122, Winter 1996:3–6. Peter Passell, "For Utilities, New Clean-Air Plan," *New York Times*, November 18, 1994:C1, C6. Peter Passell, "Economic Scene," *New York Times*, January 4, 1996:C2. Boyce Rensberger, "Clean Air Sale," *Washington Post*, August 8, 1999:W7. Schmalensee et al. (1998). EPA (2011). **http://www.epa.gov/airmarkets/so2-allowance-auctions** (viewed July 20, 2016).

Road Congestion: **http://inrix.com/scorecard/key-findings-us/** (viewed July 21, 2016).

Software Piracy: Business Software Alliance, *The Compliance Gap*, June 2014. **http://www.bsa.org/~/media/Files/StudiesDownload/BSA_GSS_US.pdf**, 2016.

Free Riding on Measles Vaccinations: Orenstein et al. (2004). Jo Craven McGinty, "How Anti-Vaccination Trends Vex Herd Immunity: Measles Outbreak Underscores Vulnerabilities Posed by Subpar Inoculation Rates," *Wall Street Journal*, February 6, 2015. **http://www.cdph.ca.gov/HealthInfo/discond/Pages/Measles.aspx** (viewed July 23, 2016).

What's Their Beef?: Gina Holland, "Top Court Considers Challenge to Beef Ads," *San Francisco Chronicle,* December 9, 2004:C1, C4. **http://www.extension.iastate.edu/agdm/articles/mceowen/McEowJuly05.htm. http://www.beefboard.org/promotion/checpromotion.asp** (December 2009, August 2012, September 2015, July 2016).

Chapter 19

Challenge

Dying to Work: Viscusi (1979). Dick Meister, "Safety First!" **http://www.truth-out.org/safety-first61799**, July 28, 2010. Kris Alingod, "Massey to Resume Operations in West Virginia Mine Despite Deaths," **http://www.allheadlinenews.com**, July 29, 2010. Tracie Mauriello and Len Boselovic, "Historic Fine Issued for Mine Disaster," *Pittsburgh Post-Gazette*, December 7, 2011. Kris Maher, "Agency Blames Massey for Fatal Mine Disaster," *Wall Street Journal*, June 29, 2011. Lateef Mungin and Farid Ahmed, "Report: 8 Arrested after Deadly Bangladesh Building Collapse," **http://www.cnn.com**, April 27, 2013. **http://www.world.time.com/2013/06/10/bangladesh-factory-collapse-uncertain-future-for-rana-plaza-survivors**, June 10, 2013. Ceylan Yeginsu, "Anger and Grief Simmer in Turkey a Year After Soma Mine Disaster," *New York Times*, June 2, 2015. **https://www.theguardian.com/world/2015/sep/12/tianjin-explosion-china-sets-final-death-toll-at-173-ending-search-for-survivors. http://www.nytimes.com/2016/04/07/us/donald-blankenship-sentenced-to-a-year-in-prison-in-mine-safety-case.html. http://www.bls.gov/iif/oshcfoi1.htm** (viewed July 27, 2016). **http://www.osha.gov** (viewed July 27, 2016).

http://www.latimes.com/world/la-fg-bangladesh-factory-fire-20160910-snap-story.html (viewed October 26, 2016).

Applications

Discounts for Data: Tara Siegel Bernard, "Giving Out Private Data for Discount in Insurance," *New York Times*, April 8, 2015. **https://www.investor.fitbit.com/press/press-releases/press-release-details/2016/Fitbit-Introduces-Fitbit-Group-Health-for-Corporate-Wellness-Weight-Management-Programs-Insurers-and-Clinical-Research/default.aspx** (viewed July 27, 2016).

Adverse Selection and Remanufactured Goods: Subramanian and Subramanyam (2012). Neto (2016).

Reducing Consumers' Information: Private Label Manufacturers Association's 2016 *Private Label Yearbook* press release **http://plma.com/share/press/resources/PLMA2016YB_COMB_RPT.pdf. http://www.plmainternational.com/industry-news/private-label-today** (viewed July 27, 2016).

Cheap Talk in eBay's Best Offer Market: Backus et al. (2015).

Chapter 20

Challenge

Clawing Back Bonuses: Matthew Goldstein, "Why Merrill Lynch Got Burned," *Businessweek*, October 25, 2007. Erin McCormick and Carolyn Said, "Investors Own about One-Fifth of Bay Area Homes in Foreclosure," *San Francisco Chronicle*, December 16, 2007. Adams et al. (2009). Elizabeth G. Olson, "Executive Pay Clawbacks: Just a Shareholder Pacifier?" **http://www.management.fortune.cnn.com/2012/08/16/**

executive-pay-clawbacks. “Morgan Stanley Defers Bonuses for High-Earners,” *Chicago Tribune*, January 15, 2013. **http://www.execcomp.org/News/NewsStories/nearly-90-pct-of-fortune-100-companies-now-disclose-clawback-policies-according-to-equilar-survey**, October 25, 2013. Andrew Ackerman, “SEC Proposes Broadened Corporate Clawback Rules,” *Wall Street Journal*, July 1, 2015. **http://www.shearman.com/~/media/Files/NewsInsights/Publications/2016/05/The-Re-Proposed-Rule-on-Incentive-Based-Compensation-at-Financial-Institutions-Overview-and-Observations-CGE-050516.pdf. http://www.realtytrac.com** (viewed on July 29, 2016). Emily Glazer, “Wells Fargo Claws Back Millions from CEO after Scandal,” *Wall Street Journal*, September 27, 2016.

Applications

Honest Cabbie?: Balafoutas et al. (forthcoming).

Health Insurance and Moral Hazard: Kowalski (2015). **http://www.hhs.gov/about/news/2016/03/03/20-million-people-have-gained-health-insurance-coverage-because-affordable-care-act-new-estimates**. Jhamb et al. (2015).

Sing for Your Supper: Personal communications with Gregg Perloff.

Capping Oil and Gas Bankruptcies: Davis (2014). Boomhower (2016).

Layoffs Versus Pay Cuts: Hall and Lilien (1979). Steven Greenhouse, “More Workers Face Pay Cuts, Not Furloughs,” *New York Times*, August 3, 2010. **http://www.bls.gov** (July 29, 2016).

References

ABC News, "High Pressure Water Drilling Touted as an Environmentally Friendly Replacement for Fracking," 2016, July 26, http://www.abc.net.au/news/rural/2016-07-26/new-drilling-technology-touted-as-fracking-replacement/7658540

Adams, William, Liran Einav, and Jonathan Levin, "Liquidity Constraints and Imperfect Information in Subprime Lending," *American Economic Review*, 99(1), March 2009:49–84.

Aigner, Dennis J., and Glen G. Cain, "Statistical Theories of Discrimination in Labor Markets," *Industrial and Labor Relations Review*, 30(2), January 1977:175–187.

Akerlof, George A., "Labor Contacts as Partial Gift Exchanges," *Quarterly Journal of Economics*, 97(4), November 1982:543–569.

Akerlof, George A., "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, 84(3), August 1970:488–500.

Alexander, Donald L., "Major League Baseball," *Journal of Sports Economics*, 2(4), November 2001:341–355.

Altonji, Joseph G., and Ernesto Villanueva, "The Marginal Propensity to Spend on Adult Children," *B.E. Journal of Economic Analysis & Policy*, Advances, 7(1), 2007, Article 14.

Alvaredo, Facundo, Anthony B. Atkinson, Thomas Piketty, and Emmanuel Saez. 2013. "The Top 1 Percent in International and Historical Perspective," *Journal of Economic Perspectives*, 27(3), Summer 2013:3–20.

Anderson, Michael, "Safety For Whom? The Effects of Light Trucks on Traffic Fatalities," *Journal of Health Economics*, 27(4), July 2008:973–989. Anderson, Michael, "As the Wind Blows: The Effects of Long-Term Exposure to Air Pollution on Mortality," NBER Working Paper, 2015.

Anderson, Michael, and Maximilian Auffhammer, "Pounds that Kill: The External Costs of Vehicle Weight," *Review of Economic Studies*, 81(2), April 2014:535–571.

Anderson, Simon P., et al., "Push-Me Pull-You: Comparative Advertising in the OTC Analgesics Industry," working paper, 2016.

Arrow, Kenneth, *Social Choice and Individual Values*. New York: Wiley, 1951.

Atalay, Enghin, "Materials Prices and Productivity," Center for Economic Studies, 2012.

Auffhammer, Maximilian, and Ryan Kellogg, "Clearing the Air? The Effects of Gasoline Content Regulation on Air Quality," *American Economic Review*, 101(6), October 2011:2,687–2,722.

Australian Competition Law, "ACCC v Colgate-Palmolive Pty Ltd & Ors," 2016, https://www.australiancompetitionlaw.org/cases/current/2013-nsd2510-colgate.html

Ayres, Ian, and Joel Waldfogel, "A Market Test for Race Discrimination in Bail Setting," *Stanford Law Review*, 46(5), May 1994:987–1047.

Backus, Matthew, Tom Blake, and Steven Tadelis, "Cheap Talk, Round Numbers, and the Economics of Negotiation," NBER Working Paper, 2015.

Balafoutas, Loukas, Rudolf Kerschbamer, and Matthias Sutter, "Second-Degree Moral Hazard in a Real-World Credence Goods Market," *Economic Journal*, forthcoming.

Bapna, Ravi, Wolfgang Jank, and Galit Shmueli, "Consumer Surplus in Online Auctions," *Information Systems Research*, 19(4), December 2008:400–416.

Barkham, Patrick, "Six Billion Plastic Bags Can't Be Wrong—So What Do We Tax Next?," *The Guardian*, October 3, 2016, https://www.theguardian.com/environment/shortcuts/2016/oct/03/six-billion-plastic-bags-cant-be-wrong-so-what-do-we-tax-next.

Basker, Emek, "Raising the Barcode Scanner: Technology and Productivity in the Retail Sector," *American Economic Journal: Applied Economics*, 4(2), July 2012:1–27.

Battalio, Raymond, John H. Kagel, and Carl Kogut, "Experimental Confirmation of the Existence of a Giffen Good," *American Economic Review*, 81(3), September 1991:961–970.

Bauer, Thomas K., and Christoph M. Schmidt, "WTP vs. WTA: Christmas Presents and the Endowment Effect," Institute for the Study of Labor Working Paper, 2008.

Baugh, Brian, Itzhak Ben-David, and Hoonsuk Park, "Can Taxes Shape an Industry? Evidence from the Implementation of the 'Amazon Tax,' " Ohio State University Dice Center working paper, 2015.

Baumeister, Christiane, and Gert Peersman, "The Role of Time-Varying Price Elasticities in Accounting for Volatility Changes in the Crude Oil Market," *Journal of Applied Econometrics*, 28(7), November/December 2013:1087–1109.

Becker, Gary S., *The Economics of Discrimination*, 2nd ed. Chicago: University of Chicago Press, 1971.

Benjamin, Daniel K., William R. Dougan, and David Buschena, "Individuals' Estimates of the Risk of Death: Part II—New Evidence," *Journal of Risk and Uncertainty,* 22(1), January 2001:35–57.

Berck, Peter, and William R. Bentley, "Hotelling's Theory, Enhancement, and the Taking of the Redwood National Park," *American Journal of Agricultural Economics*, 79(2), May 1997:287–298.

Berck, Peter, and Michael Roberts, "Natural Resource Prices: Will They Ever Turn Up?" *Journal of Environmental Economics and Management*, 31(1), July 1996:65–78.

Bezmen, Trisha L., and Craig A. Depken II, "Influences on Software Piracy: Evidence from the Various United States," *Economic Letters*, 90(3), March 2006:356–361.

Bielen, David A., Richard G. Newell, and William A. Pizer, "Who Did the Ethanol Tax Credit Benefit? An Event Analysis of Subsidy Incidence," NBER Working Paper, 2016.

Blanciforti, Laura Ann, "The Almost Ideal Demand System Incorporating Habits: An Analysis of Expenditures on Food and Aggregate Commodity Groups." Doctoral thesis, University of California, Davis, 1982.

Blau, David, and Erdal Tekin, "The Determinants and Consequences of Child Care Subsidy Receipt by Low-Income Families," in Bruce Meyer and Greg Duncan, *The Incentives of Government Programs and the Well-Being of Families*, Joint Center for Poverty Research, 2001.

Blundell, Richard, Joel L. Horowitz, and Matthias Parey, "Measuring the Price Responsiveness of Gasoline Demand: Economic Shape Restrictions and Nonparametric Demand Estimation," *Quantitative Economics*, 3(1), March 2012:29–51.

Bollinger, Bryan, Phillip Leslie, and Alan Sorensen, "Calorie Posting in Chain Restaurants," *American Economic Journal: Economic Policy*, 3(1), February 2011:91–128.

Bonoli, Giuliano and Karl Hinrichs, "Statistical Discrimination and Employers' Recruitment Practices for Low-skilled Workers," Working Papers on the Reconciliation of Work and Welfare in Europe, 2010, http://www.oecd.org/els/47089328.pdf

Boomhower, Judson, "Drilling Like There's No Tomorrow: Bankruptcy, Insurance, and Environmental Risk," Stanford Institute for Economic Policy Research Working Paper, 2016.

Borenstein, Severin, "ANWR Oil and the Price of Gasoline," U.C. Energy Institute, *Energy Notes*, 3(2), June 2005.

Borenstein, Severin, "The Private Net Benefits of Residential Solar PV: And Who Gets Them," Energy Institute at Haas Working Paper, May 2015.

Borenstein, Severin, and Nancy L. Rose, "Competition and Price Dispersion in the U.S. Airline Industry," *Journal of Political Economy*, 102(4), August 1994:653–683.

Borenstein, Severin, James Bushnell, and Matthew Lewis, "Market Power in California's Gasoline Market," University of California Energy Institute, CSEM WP 132, May 2004, http://www.ucei.berkeley.edu/PDF/csemwp132.pdf.

Borjas, George J., "The Labor Demand Curve Is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market," *Quarterly Journal of Economics*, 118(4), November 2003:1335–1374.

Bradbury, Hinton, and Karen Ross, "The Effects of Novelty and Choice Materials on the Intransitivity of Preferences of Children and Adults," *Annals of Operations Research*, 23(1–4), June 1990:141–159.

Brander, James A., and M. Scott Taylor, "The Simple Economics of Easter Island: A Ricardo-Malthus Model of Renewable Resource Use," *American Economic Review*, 88(1), March 1998:119–138.

Brander, James A., and Anming Zhang, "Market Conduct in the Airline Industry: An Empirical Investigation," *Rand Journal of Economics*, 21(4), Winter 1990:567–583.

Brown, Jennifer, Justine Hastings, Erin T. Mansur, and Sofia B. Villas-Boas, "Reformulating Competition: Gasoline Content Regulation and Wholesale Gasoline Prices," *Journal of Environmental Economics and Management*, 55(1), January 2008:1–19.

Brown, Jennifer, and John Morgan, "Reputation in Online Auctions: The Market for Trust," *California Management Review*, 49(1), Fall 2006:61–81.

Brown, Jennifer, and John Morgan, "How Much Is a Dollar Worth? Tipping versus Equilibrium Coexistence on Competing Online Auction Sites," *Journal of Political Economy*, 117(4), August 2009:668–700.

Brown, Stephen P. A., and Daniel Wolk, "Natural Resource Scarcity and Technological Change," *Economic and Financial Review* (Federal Reserve Bank of Dallas), First Quarter 2000:2–13.

Brownlee, Oswald, and George Perry, "The Effects of the 1965 Federal Excise Tax Reductions on Prices," *National Tax Journal*, 20(3), September 1967:235–249.

Bruhn, M., "Can Wage Subsidies Boost Employment in the Wake of an Economic Crisis? Evidence from Mexico," Discussion Paper No. 9995, The Institute for the Study of Labor, June 2016.

Bruich, Gregory A., "The Effect of SNAP Benefits on Expenditures: New Evidence from Scanner Data and the November 2013 Benefit Cuts," Harvard Working Paper 2014.

Brunk, Gregory G., "A Test of the Friedman-Savage Gambling Model," *Quarterly Journal of Economics*, 96(2), May 1981:341–348.

Bucks, Brian K., Arthur B. Kennickell, Traci L. Mach, and Kevin B. Moore, "Changes in U.S. Family Finances from 2004 to 2007: Evidence from the Survey of Consumer Finances," *Federal Reserve Bulletin*, 2009.

Busch, Susan H., Mireia Jofre-Bonet, Tracy Falba, and Jody Sindelar, "Burning a Hole in the Budget: Tobacco Spending and Its Crowd-out of Other Goods," *Applied Health Economics and Policy*, 3(4), 2004:263–272.

Buschena, David E., and Jeffrey M. Perloff, "The Creation of Dominant Firm Market Power in the Coconut Oil Export Market," *American Journal of Agricultural Economics*, 73(4), November 1991:1000–1008.

Caliendo, Marco Michel Clement, Dominik Papies, and Sabine Scheel-Kopeinig, "The Cost Impact of Spam Filters: Measuring the Effect of Information System Technologies in Organizations," *Information Systems Research*, 2012:1068–1080.

Callison, Kevin, and Robert Kaestner, "Do Higher Tobacco Taxes Reduce Adult Smoking? New Evidence of the Effect of Recent Cigarette Tax Increases on Adult Smoking," NBER Working Paper, 2012.

Camerer, Colin F., *Behavioral Game Theory: Experiments in Strategic Interaction.* New York: Russell Sage Foundation, 2003.

Camerer, Colin F., George Loewenstein, and Matthew Rabin, eds., *Advances in Behavioral Economics*. New York: Russell Sage Foundation, 2004.

Card, David, "Immigration and Inequality," *American Economic Review*, 99(2), May 2009:1–21.

Card, David, and Alan B. Krueger, *Myth and Measurement: The New Economics of the Minimum Wage*. Princeton, NJ: Princeton University Press, 1995.

Carden, Art, "Beliefs, Bias, and Regime Uncertainty after Hurricane Katrina," *International Journal of Social Economics*, 35(7), 2008:531–545.

Carey, Kathleen, James F. Burgess Jr., and Gary J. Young, "Economies of Scale and Scope: The Case of Specialty

Hospitals," *Contemporary Economic Policy*, 33(1), January 2015:104–117.
Carlton, Dennis W., and Jeffrey M. Perloff, *Modern Industrial Organization*, 4th ed. Reading, MA: Addison Wesley Longman, 2005.
Carman, Hoy F., "Offsetting Price Impacts from Imports with Generic Advertising and Promotion Programs: The Hass Avocado Promotion and Research Order," *Review of Agricultural Economics*, 28(4), January 2006:463–481.
Centre for Science and Environment, "Congestion on Delhi Roads has Worsened—Says New analysis by CSE of Latest Google Map Data," 2017, http://www.cseindia.org/content/congestion-delhi-roads-has-worsened-%E2%80%93-says-new-analysis-cse-latest-google-map-data
Cesarini, David, et al., "The Effect of Wealth on Individual and Household Labor Supply: Evidence from Swedish Lotteries," NBER Working Paper, 2015.
Chandra, Ambarish, and Matthew Weinberg, "How Does Advertising Depend on Competition? Evidence from U.S. Brewing," University of Toronto manuscript, 2015.
Chetty, Raj, Adam Looney, and Kory Kroft, "Salience and Taxation: Theory and Evidence," *American Economic Review*, 99(4), September 2009:1145–1177.
Cho, Daegon, Pedro Ferreira, and Rahul Telang, "The Impact of Mobile Number Portability on Price and Consumer Welfare," Carnegie Mellon University Working Paper, 2015.
Chouinard, Hayley H., David E. Davis, Jeffrey T. LaFrance, and Jeffrey M. Perloff, "Fat Taxes: Big Money for Small Change," *Forum for Health Economics & Policy*, 10(2, 2), 2007.
Chouinard, Hayley, and Jeffrey M. Perloff, "Incidence of Federal and State Gasoline Taxes," *Economic Letters,* 83(1), April 2004:55–60.
Chouinard, Hayley H., and Jeffrey M. Perloff, "Gasoline Price Differences: Taxes, Pollution Regulations, Mergers, Market Power, and Market Conditions," *The B.E. Journal of Economic Analysis & Policy*, 7(1:8) (Contributions), 2007.
Clean Energy Finance Corporation, "A Market Report by the Clean Energy Finance Corporation," 2016, November, http://www.cefc.com.au/media/222701/cefc-energy-from-waste-market-report-november-2016.pdf
Clemens, Jeffrey, and Joshua D. Gottlieb, "Do Physicians' Financial Incentives Affect Medical Treatment and Patient Health?" *American Economic Review*, 104(4), April 2014:1320–1349.
Coady, David, et al., "How Large Are Global Energy Subsidies?" CESifo Working Paper, 2016.
Coase, Ronald H., "The Problem of Social Cost," *Journal of Law and Economics*, 3, October 1960:1–44.
Cohen, Jeffrey P., and Catherine Morrison Paul, "Scale and Scope Economies for Drug Abuse Treatment Costs: Evidence for Washington State," *Applied Economics*, 43(30), 2011:4827–4834.
Collard-Wexler, Allan, "Mergers and Sunk Costs: An Application to the Ready-Mix Concrete Industry," *American Economic Journal: Microeconomics*, 6(4), November 2014:407–447.
Competition & Markets Authority, "BT Group plc and EE Limited: A Report on the Anticipated Acquisition by BT Group plc of EE Limited," 2016, January 15, https://assets.publishing.service.gov.uk/media/56992242ed915d4747000026/BT_EE_final_report.pdf
Cookson, J. Anthony, "Anticipated Entry and Entry Deterrence: Evidence from the American Casino Industry," University of Colorado Working Paper, 2015.
Cranfield, John, "Canadian Beef Demand Elasticity Study," University of Guelph Working Paper, 2012.
Crawford, David, John Del Roccili, and Richard Voith, "Comments on Proposed Tax Reforms," Econsult Corporation, June 9, 2004.
Crost, Benjamin, and Santiago Guerrero, "The Effect of Alcohol Availability on Marijuana Use: Evidence from the Minimum Legal Drinking Age," *Journal of Health Economics*, 31(1), January 2012:112–121.
Currie, Janet, and Firouz Gahvari, "Transfers in Cash and In-Kind: Theory Meets the Data," *Journal of Economic Literature*, 46(2) June, 2008:333–383.
Currie, Janet, and Reed Walker, "Traffic Congestion and Infant Health: Evidence from E-ZPass." *American Economic Journal: Applied Economics*, 3(1), January 2011:65–90.
David, Guy, and Sara Markowitz, "Side Effects of Competition: The Role of Advertising and Promotion in Pharmaceutical Markets," NBER Working Paper, 2011.
Davies, James B., Susanna Sandström, Anthony Shorrocks, and Edward N. Wolff, "Estimating the Level and Distribution of Global Household Wealth," World Institute for Development Economic Research Working Paper, 2007.
Davis, Cale Wren, "An Analysis of the Enactment of Anti-Price Gouging Laws," Master of Science thesis, Montana State University, 2008.
Davis, Lucas, "Goin' for Broke in Texas . . . Protecting the Environment without Slowing Economic Growth,"https://energyathaas.wordpress.com/2014/12/01/goin-for-broke-in-texas-protecting-the-environment-without-slowing-economic-growth, 2014.
Davis, Lucas W., "The Environmental Cost of Global Fuel Subsidies," NBER Working Paper, 2016.
Davis, Lucas W., and Alex Chun, "A Deeper Look into the Fragmented Residential Solar Market," https://energyathaas.wordpress.com/2015/06/08/a-deeper-look-into-the-fragmented-residential-solar-market/, June 8, 2015.
Davis, Lucas W., and Lutz Kilian, "The Allocative Cost of Price Ceilings in the U.S. Residential Market for Natural Gas," *Journal of Political Economics*, 119(2), April 2011:212–241.
Davis, Lucas W., and Erich Muehlegger, "Do Americans Consume Too Little Natural Gas? An Empirical Test of Marginal Cost Pricing," *RAND Journal of Economics*, 41(4), Winter 2010:791–810.
Deacon, Robert T., and Jon Sonstelie, "The Welfare Costs of Rationing by Waiting," *Economic Inquiry*, 27(2), April 1989:179–196.
de Mello, Joao, Daniel Mejia, and Lucia Suarez, "The Pharmacological Channel Revisited: Alcohol Sales Restrictions and Crime in Bogotá," Documento CEDE No. 2013-20, 2013.
DellaVigna, Stefano, "Psychology and Economics: Evidence from the Field," *Journal of Economic Literature*, 47(2), June 2009:315–372.
Devine, Hilary, Tinh Doan, and Philip Stevens, "Explaining Productivity Distribution in New Zealand Industries: The Effects of Input Quality on Firm Productivity Differences," New Zealand Ministry of Economic Development, 2012.

Dingman, Deirdre A., et al., "Does Providing Nutrition Information at Vending Machines Reduce Calories per Item Sold?" *Journal of Public Health Policy*, 36(1), February 2015:110–122.

Dixit, Avinash K., and Robert S. Pindyck, *Investment Under Uncertainty*. Princeton, NJ: Princeton University Press, 1994.

Duan, Ying, Edith S. Hotchkiss, and Yawen Jiao, "Corporate Pensions and Financial Distress," SSRN Working Paper, January 2015.

Duffy-Deno, Kevin T., "Business Demand for Broadband Access Capacity," *Journal of Regulatory Economics*, 24(3), 2003:359–372.

Easterlin, Richard A., "Happiness and Economic Growth: The Evidence," in Wolfgang Glatzer, et al., eds., *Global Handbook of Quality of Life: Exploration of Well-Being of Nations and Continents*, Springer, 2015.

Easterlin, Richard A., "Paradox Lost?" U.S.C. Dornsife Institute for New Economic Thinking Working Paper, 2016.

Eastwood, David B., and John A. Craven, "Food Demand and Savings in a Complete, Extended, Linear Expenditure System," *American Journal of Agricultural Economics*, 63(3), August 1981:544–549.

Eckel, Catherine C., and Philip J. Grossman, "Altruism in Anonymous Dictator Games, *Games and Economic Behavior*, 16, 1996:181–191.

Economides, Nicholas, "The Economics of Networks," *International Journal of Industrial Organization,* 14(6), October 1996:673–699.

Econsult Corporation, *Choosing the Best Mix of Taxes for Philadelphia: An Econometric Analysis of the Impacts of Tax Rates on Tax Bases, Tax Revenue, and the Private Economy,* Report to the Philadelphia Tax Reform Commission, 2003.

Ecouterre, "INFOGRAPHIC: What Does That $14 Shirt From Bangladesh Really Cost?" n.d., http://www.ecouterre.com/infographic-how-much-does-that-14-t-shirt-really-cost/

Edelman, Benjamin, "Adverse Selection in Online 'Trust' Certifications and Search Results," *Electronic Commerce Research and Applications*, 10(1), January–February, 2011:17–25.

Edlin, Aaron S., and Pinar Karaca-Mandic, "The Accident Externality from Driving," *Journal of Political Economy*, 114(5), October 2006:931–955.

EE Limited, "EE Results for the Year Ended 31 December 2015," 2016, February 12, http://ee.co.uk/our-company/financials/2016/02/12/EE-results-for-the-year-ended-31-December-2015

Einav, Liran, Dan Knoepfle, Jonathan D. Levin, and Neel Sundaresan, "Sales Taxes and Internet Commerce," NBER Working Paper, 2012.

Ellsberg, Daniel, "Risk, Ambiguity, and the Savage Axioms," *Quarterly Journal of Economics*, 75(4), November 1961:643–669.

Engström, Per, and Eskil Forsell, "Demand Effects of Consumers' Stated and Revealed Preferences," Uppsala University Working Paper, 2015.

Environmental Protection Agency, *The Benefits and Costs of the Clean Air Act from 1990 to 2020*, http://www.epa.gov/oar/sect812/feb11/fullreport.pdf, 2011.

Epple, Dennis, Brett Gordon, and Holger Sieg, "A New Approach to Estimating the Production Function for Housing," *American Economic Review*, 100(3), June 2010:905–924.

European Commission, "Antitrust: Commission Fines Producers of Washing Pwder € 315.2 Million in Cartel Settlement Case," 2013, April 13, Brussels, http://europa.eu/rapid/press-release_IP-11-473_en.htm

European Commission, "Weekly Oil Bulletin," June 12, 2017, http://ec.europa.eu/energy/en/data-analysis/weekly-oil-bulletin.

Evers, Michiel, Ruud De Mooij, and Daniel Van Vuuren, "The Wage Elasticity of Labour Supply: A Synthesis of Empirical Estimates," *De Economist*, 156(1), March 2008:25–43.

Fan, Victoria Y., Dean T. Jamison, and Lawrence H. Summers, "The Inclusive Cost of Pandemic Influenza Risk," NBER Working Paper, 2016.

Farrell, Joseph, and Matthew Rabin, "Cheap Talk," *Journal of Economic Perspectives*, 10(3), Summer 1996:103–118.

Farrell, Lisa, Roger Hartley, Gauthier Lanot, and Ian Walker, "The Demand for Lotto," *Journal of Business & Economic Statistics*, 18(2), April 2000:228–241.

Feng, Cong, Scott Fay, and K. Sivakumar, "Overbidding in Electronic Auctions: Factors Influencing the Propensity to Overbid and the Magnitude of Overbidding," *Journal of the Academy of Marketing Science*, 44(2), March 2016:241–260.

Financial Times, "Shipping Size Obsession Could Be Ending, Study Finds," 2016, March 8, https://www.ft.com/content/255d9394-e47a-11e5-a09b-1f8b0d268c39

Fisher, Franklin M., "The Social Cost of Monopoly and Regulation: Posner Reconsidered," *Journal of Political Economy*, 93(2), April 1985:410–416.

Flath, David, "Industrial Concentration, Price-Cost Margins, and Innovation," *Japan and the World Economy*, 23(2), March 2011:129–139.

Fox, Jeremy T., and Valerie Smeets, "Does Input Quality Drive Measured Differences in Firm Productivity?" *International Economic Review*, 52(4), November 2011:961–989.

Friedman, Milton, and Leonard J. Savage, "The Utility Analysis of Choices Involving Risk," *Journal of Political Economy*, 56(4), August 1948:279–304.

Fryer Jr., Roland G., Steven D. Levitt, John A. List, and Sally Sadoff, "Enhancing the Efficacy of Teacher Incentives through Loss Aversion: A Field Experiment," NBER Working Paper, 2012.

Fu, Chao, and Jesse Gregory, "Estimation of an Equilibrium Model with Externalities: Combining the Strengths of Structural Models and Quasi-Experiments," NBER Working Paper, 2016.

Fudenberg, Drew, and Jean Tirole, *Game Theory*. Cambridge, MA: MIT Press, 1991.

Garratt, Rod, Mark Walker, and John Wooders, "Behavior in Second-Price Auctions by Highly Experienced eBay Buyers and Sellers," *Experimental Economics*, 15(1), March 2012:44–57.

Gasmi, Farid, Jean-Jacques Laffont, and Quang H. Vuong, "Econometric Analysis of Collusive Behavior in a Soft-Drink Market," *Journal of Economics and Management Strategy*, 1(2), Summer 1992, 277–311.

Ghose, Anindya and Sang Pil Han, "Estimating Demand for Mobile Applications in the New Economy," *Management Science*, 60(6), June 2014:1470–1488.

Gibbons, Robert, *Game Theory for Applied Economists*. Princeton, NJ: Princeton University Press, 1992.

Gillen, David, and Hamed Hasheminia, "Estimating the Demand Responses for Different Sizes of Air Passenger Groups," *Transportation Research Part B*, 49, 2013:24–38.

Goldfarb, Avi, and Catherine E. Tucker, "Search Engine Advertising: Channel Substitution When Pricing Ads to Context," *Management Science*, 57(3), March 2011:458–470.

Goldfarb, Avi, "What Is Different About Online Advertising?" *Review of Industrial Organization*, 44(2), March 2014:115–129.

Goldin, Jacob, and Tatiana Homonoff, "Smoke Gets in Your Eyes: Cigarette Tax Salience and Regressivity," *American Economic Journal: Economic Policy*, 5(1), February 2013:302–336.

Golec, Joseph, and Maurry Tamarkin, "Do Bettors Prefer Long Shots Because They Are Risk Lovers, or Are They Just Overconfident?" *Journal of Risk and Uncertainty*, 11(1), July 1995:51–64.

Gonçalves, Ricardo, and Pedro Pita Barros, "Economies of Scale and Scope in the Provision of Diagnostic Techniques and Therapeutic Services in Portuguese Hospitals," *Applied Economics*, 45(4), February 2013:415–433.

González, Tanja Markus Schmid, and David Yermack, "Smokescreen: How Managers Behave When They Have Something to Hide," NBER Working Paper, 2013.

GOV.UK, "Statutory Sick Pay (SSP)" https://www.gov.uk/statutory-sick-pay/overview

Grabowski, David C., and Michael A. Morrissey, "Do Higher Gasoline Taxes Save Lives?" *Economics Letters*, 90(1), January 2006:51–55.

Green, Richard, Richard Howitt, and Carlo Russo, "Estimation of Supply and Demand Elasticities of California Commodities," manuscript, May 2005.

Grossman, Michael, and Frank Chaloupka, "Demand for Cocaine by Young Adults: A Rational Addiction Approach," *Journal of Health Economics*, 17(4), August 1998:427–474.

Gruber, Jonathan H., and Sendhil Mullainathan, "Do Cigarette Taxes Make Smokers Happier," *Advances in Economic Analysis & Policy*, 5(1), 2005:article 4.

Guinnane, Timothy W., "The Historical Fertility Transition: A Guide for Economists," *Journal of Economic Literature*, 49(3), September 2011:589–614.

Hall, Robert E., and David M. Lilien, "Efficient Wage Bargains Under Uncertain Supply and Demand," *American Economic Review*, 69(5), December 1979:868–879.

Harkness, Joseph, and Sandra Newman, "The Interactive Effects of Housing Assistance and Food Stamps on Food Spending," *Journal of Housing Economics*, 12(3), September 2003:224–249.

Haskel, Jonathan, and Raffaella Sadun, "Regulation and UK Retailing Productivity: Evidence from Micro Data," *Economica*, 79, July 2012:425–448.

Hasker, Kevin, Bibo Jiang, and Robin C. Sickles, "Estimating Consumer Surplus in eBay Computer Monitor Auctions," in R. C. Sickles and W. C. Horrace, eds., *Festschrift in Honor of Peter Schmidt: Econometric Methods and Applications*, New York: Springer, 2014, Chapter 4:83–102.

Hausman, Catherine, and Ryan Kellogg, "Welfare and Distributional Implications of Shale Gas," *Brookings Papers on Economic Activity*, 50(1), Spring 2015:71–139.

Hay, George A., and Daniel Kelley, "An Empirical Survey of Price-Fixing Conspiracies," *Journal of Law and Economics*, 17(1), April 1974:13–38.

Helliwell, John, Richard Layard, and Jeffrey Sachs, *World Happiness Report*, The Earth Institute, Columbia University, 2012, http://www.earth.columbia.edu/articles/view/2960.

Hendel, Igal, and Aviv Nevo, "Intertemporal Price Discrimination in Storable Goods Markets," *American Economic Review*, 103(7), December 2013:2722–2751.

Henderson, Jason, "FAQs about Mad Cow Disease and Its Impacts," *The Main Street Economist*, December 2003.

Herod, Roger, "Analyzing the Metrics of Global Mobility Programs," *International HR Journal*, Summer 2008:9–15.

Hill, Jason, S. Polasky, E. Nelson, D. Tilman, H. Huo, L. Ludwig, J. Neumann, H. Zheng, and D. Bonta, "Climate Change and Health Costs of Air Emissions from Biofuels and Gasoline," *Proceedings of the National Academy of Sciences*, 106(6), February 2009:2077–2082.

Ho, Jason, Tirtha Dhar, and Charles Weinberg, "Playoff Payoff: Super Bowl Advertising for Movies," *International Journal of Research in Marketing*, 26(3), September 2009:168–179.

Holt, Matthew, "A Multimarket Bounded Price Variation Model Under Rational Expectations: Corn and Soybeans in the United States," *American Journal of Agricultural Economics*, 74(1), February 1992:10–20.

Hossain, Md. Moyazzem Tapati Basak, and Ajit Kumar Majumder, "An Application of Non-Linear Cobb-Douglas Production Function to Selected Manufacturing Industries in Bangladesh," *Open Journal of Statistics*, 2, October 2012:460–468.

Hossain, Tanjim, and John A. List, "The Behavioralist Visits the Factory: Increasing Productivity Using Simple Framing Manipulations," *Management Science*, 58(12), December 2012:2151–2167.

Hotelling, Harold, "The Economics of Exhaustible Resources," *Journal of Political Economy*, 39(2), April 1931:137–175.

Howard, David H., Guy David, and Jason Hockenberry, "Selective Hearing: Physician-Ownership and Physicians' Response to New Evidence," NBER Working Paper, 2016.

Hoynes, Hilary W., and Diane Whitmore Schanzenbach, "Consumption Responses to In-Kind Transfers: Evidence from the Introduction of the Food Stamp Program," *American Economic Journal: Applied Economics*, 1(4), October 2009:109–139.

Hsieh, Wen-Jen, "Test of Variable Output and Scale Elasticities for 20 U.S. Manufacturing Industries," *Applied Economics Letters,* 2(8), August 1995:284–287.

Humphreys, Brad R., and Hyunwoong Pyun, "Monopsony Exploitation in Professional Sport: Evidence from Major League Baseball Position Players, 2000–2011," *Managerial and Decision Economics*, 2016 online version.

Imas, Alex, Sally Sadoff, and Anya Samek, "Do People Anticipate Loss Aversion?" *Management Science*, forthcoming.

Imbens, Guido W., Donald B. Rubin, and Bruce I. Sacerdote, "Estimating the Effect of Unearned Income on Labor Earnings, Savings, and Consumption: Evidence from a Survey of Lottery Players," *American Economic Review*, 91(4), September 2001:778–794.

Irwin, Douglas A., "The Welfare Cost of Autarky: Evidence from the Jeffersonian Trade Embargo, 1807–09," *Review of International Economics*, 13(4), September 2005:631–645.

Irwin, Douglas A., and Nina Pavcnik, "Airbus versus Boeing Revisited: International Competition in the Aircraft Market," *Journal of International Economics*, 64(2), December 2004:223–245.

Ito, Harumi, and Darin Lee, "Assessing the Impact of the September 11 Terrorist Attacks on U.S. Airline Demand," *Journal of Economics and Business*, 57(1), January–February 2005:75–95.

ITU, "ICT Facts and Figures 2017," https://www.itu.int/en/ITU-D/Statistics/Pages/facts/default.aspx.

Jacobson, Michael F., and Kelly D. Brownell, "Small Taxes on Soft Drinks and Snack Foods to Promote Health," *American Journal of Public Health*, 90(6), June 2000:854–857.

Jagannathan, Ravi, Ellen R. McGrattan, and Anna Scherbina, "The Declining U.S. Equity Premium," *Federal Reserve Bank of Minneapolis Quarterly Review*. 24(4), 2000:3–19.

Jaspers, J.D., "Managing Cartels: How Cartel Participants Create Stability in the Absence of Law," European Journal on Criminal Policy and Research, 23(3), 2017: 319–335, https://link.springer.com/article/10.1007/s10610-016-9329-7

Jensen, Robert T., and Nolan H. Miller, "Giffen Behavior and Subsistence Consumption," *American Economic Review*, 98(4), September 2008:1553–1577.

Jetter, Karen M., James A. Chalfant, and David A. Sumner, "Does 5-a-Day Pay?" *AIC Issues Brief*, No. 27, August 2004.

Jhamb, Jordan, Dhaval Dave, and Gregory Colman, "The Patient Protection and Affordable Care Act and the Utilization of Health Care Services among Young Adults," *International Journal of Health and Economic Development*, 1(1), January 2015:8–25.

Johnson, Ronald N., and Charles J. Romeo, "The Impact of Self-Service Bans in the Retail Gasoline Market," *Review of Economics and Statistics*, 82(4), November 2000:625–633.

Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler, "Experimental Tests of the Endowment Effect and the Coase Theorem," *Journal of Political Economy*, 98(6), December 1990:1325–1348.

Kahneman, Daniel, and Amos Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47(2), March 1979:313–327.

Kan, Kamhon, "Cigarette Smoking and Self-Control," *Journal of Health Economics*, 26(1), January 2007:61–81.

Karp, Larry, "Global Warming and Hyperbolic Discounting," *Journal of Public Economics*, 89(2–3), February 2005:261–282.

Katz, Michael L., and Carl Shapiro, "Systems Competition and Network Effects," *Journal of Economic Perspectives*, 8(2), 1994:93–115.

Keane, Michael P., "Labor Supply and Taxes: A Survey," *Journal of Economic Literature*, 49(4), December 2011:961–1075.

Kennickell, Arthur B., "An Examination of the Changes in the Distribution of Wealth from 1989 to 1998: Evidence from the Survey of Consumer Finances," Federal Reserve, 2001.

Kennickell, Arthur B., "A Rolling Tide: Changes in the Distribution of Wealth in the U.S., 1989–2001," Federal Reserve Board, 2003.

Kennickell, Arthur B., "Ponds and Streams: Wealth and Income in the U.S., 1989 to 2007," Federal Reserve Board, 2009.

Kennickell, Arthur B., "Tossed and Turned: Wealth Dynamics of U.S. Households 2007–2009," Federal Reserve Board, 2011.

Kijiji, "We're All a Part of It! The Kijiji Second-hand Economy Index 2017 Report," 2017, https://www.kijiji.ca/kijijicentral/app/uploads/2016/08/2017-Index1.pdf

Kim, DaeHwan, and J. Paul Leigh, "Are Meals at Full-Service and Fast-Food Restaurants Normal or Inferior?" *Population Health Management*, 14(6), December 2011:307–315.

Kim, Jim-Yoo, *When Are Super Bowl Advertisings Super?* University of Texas at Arlington Ph.D. dissertation, 2011.

Kleiner, Morris M., and Alan B. Krueger, "Analyzing the Extent and Influence of Occupational Licensing on the Labor Market," *Journal of Labor Economics*, April 2013, 31(2):S173–S202.

Kleiner, Morris M., and Kyoung Won Park, "Life, Limbs, and Licensing: Occupational Regulation, Wages, and Workplace Safety of Electricians, 1992–2007," *Monthly Labor Review*, January 2014.

Kleiner, Morris M., "Reforming Occupational Licensing Policies," The Hamilton Project, Discussion Paper 2015-01, Brookings, 2015.

Klemperer, Paul, *Auctions: Theory and Practice*. Princeton, NJ: Princeton University Press, 2004.

Knetsch, Jack L., "Preferences and Nonreversibility of Indifference Curves," *Journal of Economic Behavior and Organization*, 1992, 17(1):131–139.

Knittel, Christopher R., Douglas L. Miller, and Nicholas J. Sanders, "Caution, Drivers! Children Present: Traffic, Pollution, and Infant Health," *Review of Economics and Statistics*, 98(2), May 2016:350–366.

Konovalova, A.A., and E.V. Vidishcheva, "Elasticity of Demand in Tourism and Hospitality," *European Journal of Economic Studies*, 4(2), 2013.

Kowalski, Amanda E., "Estimating the Tradeoff Between Risk Protection and Moral Hazard with a Nonlinear Budget Set Model of Health Insurance, *International Journal of Industrial Organization*, 43, November 2015:122–135.

Krauss, Melissa J., Patricia Cavazos-Rehg, Andrew Plunk, Laura Bierut, and Richard Grucza, "Effects of State Cigarette Excise Taxes and Smoke-Free Air Policies on State Per Capita Alcohol Consumption in the United States, 1980 to 2009," *Alcoholism: Clinical and Experimental Research*, 38(10), October 2014:2630–2638.

Kristiansen, Tomas, "Empty Containers Cost Maersk Line USD 1 Billion a Year," ShippingWatch, December 31, 2012, http://shippingwatch.com/articles/article4891909.ece.

Kuroda, Sachiko, and Isamu Yamamoto, "Estimating Frisch Labor Supply Elasticity in Japan," *Journal of Japanese International Economies*, 22(4), December 2008:566–585.

Labonne, Julien, and Robert S. Chase, "So You Want to Quit Smoking: Have You Tried a Mobile Phone?" *Applied Economic Letters*, 18(2), 2011:103–106.

Lazear, Edward P., Kathryn L. Shaw, and Christopher T. Stanton, "Making Do with Less: Working Harder During Recessions," *Journal of Labor Economics*, 34(S1), January 2016:S333–S360.

Lazear, Edward P., Kathryn L. Shaw, and Christopher T. Stanton, "The Value of Bosses," *Journal of Labor Economics*, 33(4), October 2015:823–861.

Lee, David, and Alexandre Mas, "Long-Run Impacts of Unions on Firms: New Evidence from Financial Markets, 1961–1999," *Quarterly Journal of Economics*, 127(1), February 2012:333–378.

Lee, J. M., M. G. Chen, T. C. Hwang, and C. Y. Yeh, "Effect of Cigarette Taxes on the Consumption of Cigarettes, Alcohol, Tea and Coffee in Taiwan," *Public Health*, 124(8), August 2010:429–436.

Lee, Young Han, and Ulrike Malmendier, "The Bidder's Curse," *American Economic Review*, 101(2), April 2011:749–787.

Lee, Young-Wook, "Effects of a Universal Childcare Subsidy on Mothers' Time Allocation," *KDI Journal of Economic Policy*, 38(1), February 2016:1–22.

Leibenstein, Harvey, "Bandwagon, Snob, and Veblen Effects in the Theory of Consumers' Demand," *Quarterly Journal of Economics*, 64(2), May 1950:183–207.

Lelieveld, J., et al., "The Contribution of Outdoor Air Pollution Sources to Premature Mortality on a Global Scale," *Nature*, *525*(7569), September 2015:367–371.

Leslie, Phillip J., "A Structural Econometric Analysis of Price Discrimination in Broadway Theatre," working paper, University of California, Los Angeles, November 15, 1997.

Levitt, Steven D., and Jack Porter, "How Dangerous Are Drinking Drivers," *Journal of Political Economy*, 109(5), December 2001:1198–1237.

Liddle, Brantley, "The Systemic, Long-Run Relation among Gasoline Demand, Gasoline Price, Income, and Vehicle Ownership In OECD Countries: Evidence from Panel Cointegration and Causality Modeling," *Transportation Research Part D: Transport and Environment*, 17(4), June 2012:237–331.

List, John A., "Does Market Experience Eliminate Market Anomalies?" *Quarterly Journal of Economics*, 118(1), February 2003:41–71.

Lopez, Rigoberto A., and Emilio Pagoulatos, "Rent Seeking and the Welfare Cost of Trade Barriers," *Public Choice*, 79(1–2), April 1994:149–160.

MacAvoy, Paul W., *The Natural Gas Market*. New Haven: Yale University Press, 2000.

MacCrimmon, Kenneth R., and M. Toda, "The Experimental Determination of Indifference Curves," *Review of Economic Studies*, 56(3), July 1969:433–451.

MacEwan, Joanna P., Aaron Smith, and Julian M. Alston, "The Supplemental Nutrition Assistance Program, Energy Balance, and Weight Gain," *Food Policy*, 61, May 2016:103–120.

Machina, Mark, "Dynamic Consistency and Non-Expected Utility Models of Choice Under Uncertainty," *Journal of Economic Literature*, 27(4), December 1989:1622–1668.

Maclean's, "What Does that $14 Shirt Really Cost?" 2013, May 1, http://www.macleans.ca/economy/business/what-does-that-14-shirt-really-cost/

MacKie-Mason, Jeffrey K., and Robert S. Pindyck, "Cartel Theory and Cartel Experience in International Minerals Markets," in R. L. Gordon, H. D. Jacoby, and M. B. Zimmerman, eds., *Energy: Markets and Regulation: Essays in Honor of M. A. Adelman*. Cambridge, MA: MIT Press, 1986.

Madden, Janice F., *The Economics of Sex Discrimination*. Lexington, MA: Heath, 1973.

Madrian, Brigitte C., and Dennis F. Shea, "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior," *Quarterly Journal of Economics*, 116(4), November 2001:1149–1187.

Madrian, Brigitte C., and Dennis F. Shea, "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior: Erratum," *Quarterly Journal of Economics*, 117(1), February 2002:377.

Markowitz, Sara Erik Nesson, Eileen Poe-Yamagata, Curtis Florence, Partha Deb, Tracy Andrews, Sarah Beth L. Barnett, "Estimating the Relationship Between Alcohol Policies and Criminal Violence and Victimization," NBER Working Paper, 2012.

McPhail, Lihong Lu, and Bruce A. Babcock, "Impact of US Biofuel Policy on US Corn and Gasoline Price Variability," *Energy*, 37(1), February 2012:505–513.

Medoff, Marshall H., "Price, Restrictions and Abortion Demand," *Journal of Family and Economic Issues*, 28(4), December 2007:583–599.

Mishel, Lawrence, and Natalie Sabadish, "CEO Pay and the Top 1%," *EPI Issue Brief*, May 2, 2012.

Monteiro, Tiago, Marco Vasconcelos, and Alex Kacelnik, "Starlings Uphold Principles of Economic Rationality for Delay and Probability of Reward," *Proceedings of the Royal Society B*, 280(1756), April 2013:1–6.

Moschini, Giancarlo, and Karl D. Meilke, "Production Subsidy and Countervailing Duties in Vertically Related Markets: The Hog-Pork Case Between Canada and the United States," *American Journal of Agricultural Economics*, 74(4), November 1992:951–961.

Nash, John F., "Equilibrium Points in *n*-Person Games," *Proceedings of the National Academy of Sciences*, 36, 1950:48–49.

Nash, John F., "Non-Cooperative Games," *Annals of Mathematics*, 54(2), July 1951:286–295.

Nataraj, Shanthi, "Do Residential Water Consumers React to Price Increases? Evidence from a Natural Experiment in Santa Cruz," *Agricultural and Resource Economics Update*, 10(3), January/February 2007:9–11.

Neto, João Quariguasi Frota, Jacqueline Bloemhof, and Charles Corbett, "Market Prices of Remanufactured, Used and New Items: Evidence from eBay," *International Journal of Production Economics*, 171(3), 2016:371–380.

Neumark, David, J. M. Ian Salas, and William Wascher, "Revisiting the Minimum Wage-Employment Debate: Throwing Out the Baby with the Bathwater?" *Industrial Labor Relations Review*, 67(3 supplement), May 2014:608–648.

Nickell, S., and J. Saleheen, "The Impact of Immigration on Occupational Wages: Evidence from Britain," Staff Working Paper No. 574, Bank of England, December 2015.

Nikaj, Silda, and Frank J. Chaloupka, "The Effect of Prices on Cigarette Use Among Youths in the Global Youth Tobacco Survey," *Nicotine and Tobacco Research*, 2013.

O'Donoghue, Ted, and Matthew Rabin, "Doing It Now or Later," *American Economic Review*, 89(1), March 1999:103–124.

Orenstein, Walter A., Mark J. Papania, and Melinda E. Wharton, "Measles Elimination in the United States," *Journal of Infectious Diseases*, 189(Supplement 1), 2004:S1–S3.

Oxfam, "Dumping Without Borders," Oxfam Briefing Paper 50, 2003.

Oyebode, O., V. Gordon-Dseagu, A. Walker, et al., "Fruit and Vegetable Consumption and All-Cause, Cancer and CVD Mortality: Analysis of Health Survey for England Data," *J Epidemiol Community Health*, 68, 2014: 856–862, http://jech.bmj.com/content/jech/68/9/856.full.pdf.

Pal, Jeremy S., and Elfatih A. B. Eltahir, "Future Temperature in Southwest Asia Projected to Exceed a Threshold for

Human Adaptability," *Nature Climate Change*, 6, October 2015:197–200.
Panzar, John C., and Robert D. Willig, "Economies of Scale in Multi-Output Production," *Quarterly Journal of Economics*, 91(3), August 1977:481–493.
Panzar, John C., and Robert D. Willig, "Economies of Scope," *American Economic Review*, 71(2), May 1981:268–272.
Parry, Ian W. H., Margaret Walls, and Winston Harrington, "Automobile Externalities and Policies," *Journal of Economic Literature*, 45(2), June 2007:373–399.
Perloff, Jeffrey M., and Ximing Wu, "Tax Incidence Varies Across the Price Distribution," *Economic Letters*, 96(1), July 2007:116–119.
Perry, Martin K., "Forward Integration by Alcoa: 1888–1930," *Journal of Industrial Economics*, 29(1), September 1980:37–53.
Pesko, Michael F., et al., "The Influence of Geography and Measurement in Estimating Cigarette Price Responsiveness," NBER Working Paper, 2016.
Picchio, Matteo, Sigrid Suetens, and Jan C. van Ours, "Labor Supply Effects of Winning a Lottery," Tilburg University Working Paper 2015.
Pielke, Roger A., Jr., Joel Gratz, Christopher W. Landsea, Douglas Collins, Mark A. Saunders, and Rade Musulin, "Normalized Hurricane Damages in the United States: 1900–2005," *Natural Hazards Review*, 9(1), February 2008:29–42.
Plott, Charles R., and Kathryn Zeiler, "The Willingness to Pay—Willingness to Accept Gap, the 'Endowment Effect,' Subject Misconceptions, and Experimental Procedures for Eliciting Values," *American Economic Review*, 95(3), June 2005:530–545.
Posner, Richard A., "The Social Cost of Monopoly and Regulation," *Journal of Political Economy*, 83(4), August 1975:807–827.
Prescott, Edward C., "Why Do Americans Work So Much More Than Europeans?" *Federal Reserve Bank of Minneapolis Quarterly Review*, 28(1), July 2004:2–13.
Prince, Jeffrey T., "Repeat Purchase amid Rapid Quality Improvement: Structural Estimation of Demand for Personal Computers," *Journal of Economics & Management Strategy*, 17(1), Spring 2008:1–33.
Rabin, Matthew, "Psychology and Economics," *Journal of Economic Literature,* 36(1), March 1998:1–46.
Rao, Justin M., and David H. Reiley, "The Economics of Spam," *Journal of Economic Perspectives*, 26(3), Summer 2012:87–110.
Rawls, John, *A Theory of Justice*. New York: Oxford University Press, 1971.
Regan, Tracy L., "Generic Entry, Price Competition, and Market Segmentation in the Prescription Drug Market," *International Journal of Industrial Organization*, 26(4), July 2008:930–948.
Richards, Timothy J., and Luis Padilla, "Promotion and Fast Food Demand," *American Journal of Agricultural Economics*, 91(1), February 2009:168–183.
Roberts, Mark J., and Larry Samuelson, "An Empirical Analysis of Dynamic Nonprice Competition in an Oligopolistic Industry," *Rand Journal of Economics*, 19(2), Summer 1988:200–220.
Roberts, Michael J., and Wolfram Schlenker, "Is Agriculture Becoming More or Less Sensitive to Extreme Heat? Evidence from U.S. Corn and Soybean Yields." *The Design and Implementation of U.S. Climate Policy*, *NBER*, 2012.
Roberts, Michael J., and Wolfram Schlenker, "Identifying Supply and Demand Elasticities of Agricultural Commodities: Implications for the US Ethanol Mandate," *American Economic Review*, 103(6), October 2013:2265–2295.
Robidoux, Benoît, and John Lester, "Econometric Estimates of Scale Economies in Canadian Manufacturing," Working Paper No. 88-4, Canadian Dept. of Finance, 1988.
Robidoux, Benoît, and John Lester, "Econometric Estimates of Scale Economies in Canadian Manufacturing," *Applied Economics*, 24(1), January 1992:113–122.
Rohlfs, Jeffrey H., "A Theory of Interdependent Demand for a Communications Service," *Bell Journal of Economics and Management Science,* 5(1), Spring 1974:16–37.
Rohlfs, Jeffrey H., *Bandwagon Effects in High-Technology Industries,* Cambridge, MA: MIT Press, 2001.
Rosenberg, Howard R., "Many Fewer Steps for Pickers—A Leap for Harvestkind? Emerging Change in Strawberry Harvest Technology," *Choices,* 1st Quarter 2004:5–11.
Rossi, R. J., and T. Armstrong, Studies of Intercollegiate Athletics: Summary Results from the 1987–88 National Study of Intercollegiate Athletes (Report No. 1), Palo Alto, CA: Center for the Study of Athletics, American Institutes for Research, 1989.
Rousseas, S. W., and A. G. Hart, "Experimental Verification of a Composite Indifference Map," *Journal of Political Economy*, 59(4), August 1951:288–318.
Ruffin, R. J., "Cournot Oligopoly and Competitive Behavior," *Review of Economic Studies,* 38(116), October 1971:493–502.
Saez, Emmanuel, "Striking It Richer: The Evolution of Top Incomes in the United States (Updated with 2015 Estimates)," University of California, Berkeley Working Paper, 2016.
Salgado, Hugo, "A Dynamic Duopoly Model under Learning-by-doing in the Computer CPU Industry," U.C. Berkeley Ph.D. Dissertation Chapter, 2008.
Salop, Joanne, and Steven C. Salop, "Self-Selection and Turnover in the Labor Market," *Quarterly Journal of Economics*, 90(4), November 1976:619–627.
Salop, Steven C., "The Noisy Monopolist: Imperfect Information, Price Dispersion, and Price Discrimination," *Review of Economic Studies*, 44(3), October 1977:393–406.
Salop, Steven C., "Practices That (Credibly) Facilitate Oligopoly Coordination," in Joseph E. Stiglitz and G. Frank Mathewson, eds., *New Developments in the Analysis of Market Structure*. Cambridge, MA: MIT Press, 1986.
Salvo, Alberto, and Christian Huse, "Build It, But Will They Come? Evidence from Consumer Choice Between Gasoline and Sugarcane Ethanol," *Journal of Environmental Economics and Management*, 66(2), September 2013:251–279.
Sappington, David E. M., and Dennis L. Weisman, "Price Cap Regulation: What Have We Learned from 25 Years of Experience in the Telecommunications Industry?" *Journal of Regulatory Economics*, 38(3), 2010:227–257.
Sato, Yoshihiro and Måns Söderbom, "Are Larger Firms More Productive Because of Scale Economies? Evidence from Swedish Register Data," in Yoshihiro Sato, *Dynamic Investment Models, Employment Generation and Productivity—Evidence from Swedish Data*, University of Gothenburg Ph.D. Dissertation, 2012.

Scherer, F. M., "An Early Application of the Average Total Cost Concept," *Journal of Economic Literature*, 39(3), September 2001:897–901.

Schlenker, Wolfram, and Sofia B. Villas-Boas, "Consumer and Market Responses to Mad-Cow Disease," *American Economic Review*, 91(4), November 2009:1140–1152.

Schmalensee, Richard, Paul L. Joskow, A. Denny Ellerman, Juan Pablo Montero, and Elizabeth M. Bailey, "An Interim Evaluation of Sulfur Dioxide Emissions Trading," *Journal of Economic Perspectives*, 12(3), Summer 1998:53–68.

Schoemaker, Paul J. H., "The Expected Utility Model: Its Variants, Purposes, Evidence and Limitation," *Journal of Economic Literature*, 20(2), June 1982:529–563.

Shao, Benjamin B. M., and Winston T. Lin, "Assessing Output Performance of Information Technology Service Industries: Productivity, Innovation and Catch-Up," *International Journal of Production Economics*, 172, February 2016:43–53.

Shapiro, Carl, and Joseph E. Stiglitz, "Equilibrium Unemployment as a Worker Discipline Device," *American Economic Review*, 74(3), June 1984:434–444.

Shapiro, Carl, and Hal R. Varian, *Information Rules: A Strategic Guide to the Network Economy*. Boston: Harvard Business School Press, 1999.

Shapiro, Jesse M., "Is There a Daily Discount Rate? Evidence from the Food Stamp Nutrition Cycle," *Journal of Public Economics*, 89(2-3), February 2004:303–325.

Shapiro, Joseph S., and Reed Walker, "Why Is Pollution from U.S. Manufacturing Declining? The Roles of Trade, Regulation, Productivity, and Preferences," NBER Working Paper, 2015.

Sheehan-Connor, Damien, "Environmental Policy and Vehicle Safety: The Impact of Gasoline Taxes," *Economic Inquiry*, 53(3), July 2015:1606–1629.

Shelton, Cameron A., and Nathan Falk, "Policy Uncertainty and Manufacturing Investment: Evidence from U.S. State Elections," CESifo Working Paper, 2016.

Shiller, Ben, and Joel Waldfogel, "Music for a Song: An Empirical Look at Uniform Pricing and Its Alternatives," *The Journal of Industrial Economics*, 59(4), December 2011:630–660.

Slade, Margaret E., "Product Rivalry with Multiple Strategic Weapons: An Analysis of Price and Advertising Competition," *Journal of Economics and Management Strategy*," 4(3), Fall 1995:224–276.

Sood, Neeraj, Abby Alpert, and Jay Bhattacharya, "Technology, Monopoly, and the Decline of the Viatical Settlements Industry," NBER Working Paper, 2005.

Spence, A. Michael, *Market Signaling*. Cambridge, MA: Harvard University Press, 1974.

Stanfield, Jared, and Robert Tumarkin, "Does the Political Power of Non-Financial Stakeholders Affect Firm Values? Evidence from Labor Unions," University of New South Wales Working Paper, 2016.

Stewart, Hayden, and Rosanna Mentzer Morrison, "New Regulations Will Inform Consumers About Calories in Restaurant Foods," *Amber Waves*, February 2, 2015.

Stevenson, Betsey, and Justin Wolfers, "Subjective Well-Being and Income: Is There Any Evidence of Satiation?" *American Economic Review*, 103(3), May 2013:598–604.

Stiglitz, Joseph E., "The Theory of 'Screening,' Education, and the Distribution of Income," *American Economic Review*, 65(3), June 1975:283–300.

Stiglitz, Joseph E., "Equilibrium in Product Markets with Imperfect Information," *American Economic Review*, 69(2), May 1979:339–345.

Stiglitz, Joseph E., "The Causes and Consequences of the Dependence of Quality on Price," *Journal of Economic Literature*, 25(1), March 1987:1–48.

Subramanian, Ravi, and Ramanath Subramanyam, "Key Factors in the Market for Remanufactured Products," *Manufacturing & Service Operations Management*, 14(2), Spring 2012:315–326.

Sullivan, Jacob, "Mobile Travel Report," Kayak News, May 18, 2016, https://www.kayak.co.uk/news/mobile-travel-report-2016/.

Suzuki, Junichi, "Land Use Regulation as a Barrier to Entry: Evidence from the Texas Lodging Industry," *International Economic Review*, 54(2), 2013:495–523.

Swinton, John R., and Christopher R. Thomas, "Using Empirical Point Elasticities to Teach Tax Incidence," *Journal of Economic Education*, 32(4), Fall 2001:356–368.

Syverson, Chad, "Product Substitutability and Productivity Dispersion," *Review of Economics and Statistics*, 86(2), May 2004:534–550.

Taiki Nishimura, T., Y. Kajitani, and H. Tatano, "Damage Assessment in Tourism Caused by an Earthquake Disaster," *Journal of Integrated Disaster Risk Management*, 3(1), 2013.

Teillant, Aude, and Ramanan Laxminarayan, "Economics of Antibiotic Use in U.S. Swine and Poultry Production," *Choices*, 30(1), 1st Quarter 2015:1–11.

Tekin, Erdal, "Single Mothers Working at Night: Standard Work and Child Care Subsidies," *Economic Inquiry*, 45(2), April 2007:233–250.

Thaler, Richard H., *The Winner's Curse: Paradoxes and Anomalies of Economic Life*. Princeton, NJ: Princeton University Press, 1994.

The Australian, 2017, http://www.theaustralian.com.au/business/mergers-acquisitions/fairfax-nzme-appeal-over-blocked-merger/news-story/019ab8e6612ef2b43874060295 7af671

The Guardian, "Film and TV Streaming and Downloads Overtake DVD Sales for First Time," 2017, January 5, https://www.theguardian.com/media/2017/jan/05/film-and-tv-streaming-and-downloads-overtake-dvd-sales-for-first-time-netflix-amazon-uk

The Guardian, "London Stock Exchange and Deutsche Börse Merger Blocked by EU," 2017, https://www.theguardian.com/business/2017/mar/29/london-stock-exchange-deutsche-borse-merger-blocked-ec

The National Archives, "Dairy Retail Price Initiatives," 2011, http://webarchive.nationalarchives.gov.uk/20140402163102/http://www.oft.gov.uk/OFTwork/competition-act-and-cartels/ca98/decisions/dairy

The Telegraph, "Individual NHS Doctors Receiving £100,00 Per Year from Drug Firms," 2016, June 30, http://www.telegraph.co.uk/news/2016/06/30/individual-nhs-doctors-receiving-100000-per-year-from-drugs-firm/

The Telegraph, "Which Eurpoean Country Offers the Best Sick Leave?" 2016, http://www.telegraph.co.uk/business/0/which-european-country-pays-the-most-when-workers-take-sick-leav/

The University of Sydney, n.d., http://sydney.edu.au/education_social_work/future_students/careers/teacher_salaries.shtml

Tideman, T. Nicholaus, and Gordon Tullock, "A New and Superior Process for Making Social Choices," *Journal of Political Economy*, 84(6), December 1976:1145–1159.

Time, "Your Doctor Likely Orders More Tests Than You Actually Need," 2015, March 24, http://time.com/3754900/doctors-unnecessary-tests/

Tosun, Mehmet S., and Mark L. Skidmore, "Cross-Border Shopping and the Sales Tax: An Examination of Food Purchases in West Virginia," *The B.E. Journal of Economic Analysis & Policy*, 7(1), 2007:Article 63.

Town, Robert, Douglas Wholey, Roger Feldman, and Lawton R. Burns, "The Welfare Consequences of Hospital Mergers," NBER Working Paper, 2006.

Trabandt, Mathias, and Harald Uhlig, "The Laffer Curve Revisited," *Journal of Monetary Economics*, 58(4), May 2011:305–327.

Trabandt, Mathias, and Harald Uhlig, "How Do Laffer Curves Differ across Countries?" in A. Alesina and F. Giavazzi, eds., *Fiscal Policy after the Financial Crisis*, National Bureau of Economic Research, University of Chicago Press, 2013:211–249.

Tullock, G., "The Welfare Cost of Tariffs, Monopolies, and Theft," *Western Economic Journal*, 5(3), June 1967:224–232.

Tversky, Amos, and Daniel Kahneman, "The Framing of Decisions and the Psychology of Choice," *Science*, 211(4481), January 1981:453–458.

Tyler, John H., Richard J. Murnane, and John B. Willett, "Estimating the Labor Market Signaling Value of the GED," *Quarterly Journal of Economics*, 115(2), May 2000:431–468.

UCL News, "UCL Study Finds New Evidence Linking Fruit and Vegetable Consumption with Lower Mortality," April 1, 2014, http://www.ucl.ac.uk/news/news-articles/0414/010413-fruit-veg-consumption-death-risk.

Varian, Hal R., "Measuring the Deadweight Cost of DUP and Rent-Seeking Activities," *Economics and Politics*, 1(1), Spring 1989:81–95.

Veracierto, Marcelo, "Firing Costs and Business Cycle Fluctuations," *International Economic Review*, 49(1), February 2008:1–39.

Viscusi, W. Kip, *Employment Hazards*. Cambridge, MA: Harvard University Press, 1979.

von Neumann, John, and Oskar Morgenstern, *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press, 1944.

Waldfogel, Joel, "The Deadweight Loss of Christmas," *American Economic Review*, 83(5), December 1993:1328–1336.

Waldfogel, Joel, "Does Consumer Irrationality Trump Consumer Sovereignty?" *Review of Economics and Statistics*, 87(4), November 2005:691–696.

Waldfogel, Joel, *Scroogenomics*. Princeton, NJ: Princeton University Press, 2009.

Waldfogel, Joel, "Copyright Research in the Digital Age: Moving from Piracy to the Supply of New Products," *American Economic Review*, 102(3), May 2012:337–342.

Warner, John T., and Saul Pleeter, "The Personal Discount Rate: Evidence from Military Downsizing Programs," *American Economic Review*, 91(1), March 2001:33–53.

Webber, Douglas A., "Firm-Level Monopsony and the Gender Pay Gap," *Industrial Relations*, 55(2), April 2016:323–344.

Weinstein, Arnold A., "Transitivity of Preferences: A Comparison Among Age Groups," *Journal of Political Economy*, 76(2), March/April 1968:307–311.

Whitmore, Diane, "What Are Food Stamps Worth?" Princeton University Working Paper No. 468, July 2002.

Wiley Online Library, "Emergency Physician Perceptions of Medically Unnecessary Advanced Diagnostic Imaging," 2015, March 23, doi: 10.1111/acem.12625, http://onlinelibrary.wiley.com/doi/10.1111/acem.12625/abstract

Williamson, Oliver E., "Credible Commitments: Using Hostages to Support Exchange," *American Economic Review*, 73(4), September 1983:519–540.

Wong, L., E.A. Selvanathan, and Saroja Selvanathan, *Changing Pattern of Meat Consumption in Australia*. Queensland, Australia: Griffith Business School, 2013.

World Trade Organization, "Trade-related Aspects of Intellectual Property Rights," https://www.wto.org/english/docs_e/legal_e/27-trips_01_e.htm

Yellen, Janet L., "Efficiency Wage Models of Unemployment," *American Economic Review*, 74(2), May 1984:200–205.

Zhang, Rui Kai Sun, Michael S. Delgado, and Subal C. Kumbhakar, "Productivity in China's High Technology Industry," *Technology Forecasting and Social Change*, 79(1), 2012:127–141.

Zheng, Yuqing, Edward W. McLaughlin, and Harry M. Kaiser, "Salience and Taxation: Salience Effect vs. Knowledge Effect," *Applied Economics Letters*, 20(5), 2013.

Definitions

I hate definitions. —Benjamin Disraeli

action: a move that a player makes at a specified stage of a game, such as how much output a firm produces in the current period. (14)*

adverse selection: occurs when one party to a transaction possesses information about a hidden characteristic that is unknown to other parties and takes economic advantage of this information. (19)

asymmetric information: one party to a transaction has relevant information that another party lacks. (19)

auction: a sale in which property or a service is sold to the highest bidder. (14)

average cost (*AC*): the total cost divided by the units of output produced: $AC = C/q$. (7)

average fixed cost (*AFC*): the fixed cost divided by the units of output produced: $AFC = F/q$. (7)

average product of labor (AP_L): the ratio of output, q, to the number of workers, L, used to produce that output: $AP_L = q/L$. (6)

average variable cost (*AVC*): the variable cost divided by the units of output produced: $AVC = VC/q$. (7)

backward induction: first determine the best response by the last player to move, next determine the best response for the player who made the next-to-last move, and then repeat the process back to the move at the beginning of the game. (14)

bad: something for which less is preferred to more, such as pollution. (4)

bandwagon effect: the situation in which a person places greater value on a good as more and more other people possess it. (11)

barrier to entry: an explicit restriction or a cost that applies only to potential new firms—existing firms are not subject to the restriction or do not bear the cost. (9)

behavioral economics: by adding insights from psychology and empirical research on human cognition and emotional biases to the rational economic model, economists try to better predict economic decision making. (4)

Bertrand equilibrium (*Nash-Bertrand equilibrium* or *Nash-in-prices equilibrium*): a set of prices such that no firm can obtain a higher profit by choosing a different price if the other firms continue to charge these prices. (13)

best response: the strategy that maximizes a player's payoff given its beliefs about its rivals' strategies. (14)

bounded rationality: people have a limited capacity to anticipate, solve complex problems, or enumerate all options. (4)

budget line (or *budget constraint*): the bundles of goods that can be bought if the entire budget is spent on those goods at given prices. (4)

bundling (*package tie-in sale*): selling multiple goods or services for a single price. (12)

cartel: a group of firms that explicitly agrees to coordinate their activities. (13)

certification: a report that a particular product meets or exceeds a given standard. (19)

cheap talk: unsubstantiated claims or statements. (19)

club good: a good that is nonrival but is subject to exclusion. (18)

common knowledge (in a game): a piece of information that is known by all players, and it must be known by all players to be known by all players, and it must be known to be known to be known by all players, and so forth. (14)

(open-access) common property: a resource to which everyone has free access and an equal right to exploit. (18)

comparative advantage: the ability to produce a good at a lower opportunity cost than someone else. (10)

compensating variation (*CV*): the amount of money one would have to give a consumer to offset completely the harm from a price increase or take from a consumer to offset the benefit from a price decrease. (5)

complement: a good or service that is jointly consumed with another good or service. (2)

complete information (in a game): the situation where the payoff function is common knowledge among all players. (14)

constant returns to scale: property of a production function whereby when all inputs are increased by a certain percentage, output increases by that same percentage. (6)

consumer surplus (*CS*): the monetary difference between what a consumer is willing to pay for the quantity of the good purchased and what the good actually costs. (9)

contingent fee: a payment to a lawyer that is a share of the award in a court case (usually after legal expenses are deducted) if the client wins and nothing if the client loses.

*The numbers in the parentheses refer to the chapter in which the term was defined.

contract curve: the set of all Pareto-efficient bundles. (10)

cost (*total cost, C*): the sum of a firm's variable cost and fixed cost: $C = VC + F$. (7)

Cournot equilibrium (*Nash-Cournot equilibrium* or *Nash-in-prices equilibrium*): a set of quantities chosen by firms such that, holding the quantities of all other firms constant, no firm can obtain a higher profit by choosing a different quantity. (13)

credible threat: an announcement that a firm will use a strategy harmful to its rival and that the rival believes because the firm's strategy is rational in the sense that it is in the firm's best interest to use it. (14)

cross-price elasticity of demand: the percentage change in the *quantity demanded* in response to a given percentage change in the price of another good. (3)

deadweight loss (*DWL*): the net reduction in welfare from a loss of surplus by one group that is not offset by a gain to another group from an action that alters a market equilibrium. (9)

decreasing returns to scale: property of a production function whereby output increases less than in proportion to an equal percentage increase in all inputs. (6)

demand curve: the *quantity demanded* at each possible price, holding constant the other factors that influence purchases. (2)

demand function: the relationship between the quantity demanded, price, and other factors that influence purchases. (2)

discount rate: a rate reflecting the relative value an individual places on future consumption compared to current consumption. (16)

diseconomies of scale: property of a cost function whereby the average cost of production rises when output increases. (7)

dominant strategy: a strategy that produces a higher payoff than any other strategy the player can use for every possible combination of its rivals' strategies. (14)

duopoly: an *oligopoly* with two firms. (13)

durable good: a product that is usable for years. (7)

dynamic game: a game in which players move either sequentially or repeatedly. (14)

economic cost (*opportunity cost*): the value of the best alternative use of a resource. (7)

economic profit: revenue minus *economic cost*. (8)

economically efficient: minimizing the cost of producing a specified amount of output. (7)

economies of scale: property of a cost function whereby the average cost of production falls as output expands. (7)

economies of scope: the situation in which it is less expensive to produce goods jointly than separately. (7)

efficiency in production: a situation in which the principal's and agent's combined value (profits, payoffs), π, is maximized. (20)

efficiency in risk bearing: a situation in which risk sharing is optimal in that the person who least minds facing risk—the risk-neutral or less risk-averse person—bears more of the risk. (20)

efficiency wage: an unusually high wage that a firm pays workers as an incentive to avoid shirking. (20)

efficient contract: an agreement in which neither party can be made better off without harming the other party. (20)

efficient production (*technological efficiency*): the situation in which the current level of output cannot be produced with fewer inputs, given existing knowledge about technology and the organization of production. (6)

elasticity: the percentage change in a variable in response to a given percentage change in another variable. (3)

elasticity of demand (or *price elasticity of demand*, ε): the percentage change in the *quantity demanded* in response to a given percentage change in the price. (3)

elasticity of supply (or *price elasticity of supply*, η): the percentage change in the *quantity supplied* in response to a given percentage change in the price. (3)

endowment: an initial allocation of goods. (10)

endowment effect: people place a higher value on a good if they own it than they do if they are considering buying it. (4)

Engel curve: the relationship between the quantity demanded of a single good and income, holding prices constant. (5)

equilibrium: a situation in which no one wants to change his or her behavior. (2)

equivalent variation (*EV*): the amount of money one would have to take from a consumer to harm the consumer by as much as the price increase or give to a consumer to benefit the consumer by as much as a price increase. (5)

excess demand: the amount by which the *quantity demanded* exceeds the *quantity supplied* at a specified price. (2)

excess supply: the amount by which the *quantity supplied* is greater than the *quantity demanded* at a specified price. (2)

exclusion: others can be prevented from consuming a good. (18)

exhaustible resources: nonrenewable natural assets that cannot be increased, only depleted. (16)

expansion path: the cost-minimizing combination of labor and capital for each output level. (7)

extensive form (of a game): specifies the n players, the sequence in which they make their moves, the actions they can take at each move, the information that each player has about players' previous moves, and the payoff function over all possible strategies. (14)

externality: the direct effect of the actions of a person or firm on another person's well-being or a firm's production capability rather than an indirect effect through changes in prices. (18)

fair bet: a wager with an expected value of zero. (17)

fair insurance: a bet between an insurer and a policyholder in which the value of the bet to the policyholder is zero. (17)

firm: an organization that converts inputs such as labor, materials, energy, and capital into outputs, the goods and services that it sells. (6)

fixed cost (*F*): a production expense that does not vary with output. (7)

fixed input: a factor of production that cannot be varied practically in the short run. (6)
flow: a quantity or value that is measured per unit of time. (16)
free riding: benefiting from the actions of others without paying. (18)
game: any competition between players (firms) in which strategic behavior plays a major role. (14)
game theory: a set of tools that economists, political scientists, military analysts, and others use to analyze decision making by players who use strategies. (14)
general-equilibrium analysis: the study of how equilibrium is determined in all markets simultaneously. (10)
Giffen good: a commodity for which a decrease in its price causes the quantity demanded to fall. (5)
good: a commodity for which more is preferred to less, at least at some levels of consumption. (4)
group price discrimination (*third-degree price discrimination*): a situation in which a firm charges different groups of customers different prices but charges a given customer the same price for every unit of output sold. (12)
hidden action: an act by one party to a transaction that the other party cannot observe. (19)
hidden characteristic: an attribute of a person or thing that is known to one party but unknown to others. (19)
incidence of a tax on consumers: the share of the tax that falls on consumers. (3)
income effect: the change in the quantity of a good a consumer demands because of a change in income, holding prices constant. (5)
income elasticity of demand (or *income elasticity*): the percentage change in the *quantity demanded* in response to a given percentage change in income. (3)
increasing returns to scale: property of a production function whereby output rises more than in proportion to an equal increase in all inputs. (6)
indifference curve: the set of all bundles of goods that a consumer views as being equally desirable. (4)
indifference map (or *preference map*): a complete set of indifference curves that summarize a consumer's tastes or preferences. (4)
inferior good: a commodity of which less is demanded as income rises. (5)
interest rate: the percentage more that must be repaid to borrow money for a fixed period of time. (16)
internal rate of return (*irr*): the discount rate that results in a net present value of an investment of zero. (16)
internalize the externality: to bear the cost of the harm that one inflicts on others (or to capture the benefit that one provides to others). (18)
isocost line: all the combinations of inputs that require the same (*iso*) total expenditure (*cost*). (7)
isoquant: a curve that shows the efficient combinations of labor and capital that can produce a single (*iso*) level of output (*quantity*). (6)
Law of Demand: consumers demand more of a good the lower its price, holding constant other factors that influence consumption. (2)
learning by doing: the productive skills and knowledge that workers and managers gain from experience. (7)
learning curve: the relationship between average costs and cumulative output. (7)
Lerner Index: the ratio of the difference between price and marginal cost to the price: $(p - MC)/p$. (11)
limit pricing: a price (or, equivalently, an output level) that a firm sets so that another firm cannot enter the market profitably. (14)
limited liability: the condition whereby the personal assets of the owners of the corporation cannot be taken to pay a corporation's debts if it goes into bankruptcy. (6)
long run: a lengthy enough period of time that all inputs can be varied. (6)
marginal cost (*MC*): the amount by which a firm's cost changes if the firm produces one more unit of output. (7)
marginal product of labor (MP_L): the change in total output, Δq, resulting from using an extra unit of labor, ΔL, holding other factors constant: $MP_L = \Delta q/\Delta L$. (6)
marginal profit: the change in profit a firm gets from selling one more unit of output. (8)
marginal rate of substitution (*MRS*): the maximum amount of one good a consumer will sacrifice to obtain one more unit of another good. (4)
marginal rate of technical substitution (*MRTS*): the number of extra units of one input needed to replace one unit of another input that enables a firm to keep the amount of output it produces constant. (6)
marginal rate of transformation (*MRT*): the trade-off the market imposes on the consumer in terms of the amount of one good the consumer must give up to obtain more of the other good. (4)
marginal revenue (*MR*): the change in revenue a firm gets from selling one more unit of output. (8)
marginal revenue product of labor (MRP_L): the extra revenue from hiring one more worker. (15)
marginal utility: the extra utility that a consumer gets from consuming the last unit of a good. (4)
market: an exchange mechanism that allows buyers to trade with sellers. (1)
market failure: inefficient production or consumption, often because a price exceeds marginal cost. (9)
market power: the ability of a firm to charge a price above marginal cost and earn a positive profit. (11)
market structure: the number of firms in the market, the ease with which firms can enter and leave the market, and the ability of firms to differentiate their products from those of their rivals. (8)
microeconomics: the study of how individuals and firms make themselves as well off as possible in a world of scarcity and the consequences of those individual decisions for markets and the entire economy. (1)
minimum efficient scale (*full capacity*): the smallest quantity at which the average cost curve reaches its minimum. (13)

mixed strategy: a firm (player) chooses among possible actions according to probabilities it assigns. (14)
model: a description of the relationship between two or more economic variables. (1)
monopolistic competition: a market structure in which firms have market power but no additional firm can enter and earn positive profits. (13)
monopoly: the only supplier of a good that has no close substitute. (11)
monopsony: the only buyer of a good in a given market. (15)
moral hazard: an informed party takes an action that the other party cannot observe and that harms the less-informed party. (19)
Nash equilibrium: a set of strategies such that, if all other players use these strategies, no player can obtain a higher payoff by choosing a different strategy. (14)
Nash-Bertrand equilibrium (*Bertrand equilibrium* or *Nash-in-prices equilibrium*): a set of prices such that no firm can obtain a higher profit by choosing a different price if the other firms continue to charge these prices. (13)
Nash-Cournot equilibrium (*Cournot equilibrium* or *Nash-in-quantities equilibrium*): a set of quantities sold by firms such that, holding the quantities of all other firms constant, no firm can obtain a higher profit by choosing a different quantity. (13)
natural monopoly: the situation in which one firm can produce the total output of the market at lower cost than several firms could. (11)
network externality: the situation where one person's demand for a good depends on the consumption of the good by others. (11)
nonlinear price discrimination (*second-degree price discrimination*): the situation in which a firm charges a different price for large quantities than for small quantities, so that the price paid varies according to the quantity purchased. (12)
nonuniform pricing: charging consumers different prices for the same product or charging a customer a price that depends on the number of units the customer buys. (12)
normal form (of a game): a representation of a static game with complete information that specifies the players in the game, their possible strategies, and the payoff function that identifies the players' payoffs for each combination of strategies. (14)
normal good: a commodity of which as much or more is demanded as income rises. (5)
normative statement: a conclusion as to whether something is good or bad. (1)
oligopoly: a small group of firms in a market with substantial barriers to entry. (13)
open-access common property: a resource that is nonexclusive and rival. (18)
opportunistic behavior: one party takes economic advantage of another when circumstances permit. (19)
opportunity cost (*economic cost*): the value of the best alternative use of a resource. (7)
opportunity set: all the bundles a consumer can buy, including all the bundles inside the budget constraint and on the budget constraint. (4)
Pareto efficient: describing an allocation of goods or services such that any reallocation harms at least one person. (10)
partial-equilibrium analysis: an examination of equilibrium and changes in equilibrium in one market in isolation. (10)
patent: an exclusive right granted to the inventor to sell a new and useful product, process, substance, or design for a fixed period of time. (11)
payoffs (of a game): players' valuations of the outcome of the game, such as profits for firms or utilities for individuals. (14)
perfect complements: goods that a consumer is interested in consuming only in fixed proportions. (4)
perfect price discrimination (*first-degree price discrimination*): the situation in which a firm sells each unit at the maximum amount any customer is willing to pay for it, so prices differ across customers and a given customer may pay more for some units than for others. (12)
perfect substitutes: goods that a consumer is completely indifferent as to which to consume. (4)
pooling equilibrium: an equilibrium in which dissimilar people are treated (paid) alike or behave alike. (19)
positive statement: a testable hypothesis about cause and effect. (1)
price discrimination: practice in which a firm charges consumers different prices for the same good based on individual characteristics of consumers, membership in an identifiable subgroup of consumers, or on the quantity purchased by the consumers. (12)
price elasticity of demand (or *elasticity of demand*, ε): the percentage change in the *quantity demanded* in response to a given percentage change in the price. (3)
price elasticity of supply (or *elasticity of supply*, η): the percentage change in the *quantity supplied* in response to a given percentage change in the price. (3)
prisoners' dilemma: a game in which all players have dominant strategies that result in profits (or other payoffs) that are inferior to what they could achieve if they used cooperative strategies. (14)
private cost: the cost of production only, not including *externalities*. (18)
producer surplus (*PS*): the difference between the amount for which a good sells and the minimum amount necessary for the seller to be willing to produce the good. (9)
production function: the relationship between the quantities of inputs used and the maximum quantity of output that can be produced, given current knowledge about technology and organization. (6)
production possibility frontier: the maximum amount of outputs that can be produced from a fixed amount of input. (7)

profit (π): the difference between revenues, R, and costs, C: $\pi = R - C$. (6)
property right: the exclusive privilege to use an asset. (18)
public good: a good that is nonrival and nonexclusive. (18)
pure strategy: each player chooses an action with certainty. (14)
quantity demanded: the amount of a good that consumers are willing to buy at a given price, holding constant the other factors that influence purchases. (2)
quantity supplied: the amount of a good that firms *want* to sell at a given price, holding constant other factors that influence firms' supply decisions, such as costs and government actions. (2)
quota: the limit that a government sets on the quantity of a foreign-produced good that may be imported. (2)
rent: a payment to the owner of an input beyond the minimum necessary for the factor to be supplied. (9)
rent seeking: efforts and expenditures to gain a rent or a profit from government actions. (9)
requirement tie-in sale: a tie-in sale in which customers who buy one product from a firm are required to make all their purchases of another product from that firm. (12)
reservation price: the maximum amount a person would be willing to pay for a unit of output. (12)
residual demand curve: the market demand that is not met by other sellers at any given price. (8)
residual supply curve: the quantity that the market supplies that is not consumed by other demanders at any given price. (8)
risk: the situation in which the likelihood of each possible outcome is known or can be estimated and no single possible outcome is certain to occur. (17)
risk averse: unwilling to make a fair bet. (17)
risk neutral: indifferent about making a fair bet. (17)
risk preferring: willing to make a fair bet. (17)
risk premium: the amount that a risk-averse person would pay to avoid taking a risk. (17)
rival good: a good that is used up as it is consumed. (18)
rules of the game: regulations that determine the timing of players' moves and the actions that players can make at each move. (14)
screening: an action taken by an uninformed person to determine the information possessed by informed people. (19)
separating equilibrium: an equilibrium in which one type of people takes actions (such as sending a *signal*) that allows them to be differentiated from other types of people. (19)
shirking: a *moral hazard* in which agents do not provide all the services they are paid to provide. (20)
short run: a period of time so brief that at least one factor of production cannot be varied practically. (6)
shortage: a persistent excess demand. (2)
signaling: an action taken by an informed person to send information to a less-informed person. (19)
snob effect: the situation in which a person places greater value on a good as fewer and fewer other people possess it. (11)
social cost: the private cost plus the cost of the harms from *externalities*. (18)
standard: a metric or scale for evaluating the quality of a particular product. (19)
static game: a game in which each player acts only once and the players act simultaneously (or, at least, each player acts without knowing rivals' actions). (14)
stock: a quantity or value that is measured independently of time. (16)
strategy: a battle plan that specifies the action that a player will make conditional on the information available at each move and for any possible contingency. (14)
subgame: all the subsequent decisions that players may make given the actions already taken and corresponding payoffs. (14)
subgame perfect Nash equilibrium: players' strategies are a Nash equilibrium in every subgame. (14)
substitute: a good or service that may be consumed instead of another good or service. (2)
substitution effect: the change in the quantity of a good that a consumer demands when the good's price changes, holding other prices and the consumer's *utility* constant. (5)
sunk cost: a past expenditure that cannot be recovered. (7)
supply curve: the *quantity supplied* at each possible price, holding constant the other factors that influence firms' supply decisions. (2)
supply function: the relationship between the quantity supplied, price, and other factors that influence the number of units offered for sale. (2)
tariff (*duty*): a tax on only imported goods. (9)
tax salience: awareness of a tax.
technical progress: an advance in knowledge that allows more output to be produced with the same level of inputs. (6)
tie-in sale: a type of nonlinear pricing in which customers can buy one product only if they agree to buy another product as well. (12)
total cost (C): the sum of a firm's variable cost and fixed cost: $C = VC + F$. (7)
transaction costs: the expenses of finding a trading partner and making a trade for a good or service beyond the price paid for that good or service. (2)
two-part pricing: a pricing system in which the firm charges each consumer a lump-sum access fee for the right to buy as many units of the good as the consumer wants at a per-unit price. (12)
uniform pricing: charging the same price for every unit sold of a particular good. (12)
utility: a set of numerical values that reflect the relative rankings of various bundles of goods. (4)
utility function: the relationship between *utility* values and every possible bundle of goods. (4)
variable cost (VC): a production expense that changes with the quantity of output produced. (7)
variable input: a factor of production whose quantity can be changed readily by the firm during the relevant time period. (6)
winner's curse: auction winner's bid exceeds an item's common-value. (14)

Index

Numbers

3D printing, 230
401(k) plans, 126–127

A

Abundance, impact on price of exhaustible resources, 575
AC. *See* Average costs (*AC*)
Access fee, in two-part pricing, 429
Acid Rain Cap-and-Trade Program, 634
Actions, game
  in dynamic games, 504
  overview of, 491
  in static games, 492
Ad valorem subsidy, 134
Ad valorem tariffs, 320
Ad valorem taxes
  equilibrium effects of, 88–90
  types of sales taxes, 82–83
Adverse selection
  adverse selection and remanufactured goods, 659
  asymmetric information in, 650
  avoiding moral hazard by preventing, 700
  equalizing information, 656–657
  in insurance markets, 651–652
  overview of, 651
  with products of unknown quality, 652–654
  reducing, 656
  restricting opportunistic behavior, 660
Advertising
  benefits of advertising the price of goods or services, 664
  deciding how much to advertise, 438–439
  deciding whether to advertise, 437–438
  example (Super Bowl commercials), 439–440
  overview of, 437
  as static game, 497
  strategic advertising by monopolies, 725–726
  strategic advertising by oligopolies, 499
AFC (average fixed cost), 210
Agency problem. *See* Principal-agent problem
Allais (certainty) effect, violations of expected utility theory, 610
Allocation of goods
  efficiency vs. equity, 361–363
  endowments, 340–342
  government policies and Pareto principle, 356
  Pareto-efficient allocations, 332, 356
  social welfare functions in ranking, 356–357
  voting as means of ranking, 360–361
  voting as mechanism for, 357–359
Allocation of scarce resources
  decision making in, 26
  microeconomics as study of, 25–26
  price determining allocations, 26
  trade-offs in, 26
Amortization, of costs, 207
Annual percentage rate (APR), 556–557
Antitrust Division, of Department of Justice (DOJ), 453–455
Antitrust laws, 505
APR (annual percentage rate), 556–557
Arctic National Wildlife Refuge (ANWR), 80–81
Arrow, Kenneth, 358
Arrow's Impossibility Theorem, 358–359
Assets
  inequitable distribution of, 354–355
  nonrenewable. *See* Exhaustible resources
Assumptions, in microeconomic models, 28
Asymmetric information
  adverse selection, 651
  adverse selection in insurance markets, 651–652
  adverse selection with products of unknown quality, 652–653
  advertising and pricing and, 664
  cheap talk problem, 664–666
  consumers and, 660
  defined, 650
  education as signal of ability, 666–670
  equalizing information, 656–659
  hiring practices and, 664
  job-related risks and, 649, 672–673
  market power from price ignorance, 661–662
  overview of, 650–651
  price discrimination due to false beliefs about quality, 660–661
  principal-agent problem, 679–680
  reducing adverse selection, 656
  restricting opportunistic behavior, 660

screening in hiring, 670–672
tourist-trap model, 662
Auctions
bidder's curse, 518
bidding strategies, 516–517
elements of, 515–516
overview of, 514
winner's curse, 517–518
Auditing, to reduce moral hazard, 658
AVC. *See* Average variable cost (*AVC*)
Average costs (*AC*)
comparing competitive market with cartel, 452
deciding whether to advertise, 437–438
determining profit-maximizing output in monopolies, 374–375
impact of specific tax on short-run equilibrium, 266–267
long-run average cost as envelope of short-run average cost curves, 231–234
long-run cost curves and, 227–228
minimum of average cost curve, 717
in natural monopoly, 387
output decisions in short-run competitive firms, 256–257
shape of average cost curve, 214–215, 227
shape of average cost curve (Canadian manufacturing example), 230
short-run cost curves, 212
short-run costs, 209
taxes impacting short-run cost curves, 215–216
Average fixed cost (*AFC*), 210
Average product of labor
effect of adding extra labor, 180–181
graphing product curves, 179–180
overview of, 179
properties of average product curves, 716
relationship of product curves, 181
in total product, 177–178
Average revenue, comparing competitive firm with monopoly, 369
Average variable cost (*AVC*)
determining profit-maximizing output in monopolies, 374–375
short-run costs, 210
shutdown decisions in short-run competitive firms, 259–260
taxes impacting short-run cost curves, 215–216
Avoidability, of long-run costs, 218

B

Backward induction, in subgames perfect Nash equilibrium, 508
Bad, preferences and, 100
Ban on imports
free trade vs., 318–320
import policies and, 318
Bandwagon effect, network externalities and, 398
Bargaining ability, trades and, 344
Barrier to entry
cartels and, 457
government role in creating monopolies, 387–388
policies creating, 306
zero long-run profit with limited entry, 289–290
zoning laws as barrier to entry by hotel chains, 482–483
Barrier to exit, 306–307
Behavioral economics
defined, 98
endowment effects, 125–126
network externalities and, 398
overview of, 125
salience and bounded rationality, 127
tests of transitivity, 125
uncertainty and, 608
Behavioral game theory
overview of, 519–520
Bentham, Jeremy, 108, 360
Bertrand equilibrium. *See* Nash-Bertrand equilibrium
Bertrand, Joseph, 474
Bertrand oligopoly
with differentiated products, 475–477
with identical products, 474–475
overview of, 474
Best response
backward induction, 508
Bertrand, 474–477
Cournot algebra, 463
Cournot best-response curve, 461–462
Cournot unequal costs, 465–466
game theory, 494–496, 500–504
Stackelberg follower, 469–472
strategies in static games, 495
to a rival's action, 461
Biases
present bias causing discount rates to fall, 569–570
in risk assessment, 608
substitution bias in CPI, 157–158
Bidder's curse, auctions, 518
Bidding
formats in auctions, 515
strategy in private-value auctions, 516–517
Block pricing
calculating, 724–725
as nonlinear pricing strategy, 427–428
Bonding
efficiency wages compared with, 696
overview of, 692–693
preventing shirking, 693
problems with, 694
trade-offs between bonding and monitoring, 693–694
Bonds (financial), rate of return on, 567–568

Bonuses, contracts for avoiding moral hazard, 689–690
Boomerang generation, 503
Borda, Jean-Charles de, 359
Borlaug, Norman, 183
Borrowing, discount rate and, 555
Bounded rationality, salience and, 127
Boutique fuels, 277–278
Budget constraints, of consumers
  defined, 98
  example (water quotas), 114–115
  income change impacting opportunity set, 116
  indifference curves and, 136–137
  maximizing well-being subject to, 116–117, 712–714
  overview of, 112–113
  price change impacting opportunity set, 115–116
  slope of budget line, 114
Budget line. *See* Budget constraints, of consumers
Budgets, effects of rise in income on demand, 140
Bundle of goods
  buying happiness (more is better), 122
  competitive exchange and, 345–346
  contract curves showing set of Pareto-efficient bundles, 344
  equality of, 360–361
  maximizing well-being subject to budget constraints, 116–117
  opportunity set, 113
  optimal bundle, 117–118
  ordinal preferences, 108–109
  preference maps and, 101
  preferences and, 98
  utility and, 108
  utility function and, 108
Bundling (package tie-in sales)
  example (word processing/spreadsheet bundle), 436
  overview of, 434–436
  types of nonuniform pricing, 408
Buyers, large number in competitive markets, 247
Buying discounts, group price discrimination and, 425–426

C

California Reformulated Fuels Program (CaRFG), 277–278
Cap-and-trade markets, for pollution, 634
Capital
  as input, 177
  inputs in short-run production, 177
  inputs requiring same total cost (isocost line), 219–221
  isoquants for combining inputs for single level of output, 187
  opportunity cost of, 206–207
  technical progress increasing output of, 196–197
Capital markets
  overview of, 576–577
Cardinal preferences, 108–109
CaRFG (California Reformulated Fuels Program), 277–278
Cartels
  court action against, 456
  defined, 449
  laws preventing, 453
  maintaining, 455–456
  mergers and, 457–458
  overview of, 449, 451
  trusts, 453–455
  why they fail, 453
  why they form, 451–452
Cash, evaluating food stamp program vs. receiving cash, 122–124
CCDF (Child Care and Development Fund), 134
Certainty (Allais) effect, violations of expected utility theory, 610–611
Certifications, actions for equalizing information, 658–659
Chance nodes, in risk-neutral investing, 606
Cheap talk, 664–666
Cheating, incentive to cheat in cartels, 455–456
Chicago Mercantile Exchange, as example of perfect competition, 248
Child Care and Development Fund (CCDF), 134
Clean Air Act
  Acid Rain Cap-and-Trade Program, 634
  benefits vs. costs in pollution control, 628
  California Reformulated Fuels Program (CaRFG), 277–278
  in regulation of externalities, 624–625
Club goods, 637
Coase Theorem, 632–633
Cobb-Douglas production function
  determining shape of, 716–717
  marginal rate of technical substitution (*MRTS*), 190–192
  for paper firm, 532
  returns to scale by industry, 192–194
  tangency rule and, 222
Cobb-Douglas utility function, n. 146
COLAs. *See* Cost-of-living adjustments (COLAs)
Collusion
  comparing types of equilibria, 472–473
  mergers and, 457–458
  tacit, 454
Commissions, contracts for avoiding moral hazard, 690–691
Common knowledge, assumptions in games, 492
Common value auctions
  overview of, 516
  winner's curse where bid exceeds common value, 517–518
Comparative advantage
  benefits of trade, 349

defined, 348
marginal rate of transformation and, 348–349
number of producers and, 350–351
production possibility frontier and, 347–348
Compensating variation (*CV*), 151–152
Competition
efficiency of, 351–353
government role in encouraging, 395–396
group price discrimination compared with, 426
inefficiency of competition with externalities, 621–624
perfect. *See* Perfect competition
Competition Act (Canada), 454–455
Competitive equilibrium
comparing types of equilibria, 472–473
competitive exchange and, 345–346
competitive factor market equilibrium, 537–538
equitability of, 361
example (cost of trucking), 245, 280–281
example (vegetable oil), 279
long-run competition, 279–280
perfect price discrimination contrasted with, 414–416
short-run competition, 265–266
specific tax impacting, 266–267
Competitive exchange
competitive equilibrium and, 345–346
efficiency of competition and, 346–347
obtaining any efficient allocation using competition, 347
overview of, 344–345
Competitive factor market
competitive factor and output markets, 540–541
competitive factor market equilibrium, 537–538
factor market demand, 535–537
long-run factor demand, 534–535
overview of, 529
short-run factor demand, 529–532
Competitive market model
applying producer surplus, 299–300
consumer welfare and, 291–292
deadweight loss and market failure, 302–304
deadweight loss related to Christmas presents, 304–305
decrease in welfare due to overproduction, 302–304
effect of price change on consumer surplus, 295–297
entry barrier, 306–307
firms must maximize profit to survive, 291
free trade vs. ban on imports, 318–320
free trade vs. quotas, 323
free trade vs. tariffs, 320–323
liquor licenses, 324–326
maximization of welfare with, 300–302
measuring consumer surplus with demand curve, 292–294
measuring producer surplus with supply curve, 298–299
overview of, 287–288
policies creating barriers to entry or exit, 306–307
policies regulating imports, 318
policies that create wedge between supply and demand, 307
policies that shift supply and demand curves, 305–306
rent seeking and, 323–324
welfare effects of price ceilings, 316–318
welfare effects of price floors, 312–314
welfare effects of sales taxes, 307–310
welfare effects of subsidies, 310–312
zero long-run profit with free entry, 288–289
zero long-run profit with limited entry, 289–290
Competitive markets
cartel contrasted with, 451–452
deriving demand curve for competitive firms, 249–250
feedback between, 334–336
long-run. *See* Long-run competition
perfect. *See* Perfect competition
short-run. *See* Short-run competition
supply-and-demand model and, 33
welfare effects of pollution in, 732–734
Complement goods
cross-price elasticity of demand and, 77
overview of, 35
Complete information
in games, 492
in normal-form representation of static games, 493
Completeness property, consumer preferences, 99–100
Compounding
earning interest on interest, 555–556
example of power of, 557–558
frequency of, 556–557
stream of payments and, 558–559
Constant elasticity demand curves, 73
Constant elasticity supply curves, 79
Constant returns to scale (CRS), 191–192
Constrained consumer choice
buying happiness (more is better), 122
defined, 98
example (substituting alcohol for marijuana), 119
food stamps programs and, 122–124
optimal bundle and, 117–118

Constrained consumer choice (*continued*)
optimal bundle corner solution, 119–120
optimal bundle interior solution, 118
optimal bundle on convex sections of indifference curves, 121–122
overview of, 116–117
Consumer choice
behavioral economics and, 125
budget constraints, 112–113
buying happiness (more is better), 122
constrained, 116–117
curvature of indifference curves, 105–107
endowment effects, 125–127
example (ebooks), 97, 128–129
example (substituting alcohol for marijuana), 119
example (water quotas), 114–115
food stamps programs and, 122–124
impact of income change on opportunity set, 116
impact of price change on opportunity set, 115–116
indifference curves, 101–104
indifference curves (food/clothing example), 107
marginal rate of substitution (*MRS*) and, 111–112
marginal utility and, 110–111
optimal bundle corner solution, 119–120
optimal bundle of consumer, 117–118
optimal bundle on convex sections of indifference curves, 121–122
ordinal preferences, 108–109
overview of, 97–98
preference maps, 101–103
preference properties, 99–101
preferences, 99
salience and bounded rationality, 127
slope of budget line, 114
tests of transitivity, 125
utility, 108
utility and indifference curves, 109–110
utility function, 108
willingness to substitute, 104–105
Consumer Price Index (CPI)
adjusting for inflation, 154–157
calculating inflation indexes, 153–154
comparing prices over time, 152
overcompensating for inflation, 156–157
size of substitution bias in, 157–158
Consumer surplus (*CS*)
barriers to entry or exit lowering, 306
block pricing and, 427–428
competition maximizing welfare, 300–302
demand elasticity and, 721–722
effect of price change on, 295–297
effects of pollution, 622
effects of tariff or quotas, 321–323
entry barrier, 306–307
group price discrimination and, 426
loss from eliminating free trade, 318–320
market consumer surplus, 294
in measuring consumer welfare, 292–294
monopsonies and, 546–547
nonoptimal price regulation, 392
optimal price regulation, 390–392
perfect price discrimination and, 414–416
reducing output below competitive level, 301–304
rent seeking and, 323–324
welfare effects of price ceilings, 316–317
welfare effects of price floors, 312–314
welfare effects of sales taxes, 307–309
welfare effects of subsidies, 310–312
willingness to pay (eBay example), 294–295
Consumer theory
child care subsidies and, 134
compensating variation and equivalent variation and, 151–152
cost-of-living adjustments (COLAs), 152, 156–157
CPI adjustment for inflation, 154–157
deriving demand curves, 135–136
effects of price changes, 146
effects of rise in income on demand, 140–142
income and substitution effects applied to labor-leisure choice, 160–162
income and substitution effects with inferior goods, 150–151
income and substitution effects with normal goods, 146–150
income-consumption curves and income elasticities, 143–144
income elasticity of demand and, 142–143
income tax rates and labor supply curve, 164–166
indifference curves and rotating budget line, 136–137
inflation indexes and, 152–154
labor-leisure choice, 158–160
labor supply curves and, 158
normal goods and inferior goods and, 144–146
overview of, 134–135
price-consumption curves, 137
response of demand curve to price-consumption curve, 138–139
shape of labor supply curves, 162–164
tobacco use example, 138

Consumer welfare
effect of price change on consumer surplus, 295–296
measuring consumer welfare with demand curve, 292–295
overview of, 291–292
Consumers
asymmetric information and, 660
budget constraints. *See* Budget constraints, of consumers
role of taste in demand, 34
tax incidence and, 85–86
who pays the gasoline tax, 67–68
Consumer's optimum, 118
Contingent contracts
bonuses and options in, 689–690
commissions in, 690
example (health insurance and moral hazard), 685
piece-rate contracts, 690
profit-sharing contracts, 687
for reducing moral hazard, 685
state-contingent contracts, 685
when to choose, 700–701
Contingent fees, in contracts, 691
Contract curves, showing set of Pareto-efficient bundles, 343
Contracts, efficient contracts, 680
Contracts, for avoiding moral hazard
bonuses and options, 689–690
checks on principals in, 697–700
choosing best contract, 691, 700–701
commissions, 690
contingent contracts, 685
example (sing for your supper), 688–689
fixed-fee contracts, 684–685
overview of, 683–684
piece-rate contracts, 690
profit-sharing contracts, 687
state-contingent contracts, 685
Convention on International Civil Aviation, 456
Cooperation
in dynamic games, 505
example (airlines), 506
in static games, 505–506
Copyrights, reselling textbooks and, 422
Corporations
limited liability in, 174
ownership structures of for-profit firms, 174
Correlation, diversification and, 598–599
Cost-based monopolies, 385–387
Cost of capital, durable goods and, 206–207
Cost-of-living adjustments (COLAs)
adjusting for inflation overcompensation, 156–157
inflation indexes, 152–154
overview of, 152
Costs
average, 210
avoidable costs in long-run, 218
combining cost and production information, 221
cost advantage as cause of monopolies, 385–386
of durable goods, 206–207
estimating cost curves vs. introspection, 230–231
factor prices changes and, 224–225
fixed, variable, and total, 208
impacting supply, 42
isocost line, 219–221
last-dollar rule, 223–224
learning curve and, 235–236
long-run, 218
long-run average cost as envelope of short-run average cost curves, 231–234
long-run cost curves, 227–230
long-run expansion path and long-run cost function, 225–227
lower costs in long run, 231
lowest-isocost rule, 222
marginal, 209–210
minimizing (beer example), 221
minimizing in long-run, 218–219, 718–720
nature of, 204
nonidentical firms with unequal costs in Cournot oligopoly, 465–467
opportunity costs, 204–206
overview of, 203–204
private and social, 622
of producing multiple goods, 236–238
production functions and shape of cost curves, 212–213
profit and, 175
shape of average cost curve, 214–215
shape of marginal cost curve, 214
shape of variable cost curve, 212–213
short- and long-run expansion paths, 234–235
short-run, 208
short-run cost curves, 210–212
short-run cost curves (beer manufacturer example), 215
sunk costs, 207–208
tangency rule, 222–223
taxes impacting, 215–216
technology choices at home vs. abroad, 238–239
Coupons, group price discrimination and, 425–426
Cournot oligopoly
aircraft subsidies and, 448, 483–484
comparing types of equilibria, 472–473
equilibrium, elasticity and number of firms and, 463–465
example (airlines industry), 459–460
example of spurious product differentiation (bottled water), 477
nonidentical firms with differentiated products, 467–469
nonidentical firms with unequal costs, 465–467
CPI. *See* Consumer Price Index (CPI)

Credible threat, in sequential dynamic games, 509–510
Critical mass, as network externality, 398–399
Cross-price elasticity of demand, 77
CRS (Constant returns to scale), 191–194
Cumulative output, learning curve and, 235
Customers, sharing burden of sales tax with firm, 85–88
CV (compensating variation), 151–152

D

Deadweight loss (*DWL*)
  block pricing and, 427–428
  effects of pollution, 623
  effects of tariff or quotas, 320–323
  example (Christmas presents), 304–305
  example (gasoline tax), 309
  group price discrimination and, 426
  of monopolies, 382–383
  of monopsonies, 546–547
  nonoptimal price regulation, 393–394
  nonuniform pricing and, 407
  perfect price discrimination and, 414–416
  price caps and, 391
  reduction in welfare due to, 301–302
  rent seeking and, 323–324
  welfare effects of price ceilings, 316
  welfare effects of price floors, 312–314
  welfare effects of sales taxes, 307–309
  welfare effects of subsidies, 310–312
Decision making
  management of firms and, 175
  in resource allocation, 26
Decision nodes, in risk-neutral investing, 606
Decision trees
  in risk-averse investing, 607
  in risk-neutral investing, 606
Decreasing returns to scale (*DRS*), 191–192
Deductibles, insurance and, 605
Deferred payments, reducing moral hazard, 695
Demand
  cross-price elasticity of, 77
  demand curves, 35–36
  demand function, 38–40
  effect of factors other than price on demand, 37
  effects of rise in income on, 168–169
  excess demand, 49
  factor market demand, 535–537
  income elasticity of, 76–77
  for leisure, 159
  long-run factor demand in competitive factor market, 534–535
  market structure and factor demand, 538–539
  overview of, 34–35
  policies creating wedge between supply and demand, 307
  policies that cause demand to differ from supply, 55
  price elasticity of. *See* Price elasticity of demand
  sensitivity of quantity demanded to price, 69
  short-run factor demand in competitive market, 529–532
  summing demand curves, 41
  when supply need not equal to demand, 59–60
  when to use supply-and-demand model, 60–61
Demand curves
  choosing price vs. quantity in monopolies, 373
  deriving, 135–136
  deriving for competitive firm, 249–250
  determining market equilibrium, 47–48
  effect of factors other than price on demand, 37
  effects of price changes on, 146
  effects of rise in income on, 140–142
  elasticity along downward-sloping linear, 71–73
  elasticity along horizontal, 73
  elasticity along vertical, 73
  example (tobacco use), 138
  of firms that can price discriminate, 410–411
  income-consumption curves and income elasticities, 143–144
  income elasticity of demand and, 142–143
  indifference curves and rotating budget line, 136–137
  Law of Demand and, 36
  long-run competitive equilibrium and, 279–280
  market power and shape of, 379–380
  measuring consumer welfare with, 292–294
  movement along. *See* Movement along demand curve
  normal goods and inferior goods and, 144–146
  overview of, 35–36
  perfectly competitive markets and, 247
  price-consumption curves, 137
  relationship of linear demand curve to marginal revenue curve, 722
  residual demand curves, 720
  response to price-consumption curve, 138–139
  shift in. *See* Shift in demand curve
  significance of shape of, 68–69
  summing, 41
Demand function, 38–41
Democracy, voting as mechanism for allocation, 357–359
Department of Justice (DOJ), Antitrust Division, 453–455

Differentiated products
  in Bertrand oligopoly, 475–477
  example (food trucks), 478–479
  example (Intel/AMD), 468–469
  example of spurious product differentiation (bottled water), 477
  nonidentical firms in Cournot oligopoly, 467–468
  spurious, 463–465
  welfare and, 477
Diminishing marginal returns
  to labor, 212
  law of, 181–183
Direct/indirect effects, of network externalities, 398
Discount rates
  present bias causing rates to fall, 570–571
  relative value of present vs. future consumption, 555
  social discount rate and environment, 569
Discounting
  applying to exhaustible resources, 570
  time-varying discounting, 568–569
Discounts, group price discrimination and, 426
Diseconomies of scale, 229
Diseconomies of scope, 238
Diversification
  correlation and, 598–599
  example (mutual funds), 599–600
  example (failure to diversify), 600–601
  insurance for diversifiable risks, 604–605
  for risk reduction, 599–600
DOJ (Department of Justice), Antitrust Division, 453–455
Dominant strategies, in static games, 494–495
Downturns
  layoffs vs. pay cuts, 698–700
  productivity of labor during, 172, 198–199
DRS (decreasing returns to scale), 191–192
Duopolies
  comparing duopoly equilibria, 472–473
  Nash-Cournot Equilibrium, 459–460
  overview of, 458–459
Durable goods. *See also* Goods, 206–207
Dutch auctions
  overview of, 515
  strategies in, 517
Duties. *See* Tariffs (duty)
Dynamic entry games
  example (keeping out casinos), 513
  exclusion contracts, 510–512
  limit pricing, 513–514
  overview of, 510
Dynamic games
  actions in, 504
  defined, 504
  repeated games. *See* Repeated games
  sequential. *See* Sequential games

E

Economic cost. *See* Opportunity cost
Economic profit. *See* Profit
Economically efficient
  long-run costs and, 218–219
  output decisions, 203
Economies of scale
  learning by doing and, 235–236
  overview of, 229
Edgeworth boxes
  competitive exchange and, 345–346
  showing endowments with, 341–342
Edgeworth, Francis Ysidro, 341–342
Education, as signal of ability
  efficiency of, 669–670
  multiple equilibria, 668–669
  pooling equilibrium, 667–668
  separating equilibrium, 667, 669
Efficiency
  of contracts, 680
  overview of, 353
  Pareto-efficient allocations, 356
  of perfect price discrimination, 414–416
  of product mixes, 351
  of production, 175
  role of government, 353
  social welfare functions for equity, 356–357, 360–361
  voting as allocation method, 357–359
  vs. equity, 361
Efficiency in production, 680
Efficiency in risk bearing, 680
Efficiency wages, reducing moral hazard, 695–697
Efficient contracts, 680
Effluent charge, regulating externalities, 626
Elasticity
  Cournot oligopoly and, 463–465
  defined, 69
  of demand. *See* Price elasticity of demand
  of supply. *See* Price elasticity of supply
Employees
  knowledge of injury rates in industry, 649–650
  monitoring for avoiding moral hazard, 691–692
  paying to relocate, 156
  screening out bad workers, 700–701
Employment
  efficiency wages and unemployment, 696–697
  hiring practices. *See* Hiring practices
  job-related risks and, 649–650, 672–673
  layoffs vs. pay cuts, 698–700
Endowments
  as basis of trades between two people, 340–342
  competitive exchange and, 344–345
  endowment effects, 125–126

Endowments (*continued*)
- mutually beneficial trades and, 342–344
- redistribution of wealth, 353

Engel curves
- fast-food example, 145–146
- relationship of income to quantity demanded, 141–142

English auctions
- overview of, 515
- strategies in, 517

Entry of firms, into market
- barriers to. *See* Barrier to entry
- dynamic games. *See* Dynamic entry games
- free in competitive markets, 248
- in long-run competition, 269
- market supply when entry is limited, 272
- market supply with identical firms and free entry, 271–272
- policies creating barriers to entry or exit, 306–307
- zero long-run profit with free entry, 288–289
- zero long-run profit with limited entry, 289–290

Environment
- pollution and. *See* Pollution
- social discount rate and, 569–570

Equilibrium
- capital market equilibrium, 576–577
- comparing competitive, Stackelberg, Cournot, and collusive equilibria, 472–473
- competitive. *See* Competitive equilibrium
- Cournot oligopoly and, 463–465
- defined, 47
- effects of ad valorem taxes, 88–90
- effects of government interventions, 52
- effects of specific tax, 83–84
- effects of specific tax on, 709–710
- general. *See* General equilibrium
- market. *See* Market equilibrium
- monopolistic competition and, 479–480
- Nash-Bertrand equilibrium. *See* Nash-Bertrand equilibrium
- Nash-Cournot equilibrium. *See* Nash-Cournot equilibrium
- Nash equilibrium. *See* Nash equilibrium
- Nash-Stackelberg equilibrium. *See* Nash-Stackelberg equilibrium
- in oligopoly, 458–459

Equilibrium price, in market equilibrium, 47

Equilibrium quantity, in market equilibrium, 47

Equity
- efficiency vs., 361
- overview of, 353
- Pareto-efficient allocations, 356
- role of government, 353
- social welfare functions for, 356–357, 360–361
- voting as allocation method, 357–359

Equivalent variation (*EV*), 151–152

Essential facility (scarce resource), causes of monopolies, 386

Estimates, using regression
- confidence in, 708–709
- estimating economic relations, 707–708

E*U* (expected utility). *See* Expected utility (E*U*)

European Union antitrust laws, 454–455

EV (equivalent variation), 151–152

Excess demand, market equilibrium and, 49

Excess supply. *See also* Residual supply curves
- international trade and, n. 275
- market equilibrium and, 49

Exclusion. *See also* Rivalry/exclusion, 635

Exclusion contracts, in dynamic entry games, 510–512

Executives, monitoring for avoiding moral hazard, 691–684

Exhaustible resources
- constant or falling prices, 574–576
- overview of, 570
- pricing scarce exhaustible resources, 571–574
- real prices (redwood trees example), 574
- when to sell, 570–571

Exit of firms, from market
- entry and exit of solar power firms, 270–271
- free in competitive markets, 248
- in long-run competition, 269–270
- policies creating barriers to entry or exit, 306–307

Expansion paths
- long-run cost curve and, 226
- minimizing cost of labor and capital, 225
- short- and long-run costs and, 234–235

Expected utility (E*U*)
- comparing prospect theory with, 611–612
- including risk in utility maximizing function, 591–592
- of an investment, 607
- risk aversion and, 592–594
- risk neutrality and, 595–596
- violations of expected utility theory, 609–611

Expected value
- in risk assessment, 588–589
- variance and standard deviation from, 590–591

Extensive-form diagram. *See* Game trees

Externalities
- allocating property rights to reduce, 631–632
- benefits vs. costs in pollution control, 628
- Coase Theorem and, 632–633
- emissions fees, 626–628

emissions standards, 626
inefficiency of competition with, 621–624
market structure and, 629
markets for pollution, 634
markets for positive externalities, 635
monopolies and, 629–630
monopoly vs. competitive welfare with, 630–631
overview of, 619–621
regulating, 624–626
taxing in noncompetitive markets, 631
trade and pollution and, 643–644

F

F. *See* Fixed costs (*F*)
Face value, of bond, 567–568
Factor markets
competitive, 529
competitive factor and output markets, 539–540
competitive factor market and monopolized output markets, 540–541
competitive factor market equilibrium, 537–538
effects of monopolies on, 538
example (athletes' salaries and ticket prices), 528–529, 548–549
long-run factor demand in competitive firms, 534–535
market demand, 535–537
market power in input and output markets, 539
market structure and factor demands, 538–539
monopolized factor market and competitive output markets, 541–542
monopoly in both input and output markets, 542–544
overview of, 528
short-run factor demand in competitive firms, 529–533
Factors of production (inputs)
all inputs that require same cost (isocost line), 219
competitive factor and output markets, 539–540
competitive factor market and monopolized output markets, 540–541
converting into outputs, 173, 176
demand for labor, 536, 730–731
fixed and variable, 177
isoquants for combining inputs for single level of output, 187
labor demand (paper firm example), 532
marginal revenue product of, 530
market power in input markets, 539
monopolized factor market and competitive output markets, 541–542
monopolized market for, 542–544
monopoly in both input and output markets, 542–544
price changes and, 224–225
production function and, 176–177
profit maximization for labor, 532–533
rents on scarce resources, 573
returns to scale and, 191
substitutability and marginal products of, 189–190
substitutability of, 186
substitutability varies along isoquant, 189
substituting, 188–190
taxes on impacting output from, 534
variability over time, 183–184
wage or price change impacting demand, 533–534
Fair bet, expected utility and, 591–592
Fair insurance, 601–603
FDA (Food and Drug Administration), 38
Federal Motor Carrier Safety Administration (FMCSA), 245
Federal Trade Commission Act of 1914, 453
Federal Trade Commission (FTC)
antitrust regulation, 453–454
policy on advertising prices, 664
Financial crisis 2007–2009, excessive risk taking leading to, 677–678
Firm supply
in increasing-cost markets, 274
long-run supply curve, 268
short-run supply curve, 262
Firms
competitive firms as price takers, 246–247
deriving demand curve for competitive firm, 249–250
entry and exit in long-run competition, 269–270
entry barrier, 306–307
management of, 175
market supply when entry is free, 271–272
market supply when entry is limited, 272
market supply when firms differ, 272
maximizing profit to survive, 291
overview of, 173
ownership of for-profit firms, 174
policies creating barriers to entry or exit, 306–307
private, public, and nonprofit, 173–174
sharing burden of sales tax with customers, 85–88
uses of models, 31
what owners want from, 175
which firms can price discriminate, 410–411
who pays the gasoline tax, 67–68
First degree price discrimination. *See* Perfect price discrimination

First-sale doctrine, 422
First Theorem of Welfare
Economics, 345, 346
Fixed costs (*F*)
avoidable in long-run, 218
of entry, 306
monopolistic competition and, 481–483
overview of, 208
Fixed-fee contracts
overview of, 681–682
for reducing moral hazard, 684–685
when to choose, 700–701
Fixed inputs, production, 177
Fixed-proportions production function
long-run costs and, 225
production function, 186
Flows, of goods, 554
FMCSA (Federal Motor Carrier Safety Administration), 245
Food and Drug Administration (FDA), 38
Food stamp programs
evaluating vs. cash, 122–124
example of present bias, 569–570
redistribution of wealth via, 353
For-profit (Private) firms. *See* Private firms
Ford, Henry, 197
Framing effect, violations of expected utility theory, 609–610
Franchise taxes, impacting short-run cost curves, 216
Free riding
public goods and, 638–639
reducing, 641
Free trade
import policies and, 318
loss from eliminating, 318
Pareto efficiency of, 340
vs. ban on imports, 318–320
vs. quotas, 323
vs. tariffs, 320–323
Frequency of events, probability of risk and, 587
FTC (Federal Trade Commission). *See* Federal Trade Commission (FTC)
Future value (*FV*)
interest rates connecting present and future value, 557–558
of stream of payments, 561

G

Gambler's fallacy, 608
Gambling, risk attitudes and, 596–597
Game theory
applying to strategic advertising by oligopolies, 499
auctions and, 514
behavioral game theory, 519–520
best response and Nash Equilibrium, 495–497
bidder's curse in auctions, 518
bidding strategies in auctions, 516–517
cooperation (airlines example), 506
cooperation in repeated prisoner's dilemma game, 505–506
credible threat and, 509–510
defined, 491
dominant strategies, 494–495
dynamic entry games, 510–512
dynamic entry games (keeping out casinos example), 513
dynamic games, 504
elements of auctions, 515–516
example (Intel, AMDs advertising strategies), 490, 520–521
in explanation of parent/child interaction, 503
game trees and, 506–507
limit pricing and, 513
mixed strategies, 501–502
multiple Nash equilibria and, 500
normal-form representation of static games, 493–494
overview of, 490–492
predicting game outcome, 493–494
sequential dynamic games, 506
static games, 492–493
strategies in dynamic games, 504
subgames perfect Nash equilibrium, 508–509
winner's curse in auctions, 517–518
Game trees
for dynamic entry game, 510
sequential dynamic games, 506–507
Stackelberg game tree, 507
Games, defined, 491
GDP (gross domestic product), private sector and, 173
General equilibrium
bargaining ability and, 344
competitive exchange and, 344–347
endowment as basis of trades, 340–342
example (urban flight), 340
feedback between competitive markets (corn/soybean example), 334–336
feedback between competitive markets (tea/coffee example), 336–337
minimum wages with incomplete coverage, 337–339
mutually beneficial trades, 342–344
overview of, 333–334
in trade between two people, 340
General-equilibrium analysis
applied to feedback between competitive markets, 334–336
overview of, 333
Giffen goods, income and substitution effects and, 150–151
Giffen, Robert, n. 150
Goods
allocation of. *See* Allocation of goods
budget constraints and, 112–113
bundle of. *See* Bundle of goods
club goods, 637
endowment effects, 125–127

inferior. *See* Inferior goods
marginal utility of, 110–111
nominal vs. real prices, 152–153
normal. *See* Normal goods
open-access common property as rival good, 635–636
Pareto-efficient allocations, 332, 356
preferences and, 100
price determining allocations, 26
producing multiple goods, 236–238
public goods, 638–639
quality goods. *See* Quality goods
rival goods, 635
stocks and flows, 554
substitutes and complements, 34–35
trade-offs in resource allocation, 26

Government
role in efficiency and equity, 353–356
support for cartels, 456
uses of models, 31

Government policies and regulation
aircraft subsidies, 448, 483–483
allocating property rights to reduce externalities, 631–635
benefits vs. costs in pollution control, 628
causing demand to differ from supply, 55
causing shift in supply curves, 43–44, 52–55
emissions fees, 626–627
emissions standards, 626
encouraging competition, 395
equilibrium effects of government interventions, 52
of externalities, 624–626
impact of import policy on supply curve, 45–47
impact on demand, 35
impact on supply, 42
monopolies created by, 387–390
natural gas regulation, 394–395
nonoptimal price regulation, 392–394
price ceilings, 390–392
problems with regulation, 392
resale prevention, 411

Graphical models
applying Nash-Cournot Equilibrium to airlines industry, 460–462
determining market equilibrium, 47–48
to profit maximization in monopoly, 373–375
of Stackelberg oligopoly (American Airlines example), 471–472

Green Revolution, diminishing marginal returns and, 182–183
Gross domestic product (GDP), private sector and, 173
Group price discrimination
calculating, 724
example (buying discounts), 425–426
example (*Harry Potter* DVD), 419–420
example (textbook resale), 422
identifying groups, 424–425
overview of, 418–419
with two groups, 418–419
types of price discrimination, 412
welfare effects of price discrimination, 426

H

Hatch-Waxman Act, n. 367
Hidden actions
moral hazard of, 678
types of asymmetric information, 650–651
Hidden characteristics, types of asymmetric information, 650
Hiring practices
cheap talk problem, 664–666
education as signal of ability, 666–668
overview of, 664
screening in hiring, 670–672
Human capital
defined, 553
example (decision to go to college), 578–580
Hypotheses
positive vs. normative statements and, 29–31
testing theories, 28–29

I

IATA (International Air Transport Association), 456
Identical products
Bertrand oligopoly with, 474–475
in competitive markets, 247
Imports
free trade vs. ban on, 318–320
free trade vs. quotas, 323
free trade vs. tariffs, 320–323
impact of import policies on supply curves, 45–47
policies regulating, 318
Incidence of taxes
on consumers, 85–86
dependence on elasticities, 86
impact of specific tax on monopolies, 722–723
impact of specific tax on short-run equilibrium, 266–267
Income
budget constraints. *See* Budget constraints, of consumers
effects of rise in income on demand, 140–142, 168–169
example (apple pie), 148–150
impact on demand, 35
impact on opportunity set, 116
inequitable distribution of wealth, 354–355
with inferior goods, 150–151
labor-leisure choice and, 158–160
with normal goods, 146–150
overview of, 146

Income-consumption curves
income elasticity of demand and, 143–144
as income rises, consumption rises (beer/wine example), 141
Income elasticity of demand
consumer theory and, 142–143
income-consumption curves and, 143–144
Slutsky equation applied to, 714–715
in supply-and-demand model, 76
Income elasticity of leisure, 162–163
Income taxes
history of, n. 164
labor supply curves and, 164–166
progressive system of, n. 165
Income threshold model, Chinese example, 27–28
Increasing returns to scale (IRS)
learning curve and, 235
overview of, 191–192
Indifference curves
applying to Coke/Pepsi choice, 141–142
budget constraints and, 136–137
bundle of goods and, 101–104
curvature of, 105–107
example (food/clothing), 107
example of impossible curve, 103
isoquants compared with, 187
marginal rate of substitution (*MRS*) and, 104–107
mutually beneficial trades and, 342–344
optimal bundle of consumer and, 117–119
optimal bundle on convex sections of indifference curves, 121–122
thickness of, 104
utility and, 109–110, 710–712
Indifference maps. *See* Preference maps
Industry, returns to scale by, 192–194
Inferior goods
fast-food example, 145–146
income and substitution effects and, 150–151
income elasticities of, 143
not possible for all goods to be inferior, 144
Inflation
adjusting CPI for, 154–158
calculating inflation indexes, 153–154
interest rates and, 562–563
lottery pay-out options example, 564
nominal vs. real prices and, 152–153
rate of return on bonds and, 567–568
Information
asymmetric. *See* Asymmetric information
in competitive markets, 247
equalizing, 656
in games, 492
impact on demand, 34
impact on demand (Starbucks example), 38
market equilibrium with symmetric information, 653–654
obtaining as risk reduction method, 598
screening, 656–657
signaling, 657–658
symmetric information in efficient contracts, 680–681
third-party comparisons, 658–659
Innovations
productivity, 194–195
robots and the food you eat, 197
technical progress and, 196–197
Inputs. *See* Factors of production (inputs)
Instant runoff voting, 359
Insurance
adverse selection in insurance markets, 651–652
determining amount to buy, 602–603
for diversifiable risks, 604
example (flight insurance), 603–604
fairness and, 603
natural disasters and, 604–605
overview of, 601
universal coverage controlling opportunistic behavior, 660
Insurance companies, monitoring outcomes as means of avoiding moral hazard, 697
Insurance premiums, 601–602
Interdependence, mutually beneficial trades and, 342
Interest rates
capital markets and investments and, 576–577
compounding, 555–556
connecting present and future value of money, 557–558
discount rate, 555
example (car purchase vs. lease), 560–561
example (retirement savings), 561–562
frequency of compounding, 556–557
inflation and, 562–563
overview of, 554–556
pricing scarce exhaustible resources and, 571–572
real and nominal, 562–563
on stream of payments, 558–561
Internal rate of return (*irr*)
approach to investments, 566–567
example (Golden State Warriors), 567
Internalize the externality, 627
International Air Transport Association (IATA), 456
Internet, compensating variation and equivalent variation for giving up, 152
example (what's your smart phone worth to you?), 152

Investments
capital markets and interest rates and, 576–577
diversification of, 599–600
example (deciding to go to college), 578–580
government needs crowding out private investments, 577–578
internal rate of return, 566–567
net present value approach to, 565–566
overview of, 564–565
rate of return on bonds, 567–568
risk-averse, 607
risk-neutral, 606
time-varying discounting and, 568–569
Invisible hand analogy, forces that drive market to equilibrium, 49
irr (Internal rate of return)
approach to investments, 566–567
example (Golden State Warriors), 567
IRS (increasing returns to scale)
learning curve and, 235
overview of, 191
Isocost line
inputs requiring same total cost, 219–221
minimizing costs and, 222
Isoquants
combining with isocosts in decision making, 221
example (semiconductor integrated circuit), 187
overview of, 184–187
properties of, 184–185
slope or shape of, 185–186, 716
substitutability of inputs varies along, 189
substituting inputs and, 188

L

Labor
average product of, 179
changes that save, 196–197
competitive factor and output markets, 539–540
competitive factor market and monopolized output markets, 540–541
demand (paper firm example), 535
demand varying by market structure, 539
diminishing marginal returns to, 212
firm and market demand for, 536
hiring practices. *See* Hiring practices
as input, 177
inputs combined for single level of output (isoquants), 187
inputs of short-run production, 177
inputs requiring same total cost (isocost line), 219–221
marginal product of, 178–179
monopolized factor market and competitive output markets, 541–542
monopoly in both input and output markets, 542–544
monopsonies and, 547–548
production relationships with variable labor, 179–180
productivity of labor during recession, 172, 198–199
profit maximization, 532–533
profit maximization in monopsonies, 544–545
relationship between labor market and output market, 531
short-run factor demand, 529–530
taxes on impacting output from, 534
technical progress increasing output of, 196–197
wages impacting demand, 533–534, 730–731
Labor-leisure choice
calculations in, 715–716
example (scholarship), 161–162
example (winning the lottery), 163–164
income and substitution effects applied to, 160–162
overview of, 158–160
Labor supply curves
consumer theory, 158
income and substitution effects applied to labor-leisure choice, 160–162
income tax rates and, 164–166
labor-leisure choice, 158–160
shape of, 162–164
Last-dollar rule, in minimizing costs, 223–224
Law of Demand, 710
elasticity of demand and, n. 70
Giffen goods and, 150
overview of, 36
Law of diminishing marginal returns, 181–183
Laws
baseball exemption from antitrust laws, 453
cartels prevented by, 453–455
court action against cartels, 456
Layoffs, vs. pay cuts, 698–700
Learning curve
example (Intel), 236
increasing returns to scale (IRS), 235
overview of, 235–236
Leisure
demand for, 159
income and substitution effects applied to labor-leisure choice, 160–162
income elasticity of, 162–163
labor-leisure choice, 158–160
Lemons
adverse selection and, 651–652
market for, 653
Lending, discount rate and, 555
Lerner Index
of Cournot oligopoly, 464
example (Surface Pro 28 laptop and tablet), 381
market power and shape of demand curve, 379
relationship of price to marginal cost in, 381

Licenses, as barrier to entry, 387–388
Licensing fee
for reducing moral hazard, 684–685
taxes impacting short-run cost curves, 216–217
Limit pricing, in dynamic entry games, 513
Limited liability
in corporations, 174
example (British Petroleum), 585
Limited liability company (LLC), n. 174
Loans, capital markets and interest rates and, 576–577
Long-run average cost curves (*LRAC*), 231–234
Long-run competition
competitive equilibrium, 279–280
entry and exit of firms, 269–270
market supply curve, 269–270
market supply when entry is limited, 272
market supply when firms differ, 272
market supply when input prices vary with output, 273–275
market supply with identical firms and free entry, 271–272
market supply with a large buyer, 275–277
profit maximization, 267
supply curve for, 268
Long-run cost curves
estimating cost curves vs. introspection, 230–231
expansion path and, 226
long-run average cost as envelope of short-run average cost curves, 231–234
overview of, 227–230
Long-run cost function, 225–227
Long-run costs
avoidable nature of, 218
combining cost and production information, 221
estimating cost curves vs. introspection, 230–231
factor price changes and, 224–225
isocost line, 219–221
last-dollar rule, 223–224
learning curve and, 235–236
long-run average cost as envelope of short-run average cost curves, 231–234
long-run cost curves, 227–230
long-run expansion path and long-run cost function, 225–227
lower costs in long run, 231
lowest-isocost rule, 222
minimizing, 218–219
overview of, 218
short- and long-run expansion paths, 234–235
tangency rule, 222–223
Long-run demand
demand elasticity over time, 75–76
factor demand in competitive factor market, 534–535
Long-run expansion path, 225–227
Long-run planning, 231
Long-run production
defined, 177
isoquants, 184–188
overview of, 183–184
substituting inputs and, 188–190
Lowest-isocost rule, minimizing costs and, 222
LRAC (Long-run average cost curves), 231–232
Lump-sum subsidies, vs. per-hour subsidies, 134, 166–167

M

Magnuson-Moss Warranty Improvement Act, 433–434
Malthus, Thomas, 182–183
Management
of firms, 175
good bosses raising productivity, 198
innovations improving output, 196
Marginal benefit (*MB*), of advertising, 439
Marginal costs (*MC*)
comparing competitive market with cartel, 452
deciding whether to advertise, 437–438
determining profit-maximizing output in monopolies, 374–375
in duopoly Nash-Cournot Equilibrium, 466
impact of specific tax on short-run equilibrium, 266–267
long-run cost curves and, 227–228
market failure and, 302
market failure when price exceeds marginal cost, 382
market power and shape of demand curve, 379–380
in natural monopoly, 387
output decisions in short-run competitive firms, 256–257
output rules in profit maximization, 253–254
overview of, 209–210
perfect price discrimination and, 413–416
relationship to price in Lerner Index, 381
shape of marginal cost curve, 212
shape of marginal cost curve depends on shape of long-run cost curve, 227
shift in marginal cost curve effecting monopolies, 377–378
short-run cost curves, 210–212
taxes impacting short-run cost curves, 215–217
Marginal expenditure (*ME*), in monopsonies, 546
Marginal firms, entry and exit of firms and, 269–270
Marginal product of labor (MP_L)
graphing product curves, 179–180

law of diminishing marginal returns, 181–183
long-run factor demand, 531
overview of, 178–179
properties of marginal product curves, 716
relationship of product curves, 181
short-run factor demand, 531–532
substitutability of inputs and, 189–190
in total product, 177-178
Marginal profit, output decisions and, 253–254
Marginal rate of substitution (*MRS*)
efficiency of competition and, 346–347, 351–353
efficient product mixes and, 351
example, 106
expressing in terms of marginal utility, 111–112
labor-leisure choice and, 715
mutually beneficial trades and, 342–344
perfect complements, 106
perfect substitutes, 105–106
willingness to substitute, 104–105
Marginal rate of technical substitution (*MRTS*)
Cobb-Douglas production function, 190–192
last-dollar rule and, 223–224
substitutability of inputs and, 189–190
tangency rule and, 222–223
Marginal rate of transformation (*MRT*)
changes in budget constraints and, 115
comparative advantage, 347–349
efficiency of competition and, 351–353
efficient product mixes and, 351
labor-leisure choice and, 715
number of producers and, 350–351
trade-off between goods, 114
Marginal revenue (*MR*)
comparing competitive firm with monopoly, 369
comparing competitive market with cartel, 451–452
deciding whether to advertise, 437–438
determining profit-maximizing output in monopolies, 374–375
marginal revenue curves in monopolies, 370–372
market power and shape of demand curve, 379–380
in monopolies, 369
output decisions in short-run competitive firms, 256–257
output rules in profit maximization, 254
perfect price discrimination and, 413–416
price elasticity of demand in monopolies and, 372–373
price in monopolies and, 369–370
relationship of linear demand curve to marginal revenue curve, 722
Marginal revenue product of labor (MRP_L)
factor market demand, 535–536
long-run factor demand, 534–535
market structure and factor demand, 538–539
short-run factor demand, 529–532
Marginal tax rate
impact on labor-leisure choice, 164–165
relationship of tax revenue to, 165–166
Marginal utility
expressing marginal rate of substitution as, 111–112
of good, 110–111
labor-leisure choice and, 715
mutually beneficial trades and, 342–344
Marginal willingness to pay
consumer surplus (*CS*) and, 292–294
example (eBay), 294–295
overview of, 292
Market basket. *See* Bundle of goods
Market consumer surplus, 294
Market equilibrium
with asymmetric information, 654–655
for competitive factor market, 537–538
effects of government interventions, 52
forces that drive market to equilibrium, 49–50
graphical and mathematic determination of, 47–48
overview of, 47
shocks to, 50–52
with symmetric information, 653–654
Market failure
deadweight loss reflecting, 302
due to monopoly pricing, 382–385
when price exceeds marginal cost, 382–383
Market makers, impact on market equilibrium, 49
Market power
example (mergers to monopolize), 458
example (San Francisco cable cars), 380
how firms price discriminate, 413–414
impact on price of exhaustible resources, 575–576
in input and output markets, 539
nonuniform pricing, 406
overview of, 378–379
from price ignorance, 661–662
shape of demand curve and, 379–380
sources of, 381–382
which firms can price discriminate, 410–411

Market structures (*continued*)
  Bertrand oligopoly. *See* Bertrand oligopoly
  cartels. *See* Cartels
  comparing, 450–451
  competitive. *See* Competitive markets
  Cournot oligopoly. *See* Cournot oligopoly
  externalities and, 629
  factor demand and, 538–539
  labor demand varying by, 539
  monopolies. *See* Monopolies
  monopolistic competition. *See* Monopolistic competition
  monopsonies. *See* Monopsonies
  Oligopolies. *See* Oligopolies
  Stackelberg oligopoly. *See* Stackelberg oligopoly
Market supply, long-run
  example (vegetable oil), 271
  in increasing-cost markets, 274
  long-run supply curve, 269–270
  supply curve, 268
  supply curve (cotton example), 273
  supply when entry is limited, 272
  supply when firms differ, 272
  supply when input prices vary with output, 273–275
  supply with identical firms and free entry, 271–272
  supply with a large buyer, 275–277
Market supply, short-run
  with firms that differ, 264–265
  with identical firms, 263–264
  impact of specific tax on short-run equilibrium, 266–267
Markets
  as exchange mechanism between buyers and sellers, 26
  market structures. *See* Market structures
  for pollution externalities, 634
  for positive externalities, 635
  price taking in, 246–247
  taxing externalities in noncompetitive markets, 631
Mathematical approach
  applying Nash-Cournot Equilibrium to airlines industry, 462–463
  determining market equilibrium, 47–48
  to profit maximization in monopoly, 375–377
MB (marginal benefit), of advertising, 439
MC. *See* Marginal costs (*MC*)
ME (marginal expenditure), in monopsonies, 546
Medical economies of scope
  cost of producing multiple goods, 236
  examples, 238
Mergers
  antitrust laws and, 457–458
  example (mergers to monopolize), 458
Microeconomics, introduction
  decision making, 26
  models, 27
  overview of, 25
  positive vs. normative statements, 29–31
  price determining allocations, 26
  simplification by assumption in microeconomic models, 28
  testing theories, 28–29
  trade-offs in resource allocation, 26
  uses of models, 31
Mill, John Stuart, 108, 360
Minimizing long-run costs, 218–219
Minimum efficient scale
  for monopolistic competition, 480
  overview of, 479
Minimum wages
  with incomplete coverage, 337–339
  price floors, 58
Mixed strategies, in static games, 501–502
Mobile Number Portability (MNP), 465
Models
  income threshold model (China), 27–28
  overview of, 27
  positive vs. normative statements, 29–31
  simplification by assumption, 28
  testing theories, 28–29
  uses of, 31
Money
  comparing money today with money in future, 554–555
  future value of, 557–558
  present value of, 558
Monitoring, to reduce moral hazard
  bonding to prevent shirking, 693
  checks on principals in contracts, 697–698
  deferred payments, 695
  efficiency wages, 696–697
  outcome monitoring, 697
  overview of, 691–692
Monopolies
  advertising and, 407, 437
  choosing price vs. quantity, 373
  comparing market structures, 450
  cost-based, 385–387
  effects of shift of demand curve, 377–378
  example (Botox patent), 388–390
  example (brand name vs. generic drugs), 367–368, 400–401
  example (iPad), 376
  example (3D printer), 395–396
  externalities and, 629–630
  government creation of, 387–390
  government regulation of, 390–395
  government regulation of (natural gas example), 394–395
government role in encouraging competition, 395–397

graphical approach to profit maximization, 373–375
impact of specific tax on, 384–385, 722–723
Lerner Index and, 381
marginal revenue, 369–373
market failure due to monopoly pricing, 382–385
market power and, 378–379
market power and shape of demand curve, 379–380
mathematical approach to profit maximization, 375–377
network externalities and, 397–398
nonuniform pricing. *See* Nonuniform pricing
overview of, 368
profit maximization in, 368
sources of market power, 381–382
two-period monopoly model, 399–400
vs. competitive welfare with externalities, 630

Monopolies, in factor markets
competitive factor and output markets, 539–540
competitive factor market and monopolized output markets, 540–541
market power in input and output markets, 539
market structure and factor demands, 538–539
monopolized factor market and competitive output markets, 541–542
monopoly in both input and output markets, 542–544
overview of, 538

Monopolistic competition
comparing market structures, 450
defined, 449
equilibrium effects of, 479–480
example (airlines), 481
example (food truck), 478–479
example (zoning laws as barrier to entry), 482–483
fixed costs and number of firms, 481–482
overview of, 449, 478

Monopsonies
defined, 529
overview of, 544
profit maximization in, 544–546, 731–732
welfare effects of, 546–547

Monopsony power
overview of, 545
Monopsony and the gender wage gap example, 546

Monotonicity property, consumer preferences, n. 100

Moral hazard
asymmetric information and, 681–682
contracts for avoiding. *See* Contracts, for avoiding moral hazard
example (clawing back bonuses), 677–678, 701–702
in financial crisis 2007–2009, 677–678
monitoring for avoiding. *See* Monitoring, to reduce moral hazard
opportunistic behavior, 650
overview of, 678–679
principal-agent problem. *See* Principal-agent problem

More-is-better property, consumer preferences, 100–101

Morgenstern, Oskar, 591

Movement along demand curve
caused by shift in supply curve, 68
effect of factors other than price on demand, 37–38
labor example, 533

Movement along indifference curve, 148

Movement along supply curve
price and, 43–44
shocks to equilibrium, 50–51

MP_L. *See* Marginal product of labor (MP_L)

MR. *See* Marginal revenue (*MR*)

MRP_L. *See* Marginal revenue product of labor (MRP_L)

MRS. *See* Marginal rate of substitution (*MRS*)

MRT. *See* Marginal rate of transformation (*MRT*)

MRTS. *See* Marginal rate of technical substitution (*MRTS*)

Multiple regression, 709

Mutual funds, diversification via, 599–600

Mutually beneficial trades, between two people, 342–344

N

Nash-Bertrand equilibrium
best-response curves, 474–475
calculating, 729–730
compared with Nash-Cournot equilibrium, 475
with differentiated products, 476–477
example (cola market), 476
overview of, 474

Nash-Cournot equilibrium
applied to duopoly, 459–460
compared with Nash-Bertrand equilibrium, 475
comparing two firms with different marginal costs, 728
comparing types of equilibria, 472–473
general model, 726–727
graphical approach in applying to airlines industry, 460–462
linear example, 727
marginal costs in Duopoly Nash-Cournot Equilibrium, 466
mathematical approach in applying to airlines industry, 462–463
varying with number of firms in Cournot oligopoly, 464

Nash equilibrium
applying to cooperation in static games, 505–506

Nash equilibrium
applying to dynamic entry game, 510–512
equilibrium in oligopolies and, 458–459
example (explanation of parent/child interactions), 503
mixed strategies in static games, 501–502
multiple Nash equilibria applied to static games, 500
strategies in static games, 495–496
subgames perfect Nash equilibrium, 508–509
Nash, John, 459, 495
Nash-Stackelberg equilibrium
algebraic solution (airline example), 471–472
calculating, 728–729
comparing with Cournot and collusive equilibria, 472–473
graphical model (airline example), 470–472
why sequential movement is essential, 472
Natural disasters
anti-price gouging laws, 331, 362–363
impact on equilibrium price, 332
insurance and, 604–605
Natural monopolies, 386–387
Negative externalities. *See also* Externalities
defined, 620
Spam example, 621
Net present value (*NPV*)
approach to investments, 565–566
example (Golden State Warriors), 566
internal rate of return and, 566–567
Network externalities
behavioral economics and, 397–398
eBay example, 399
as explanation for monopolies, 398–399
overview of, 397
Nominal prices, inflation and, 562–563
Nonlinear price discrimination
example of uniform pricing (iTunes), 432–433
overview of, 427–428
two-part pricing, 429
two-part pricing with identical consumers, 429–430
two-part pricing with nonidentical consumers, 430–432
types of price discrimination, 412
Nonprofit firms, 173
Nonsatiation property, consumer preferences, n. 100
Nonuniform pricing
market power and, 406
price discrimination. *See* Price discrimination
tie-in sales, 433–436
two-part pricing, 429–433
Normal-form, representation of static games, 493–494
Normal goods
fast-food example, 145–146
income effects, 146–149
income elasticities of, 143
some goods must be normal, 144
substitution effects, 146–149
Normative statements, in microeconomics, 29–31
NPV. *See* Net present value (*NPV*)

O

Observable characteristics, identifying groups by, 424–425
Oligopolies
applying static game to strategic advertising by oligopolies, 504
Bertrand oligopoly. *See* Bertrand oligopoly
cartels. *See* Cartels
comparing market structures, 450–451
Cournot oligopoly. *See* Cournot oligopoly
defined, 448
overview of, 448–449
Stackelberg oligopoly. *See* Stackelberg oligopoly
OPEC (Organization of Petroleum Exporting Countries)
as example of cartel, 451
market power and price of exhaustible resources and, 575–576
Open-access common property, 635–636
Opportunistic behavior
adverse selection. *See* Adverse selection
asymmetric information and, 650
moral hazards and, 650–651
restricting, 660
Opportunity costs
of durable goods, 206–207
example problem, 206
of MBA degree, 205–206
overview of, 204–205
Opportunity sets
bundle of goods and, 113
impact of income change on, 116
impact of price change on, 115–116
impact of water quotas on, 114–115
Optimal bundle
on convex sections of indifference curves, 121–122
corner solution, 119–120
Giffen goods and, 150–151
interior solution, 118
more-is-better and, 122
overview of, 117–118
Options (stock), contracts for avoiding moral hazard, 689–690
Ordinal preferences, utility and, 108–109
Organization of Petroleum Exporting Countries (OPEC)
as example of cartel, 451

market power and price of exhaustible resources and, 575–576
Organizational change
productivity and, 172–173
Outputs
competitive factor and output markets, 539–540
competitive factor market and monopolized output markets, 540–541
converting inputs into, 173, 176
cost of producing multiple goods, 236
decisions in short-run competition, 256–257
determining profit-maximizing output in monopolies, 375–377
effect of adding extra labor, 180–181
graphing product curves, 179–180
isoquants for combining inputs for single level of output, 187
law of diminishing marginal returns, 181–183
learning curve and cumulative output, 235
long-run market supply when input prices vary with output, 273–275
market power in output markets, 539
monopolized factor market and competitive output markets, 541–542
monopoly in both input and output markets, 542–544
production function and, 176
profit maximization and, 253–254
profit maximization for labor, 531
reducing below competitive level lowers welfare, 301–304
relationship between labor market and output market, 531
relationship of product curves, 181
returns to scale varying by output levels, 194–195
of short-run production, 177–178
technical progress increasing levels of, 196–197
technological and economic efficiency and, 203
Overconfidence, biases in assessing risk probability, 608
Owners
management of firms and, 175
what owners want, 175
Ownership, of for-profit firms, 174

P

Package tie-in sales. *See* Bundling (package tie-in sales)
Pareto-efficient allocations
competitive exchange and, 346–347
contract curves and, 343
efficiency of competition and, 352
government policies and, 356
obtaining any efficient allocation using competition, 347
Pareto principle
government policies and, 356
in study of welfare, 332
Pareto, Vilfredo, 332
Partial-equilibrium analysis
applied to effects of minimum wage, 58–60
applied to feedback between competitive markets, 334
overview of, 333
Partnerships, ownership of for-profit firms, 174
Patents
brand name vs. generic drugs, 367–368
example (Botox), 388–390
government role in creating monopolies, 387
as monopoly, 368
Pay cuts, vs. layoffs, 698–700
Payoff function
cooperation and, 497
in games, 491
Payoff (or profit) matrix, in static games, 493
Payoffs, in games, 491
Per-hour subsidies, for childcare, 134, 166–167
Perfect competition
characteristics of, 247
comparing market structures, 450–451
deriving demand curve for competitive firm, 249–250
deviations from, 248–249
overview of, 246
price taking and, 246–247
reasons for studying, 251
Perfect complements, marginal rate of substitution and, 106
Perfect monitoring, 681
Perfect price discrimination
calculating, 723–724
efficiency of, 414–416
example (Botox), 416–417
overview of, 413–414
transaction costs and, 418
types of price discrimination, 412
Perfect substitutes
Coke/Pepsi choice, 139, 141–142
marginal rate of substitution and, 105–106
Perfectly elastic demand, 72
Perfectly elastic supply, 78
Perfectly inelastic demand, 71–72
Perfectly inelastic supply, 78
Physical products. *See* Products
Piece-rate contracts, 690
Piracy, of software, 637–638
Policies. *See also* Government policies and regulation
creating barriers to entry or exit, 306–307
entry barriers, 306–307
free trade vs. ban on imports, 318–320
free trade vs. quotas, 323
free trade vs. tariffs, 320–323

Policies (*continued*)
regulating imports, 318
rent seeking and, 323–324
that create wedge between supply and demand, 307
that shift supply and demand curves, 305–306
welfare effects of price ceilings, 316–318
welfare effects of price floors, 312–314
welfare effects of sales taxes, 307–309
welfare effects of subsidies, 310–312
Pollution
Coase Theorem, 632–633
controlling by allocating property rights, 631–632
emissions fees, 626–627
emissions standards, 626
markets for pollution externalities, 634
regulating externalities, 624–626
regulating externalities (pulp and paper mill example), 626
taxes in controlling, 627
trade and, 643–644
welfare effects in competitive market, 732–734
Pooling equilibrium, 667–668
Positive externalities. *See also* Externalities
defined, 621
markets for, 635
Positive monotonic transformation, n. 109
Positive statements, in microeconomics, 29–31
PPF. *See* Production possibility frontier (*PPF*)
Predictions
of game outcome, 493–494
positive statements, 30
Preference maps
for bundle of goods, 101
indifference curves and, 101
properties, 101
Preferences, consumer
curvature of indifference curves, 105–106
defined, 98
indifference curves, 101–104
indifference curves (food/clothing example), 107
ordinal preferences, 108–109
overview of, 99
preference maps, 101
properties, 99–101
willingness to substitute, 104–105
Preferences, social, 358–359
Present biases
food stamp example, 569–570
time-varying discounting and, 568–569
Present value (*PV*)
calculating real present value of payment, 563–564
interest rates connecting present and future value, 557–558
of stream of payments, 558–559, 732
Present value of payments over time, 567–568, 732
Price ceilings
anti-price gouging laws, 331, 362–363
government regulation of monopolies, 390–394
nonoptimal price regulation, 392–394
social costs of, 317–318
supply and demand and, 55–58
welfare effects of, 316–317
Price-consumption curves
demand curve corresponding to, 138
example (tobacco use), 138
taste (indifference curves) impacting, 137–138
Price discrimination
due to false beliefs about quality, 660–661
example (Disneyland), 410
group. *See* Group price discrimination
nonlinear. *See* Nonlinear price discrimination
nonuniform pricing, 406–407
not all price differences are price discrimination, 412
perfect. *See* Perfect price discrimination
preventing resale, 411–412
types of, 412
welfare effects of, 426
which firms can price discriminate, 410–411
why it pays, 408–409
Price elasticity of demand
downward-sloping linear demand curve, 71–73
example (oil drilling in Arctic National Wildlife Refuge), 80–81
example (price of corn), 70–71
group price discrimination and, 421
horizontal demand curve, 73
Lerner Index and, 381
monopolies and, 372–373
over time, 75–76
overview of, 69–71
residual demand curves, 720
revenue and, 74–75
Slutsky equation applied to, 714–715
surplus and, 721–722
total, average, and marginal revenue and, 371
vertical demand curve, 73
Price elasticity of supply
along supply curve, 79
example (oil drilling in Arctic National Wildlife Refuge), 80–81
over time, 79
overview of, 78–79
Price floors
agricultural, 312–314
supply and demand and, 58–59
welfare effects of, 312
Price supports, redistribution of wealth and, 353
Price taking, in competitive markets, 246–247
Prices/pricing
advertising and, 664
allocations determined by, 26

anti-price gouging laws, 331, 362–363
block pricing, 427–428, 724–725
bundling tie-in sales, 434–436
choosing price vs. quantity in monopolies, 373
compensating variation and equivalent variation and, 151–152
effect of price change on consumer surplus, 295–296
effect of price change on demand, 146, 169
effect of price change on factor demand, 533
effect of price change on opportunity set, 115–116
effect of price change on revenue, 74
efficiency of perfect price discrimination, 414–416
example (buying discounts), 425–426
Giffen goods and, 150–151
group price discrimination, 418–419, 724
group price discrimination (textbook resale example), 422
group price discrimination with two groups, 418–419
identifying groups for price discrimination, 424–425
impact on quantity demanded, 36–37
impact on quantity supplied, 42–43
long-run market supply when input prices vary with output, 273–275
marginal revenue and price in monopolies, 369–370
market failure due to monopoly pricing, 382–385
market failure when marginal costs are not equal to price, 302
market failure when price exceeds marginal cost, 382
market power from price ignorance, 661–662
nominal vs. real prices, 152–153
nonlinear price discrimination, 427–428
not all price differences are price discrimination, 412
overview of, 406–407
perfect price discrimination, 413–416, 723–724
preventing resale, 411–412
price discrimination, 412
price discrimination (Botox example), 416–417
price taking in competitive markets, 246–247
quantity vs. price in monopolies, 725–726
relationship to marginal cost in Lerner Index, 381
requirement tie-in sales, 433
sale prices, 440–441
sales prices, 406
sensitivity of quantity demanded to, 69
sensitivity of quantity supplied to, 78
short-run firm supply curve and, 262–263
tie-in sales, 433
transaction costs and perfect price discrimination, 418
two-part pricing, 429
two-part pricing with identical consumers, 429–430
two-part pricing with nonidentical consumers, 430–432, 725–726
uniform pricing (iTunes example), 432–433
welfare effects of price discrimination, 426
which firms can price discriminate, 410–411
why price discrimination pays, 408–409

Prices/pricing, of exhaustible resources
constant or falling prices, 574–576
overview of, 571–573
scarce exhaustible resources (redwood tree example), 574

Principal-agent problem
asymmetric information, 679–680
efficiency of contracts, 680
example (honest cabbies?), 683, 684–685
example (subprime loans), 682–683
overview of, 679–680
symmetric information, 680–681

Principals, checks on principals in contracts, 697–698

Prisoner's dilemma game
cooperation in repeated prisoner's dilemma game, 505
dominant strategies in, 494–495

Private costs, cost of production without externalities, 621

Private firms
overview of, 173–174
ownership of, 174

Private goods, rival goods compared with, 635

Private value auctions, 516

Probability of risk
biases in assessing, 608
frequency of events and, 587
overview of, 587
probability distribution, 588
subjective probability, 587–588

Producer surplus (PS)
applying, 299
block pricing and, 419–420
competition maximizing welfare, 300–301
effects of pollution, 622–623
effects of tariff or quotas, 320–323
entry barrier, 306–307
example (rose production), 300
group price discrimination and, 426
loss from eliminating free trade, 318–320

Producer surplus (PS) (*continued*)
measuring with supply curve, 298–299
monopsonies and, 547
nonoptimal price regulation, 392
optimal price regulation, 390–391
overview of, 297–298
perfect price discrimination and, 414–416
reducing output below competitive level lowers welfare, 301–304
rent seeking and, 323–324
welfare effects of price ceilings, 316
welfare effects of price floors, 312
welfare effects of sales taxes, 307–309
welfare effects of subsidies, 310
Producer welfare. *See* Producer surplus (*PS*)
Producers, number of producers and comparative advantage, 350–351
Product curves, graphing, 179–181
Production
average product of labor, 179
combining cost and production information, 221
constant, increasing, and decreasing returns to scale (*CRS*), 191–192
graphing product curves, 179–181
isoquants, 184–187
law of diminishing marginal returns, 181–183
long-run, 183–184
marginal product of labor, 178–179
overview of, 172–173
production function, 177
quantity vs. price in monopolies, 725–726
returns to scale by industry, 192–194
returns to scale varying by output levels, 194–195
short-run, 177
substituting inputs and, 188–190
total product in short-run production, 177–178
Production and trading
comparative advantage, 347–351
competition and, 351–353
efficient product mixes, 351
overview of, 347
Production functions
Cobb-Douglas production function. *See* Cobb-Douglas production function
determining shape of cost curves, 212
example (beer manufacturer), 717–718
fixed-proportions production function, 185–186
relationship of inputs to outputs, 176–177
returns to scale and, 194
Production possibility frontier (*PPF*)
comparative advantage and, 347–351
cost of producing multiple goods, 236–237
efficiency of competition and, 352
efficient product mixes, 351
example, 349–350
number of producers and, 350–351
optimal product mix, 350
Productivity
innovations and technical progress and, 196–197
of labor during recession, 172, 198–199
organizational change and, 197–198
relative productivity, 195–196
Products. *See also* Outputs
adverse selection with products of unknown quality, 652–653
Bertrand oligopoly with differentiated products, 475–477
Bertrand oligopoly with identical products, 474
efficient product mixes, 351
example of spurious product differentiation (bottled water), 477
identical or homogeneous in competitive markets, 247
nonidentical firms with differentiated products in Cournot oligopoly, 467–469
as output, 176
Profit
firms must maximize profit to survive, 291
formula for, 251–252
graphical approach to profit maximization, 373–375
mathematical approach to profit maximization, 375–377
maximizing, 175, 721
maximizing for output, 531–532
maximizing for output (airlines example), 460
maximizing in long-run competition, 267
maximizing in monopolies, 368, 725–726
maximizing in monopsonies, 544–546
maximizing in short-run competition, 251–252
output decision rules, 253–254
shutdown decision rule, 254–255
two decisions for maximizing profit, 253
zero long-run profit with free entry, 288–289
zero long-run profit with limited entry, 289–290
Profit-sharing contracts, 687–688
Properties
consumer preferences, 99–101

preference maps, 101
Property rights, allocating to reduce externalities, 631–632
Prospect theory
comparing with expected utility theory, 611–612
overview of, 611
properties of, 612–613
PS. *See* Producer surplus (*PS*)
Public firms
overview of, 173–174
state-owned enterprises in China, 174
Public goods
example (free riding on measles vaccination), 640
free riding and, 638–639
overview of, 638
reducing free riding, 641
valuing, 641–642
Public utilities
block pricing by, 427–428
as natural monopoly, 386
PV. *See* Present value (*PV*)

Q

Quality goods
price discrimination due to false beliefs about quality, 660–661
shipping fees and, 150
varying quality with asymmetric information, 655–656
warranties as signal of quality, 657
Quantity demanded
cross-price elasticity of demand and, 77
demand curves and, 35–36
effect of prices on, 36–37
income elasticity of demand and, 76–77, 142–143
relationship to income, 140, 141–142
sensitivity to price, 69
Quantity supplied
price and, 42
supply elasticity and, 78–79
Quotas
effects of, 322
free trade vs., 323
impact on supply curves, 46
import policies and, 318

R

Rate of return
on bonds, 567–568
internal rate of return approach to investing, 566–567
Rawls, John, 360
Rawlsian welfare functions, 360
Real prices, inflation and, 562–563
Rebates, group price discrimination and, 426
Reformulated Fuels Program (RFG), 277
Regression
confidence in estimates, 708–709
estimating economic relations, 707–708
multiple regression, 709
overview of, 707
Regulations. *See* Government policies and regulation
Relative productivity, 195–196
Rent-seeking policies, 323–324
Rents
celebrities collecting, 290
example of zero long-run profit with limited entry, 289–290
on scarce resources, 573
Repeated games
cooperation example (airlines), 506
cooperation in repeated prisoner's dilemma game, 505–506
overview of, 504
strategies in dynamic games, 504
Reputation
of employees, 698
screening by, 657
Requirement tie-in sales, 433–434
Resale prevention
example (designer bags), 411–412
example (textbooks), 422
price discrimination and, 422–424
Reservation prices
of lemon owners, 652–653
in perfect price discrimination, 413
Reserves, of exhaustible resources, 575
Residual demand curves
deriving for competitive firm, 249–250
elasticity of, 720
Residual supply curves
example (cotton in Japan), 275–276
international trade and, n. 275
long-run market supply curve with a large buyer, 275–277
Resources
allocating scarce. *See* Allocation of scarce resources
exhaustible. *See* Exhaustible resources
failure to diversify, 600–601
opportunity cost of using, 204
Returns to scale
constant, increasing, and decreasing, 191–194
by industry, 192–194
long-run costs and, 229
varying by output levels, 194–195
Returns to specialization, 194
Revenue (*R*)
comparing competitive firm with monopoly, 369
demand elasticity and, 74–75
demand elasticity over time and, 75–76
marginal. *See* Marginal revenue (*MR*)
marginal revenue product of labor. *See* Marginal revenue product of labor (MRP_L)
profit and, 175
Reverse auctions, 425–427
RFG (Reformulated Fuels Program), 277–278
Risk assessment
biases in, 608

Risk assessment (*continued*)
expected value, 588–589
overview of, 587
probability of risk, 587–588
prospect theory, 611
variance and standard deviation from expected value, 590–591
violations of expected utility theory, 609–611
Risk attitudes
expected utility theory and, 591–592
gambling and, 596–597
overview of, 591
risk aversion, 592–595
risk neutrality, 595–596
risk preference, 596
Risk aversion
attitudes toward risk, 592–594
investing and, 607
unwillingness to take a fair bet, 592
Risk preferring
attitudes toward risk, 595–596
indifferent regarding fair bet, 592
investing and, 605–606
overview of, 596
willing to take a fair bet, 592
Risk premiums, 594–595
Risk reduction
diversification, 598–600
flight insurance example, 603–604
insurance, 601–603
natural disaster insurance, 604–605
obtaining information, 598
overview of, 597–598
Rival goods, 635
Rivalry/exclusion
club goods, 637
free riding and, 638–639
open-access common property, 635–636
overview of, 635
piracy and, 637–638
public goods, 638–640
valuing public goods, 641–642
Robinson-Patman Act, n. 408
Rules, of games, 491

S

Sales, of exhaustible resources, 570–571
Sales prices, 406
Sales taxes
burden shared by firms and customers, 85–88
effect on supply-and-demand, 82–83
effects of ad valorem and specific taxes, 88–90
equilibrium effects of, 83–85
example (gasoline tax), 91–92
example (prevention of obesity), 87
impact on demand, 35
welfare effects of, 307–309
Salience, bounded rationality and, 127
Satiation
more-is-better and, 122
mutually beneficial trades and, 342–344
Scale, return to. *See* Returns to scale
Scarce resources
allocation of. *See* Allocation of scarce resources
causes of monopolies, 386
pricing, 571–573
Screening
actions for equalizing information, 656–657
bad workers, 700
Sealed-bid auction
overview of, 515
strategies in, 517
Second-degree price discrimination. *See* Nonlinear price discrimination
Second-price strategy
in auctions, 516–517
example (eBay), 294–295
Second Theorem of Welfare Economics, 345, 347
Self-control problems, resulting in falling discount rate, 569–570
Sellers, in competitive markets, 247
Separating equilibrium, education as signal of ability and, 667
Sequential games
credible threat and, 509–510
dynamic entry games, 510–512
dynamic entry games (keeping out casinos example), 513
game trees, 506–507
limit pricing, 513
overview of, 504, 506
subgames perfect Nash equilibrium, 508–509
Sequential movement, in Stackelberg oligopoly, 472
Services
as output, 176
price determining allocations, 26
trade-offs in resource allocation, 26
Shareholders, 174
Shares, corporate, 174
Sherman Antitrust Act in 1890, 453–455
Shift in demand curve
comparing competitive firm with monopoly, 378
effect of factors other than price on demand, 37
effects of rise in income on demand, 140–142
effects on monopolies, 377–378
equilibrium effects of, 50
labor example, 533
policies causing, 304–305
shocks to equilibrium, 50–52
Shift in supply curve
causing movement along demand curve, 68
equilibrium effects of, 50
example, 44–45
policies causing, 304–305
shocks to equilibrium, 50–51
variables causing, 43–44
Shirking
bonding to prevent, 693
contracts for avoiding moral hazard, 700–701
deferred payments to prevent, 695
efficiency wages to prevent, 696–697, 734
monitoring for avoiding, 691–692

Shocks
effects of supply shock depend on shape of demand curve, 68–69
feedback between competitive markets, 334–336
to market equilibrium, 50–51
Short- and long-run expansion paths, 234–235
Short-run average cost curves (*SRAC*)
laser printer example, 233–234
Short-run average cost curves (*SRAC*) (*continued*)
long-run average cost as envelope of short-run average cost curves, 231–232
Short-run competition
competitive equilibrium, 265–267
firm supply curve, 262
market supply curve, 262–265
output decisions, 256–259
overview of, 255–256
shutdown decisions, 259–262
Short-run cost curves
example (beer manufacturer), 215, 717–718
long-run average cost as envelope of short-run average cost curves, 231–234
overview of, 210–212
production functions determining shape of, 212
shape of average cost curve, 214–215
shape of marginal cost curve, 214
shape of variable cost curve, 212–213
Short-run costs
average, 210
fixed, variable, and total, 208
marginal, 209–210
overview of, 208
production functions and shape of cost curves, 212
shape of average cost curve, 214–215
shape of marginal cost curve, 214
shape of variable cost curve, 214–215
short-run cost curve (beer manufacturer example), 215
short-run cost curves, 210–212
taxes impacting, 215–217
Short-run demand
demand elasticity over time, 75
factor demand in competitive factor market, 529–533
Short-run production
average product of labor, 179
defined, 177
graphing product curves, 179–181
law of diminishing marginal returns, 181–183
marginal product of labor, 178–179
overview of, 177
total product, 177–178
Shutdown
decisions in monopolies, 375
decisions in short-run competition, 259–261
example (ethanol processing plants), 269
fracking and shutdowns, 261
rules in profit maximization, 254–255
Signaling
actions for equalizing information, 657
education as signal of ability, 666–670
employees signaling regarding productivity, 700
strategies in dynamic games, 505
Simplification, models and, 28
Single-price monopoly
comparing welfare effects of group price discrimination, 426
group price discrimination compared with, 426
perfect price discrimination, 415–416
Slutsky equation, n. 146, 714–715
Slutsky, Eugene, 714
Snob effect, network externalities and, 398
Social costs
cost of production with externalities, 622
social marginal costs, 623
Social producer surplus, 623
Social welfare functions
efficiency vs. equity, 361
in ranking allocations, 356–357
voting as means of ranking allocations, 360–361
SOEs (state-owned enterprises), in China, 174
Sole proprietorships, ownership of for-profit firms, 174
Specialization, returns to, 194
Specific taxes (unit taxes)
equilibrium effects of, 88–90, 709–710
example (roses), 308
free trade vs. tariffs, 320–323
impact on monopolies, 384, 722–723
impact on short-run equilibrium, 266–267
types of sales taxes, 82–83
Spurious product differentiation
Cournot oligopoly and, 467–468
example (bottled water), 477
SRAC (short-run average cost curves)
laser printer example, 234–235
long-run average cost as envelope of short-run average cost curves, 231–232
Stackelberg game tree, 507
Stackelberg oligopoly
comparing types of equilibria, 472–473
as dynamic entry game, 510–512
graphical model of, 470–471
limit pricing and, 513
overview of, 469
as sequential game (airline example), 506–510
sequential movement in, 472

Standard deviation, from expected value, 590–591
Standards, actions for equalizing information, 658
State-contingent contracts, 685
State-owned enterprises (SOEs), in China, 174
States of nature, decision making and, 586
Static games
  applying to strategic advertising by oligopolies, 497–499
  best response and Nash Equilibrium, 495–496
  cooperation and, 505–506
  dominant strategies, 494–495
  in explanation of parent/child interaction, 503
  mixed strategies, 501–502
  multiple Nash equilibria and, 500–501
  normal-form representation of, 492–493
  overview of, 492–493
  predicting outcomes, 494
Statistical discrimination, in hiring, 670–672
Stocks (investment)
  failure to diversify, 600–601
  investing in corporations, 174
  risk premium of, 594–595
Stocks (of durable goods), 554
Strategies, in game theory
  auction strategies, 516–517
  best response and Nash Equilibrium, 495–496
  defined, 491
  dominant strategies in static games, 494–495
  in dynamic games, 504
  mixed strategies, 501–502
  multiple Nash equilibria and, 500–501
  in normal-form representation of static games, 504
  pure strategy and mixed strategy, 501–502
  in repeated prisoner's dilemma game, 505–506
Stream of payments
  interest rates on, 558–561
  present value of payments over time, 732
Subgames, in sequential games, 507
Subgames perfect Nash equilibrium, 508–509
Subjective probability, 587–588
Subsidies
  effects on supply and demand, 90–92
  example (aircraft construction), 448, 483–484
  example (childcare), 134, 166–167
  example (Ethanol), 90–91
  as negative tax, n. 87
  welfare effects of, 310–312
Substitute goods
  cross-price elasticity of demand and, 77
  demand and price and, 34–35
  example (substituting alcohol for marijuana), 119
  imperfect substitutes, 107
  marginal rate of substitution (*MRS*) and, 105–106
  market power and lack of, 381–382
  perfect complements, 106
  perfect substitutes, 105–106
Substitution bias, in CPI, 157–158
Substitution effects
  example (apple pie), 148–149
  example (plate prices at two types of stores), 149–150
  inferior goods and, 150–151
  labor-leisure choice and, 158–160
  with normal goods, 146–150
  overview of, 146
  Slutsky equation applied to, 714–715
Sunk costs
  as barrier to entry, 306
  overview of, 207–208
Supply
  excess supply, 49
  impact of import policy on supply curves, 45–46
  overview of, 42
  policies creating wedge between supply and demand, 307
  policies that cause demand to differ from supply, 55
  policies that shift supply curves, 52–55
  shock impact depending on shape of demand curve, 68–69
  summing supply curves, 45
  supply curve, 42–44
  supply function, 44
  when to use supply-and-demand model, 60–62
  why supply need not equal demand, 59–60
Supply and demand
  demand curves, 35–36
  demand function, 38–41
  effect of factors other than price on demand, 37–38
  equilibrium effects of government interventions, 52
  example (genetically modified food), 33
  import policy impacting supply curves, 45–47
  market equilibrium, 47–50
  overview of, 34
  policies that cause demand to differ from supply, 55
  policies that shift supply curves, 52–55
  price ceilings, 55–58
  price floors, 58–59
  shocks to equilibrium, 50–52
  summing demand curves, 41–42
  summing supply curves, 45
  supply curves, 42–44
  supply function, 44–45
  why supply need not equal demand, 59–60
Supply-and-demand model
  applying demand and supply elasticity (oil drilling example), 80–81
  cross-price elasticity of demand, 77

demand elasticity along downward-sloping linear curve, 71–73
demand elasticity along horizontal curve, 73
demand elasticity along vertical curve, 73
demand elasticity and revenue, 74–75
demand elasticity over time, 75–76
effects of ad valorem and specific taxes, 88–89
effects of sales taxes, 82–83
effects of subsidies, 90–91
equilibrium effects of specific sales tax, 83–85
example (gasoline tax), 67

Supply-and-demand model (*continued*)
income elasticity of demand, 76–77
overview of, 68
price elasticity of demand, 69–71
price elasticity of supply, 78–79
quantity demanded sensitivity to price, 69
quantity supplied sensitivity to price, 78
sales tax burden shared by firms and customers, 85–88
shapes of supply and demand curves, 68
supply elasticity along supply curve, 78–79
supply elasticity over time, 78–79
use by economists, 34
when to use, 60–62

Supply curves
defined, 42
determining market equilibrium, 47
effects of specific tax on short-run equilibrium, 266–267
example (coffee), 42–43
example (cotton), 273
example (gasoline), 277–278
import policy impacting, 45–47
long-run competitive equilibrium and, 279–280
long-run firm supply curve, 268
long-run market supply curve, 269–270
measuring producer surplus, 298–299
monopsonies and, 544
movement along. *See* Movement along supply curve
overview of, 42–44
shape of, 68
shift in. *See* Shift in supply curve
short-run firm supply curve, 262
short-run market supply curve, 262–265
summing, 45
supply elasticity along, 79

Supply curves, for labor
deriving, 158
income and substitution effects applied to labor-leisure choice, 160–162
income tax rates and, 164–166
labor-leisure choice, 158–160
shape of, 162–164

Supply function, 44–45

Surplus
consumer surplus. *See* Consumer surplus (*CS*)
producer surplus. *See* Producer surplus (*PS*)

Symmetric information, in efficient contracts, 680–681

T

Tangency rule, minimizing costs and, 222

Tariffs (duty)
free trade vs., 320–323
import policies and, 318
resale prevention, 411

Tastes, consumer. *See also* Preferences, consumer, 34

Taxes
bounded rationality and tax salience, 127
burden of sales tax shared by firms and customers, 85–88
effect of sales tax on supply-and-demand, 82–83
effects of ad valorem and specific taxes, 88–90
effects of specific tax on monopolies, 384–385
effects of specific tax on short-run equilibrium, 266–267
effects of tax on output on demand for labor, 533–534
equilibrium effects of specific sales tax, 83–84
example (gasoline tax), 91–92
example (tobacco use), 138
on externalities in noncompetitive markets, 631
impact of sales taxes on demand, 35
incidence of, 85–86
income tax rates and labor supply curve, 164–166
for pollution control, 627
redistribution of wealth and, 353
shifting burden to middle class, 354–355
short-run costs impacted by, 215–217
subsidies as negative tax, n. 87
urban flight due to tax on wages, 340
welfare effects of sales taxes, 307–309
who pays the gasoline tax, 67–68

Technical progress
defined, 196
impact on price of exhaustible resources, 575
productivity and, 196–197

Technological efficiency
choices at home vs. abroad, 238–239
long-run costs and, 218
maximizing profits, 175

Technological efficiency (*continued*)
output decisions, 203
Testing theories, with microeconomic model, 28–29
Tests of transitivity, 125
Thales, 368
Third-degree price discrimination. *See* Group price discrimination
Third-party information, actions for equalizing information, 658–659
Threatening to punish
credibility of in sequential dynamic games, 509–510
strategies in dynamic games, 504–506
Tie-in sales
bundling, 434–436
overview of, 433
requirement tie-in sales, 433–434
Time
choices over, 564–565
investing and. *See* Investments
price elasticity of demand over, 75–76
supply elasticity over, 80
time-varying discounting, 568–569
Total costs (*C*)
long-run cost curves, 227–228
overview of, 208
short-run cost curves, 210–211
Total effects
apple pie example, 148–150
labor-leisure choice and, 158–160
as sum of income and substitution effects, 148
Total product. *See* Outputs
Tourist-trap model
monopoly pricing, 663
overview of, 662
uncompetitive pricing, 662–663
Trade
benefits of, 349
long-run market supply curve with a large buyer, 275–277
pollution and, 643–644
Trade laws, resale prevention, 411–412
Trade-offs
marginal rate of transformation measuring, 114
resource allocation and, 26
Trades, between two people
bargaining ability and, 344
endowment as basis of, 340–342
mutually beneficial trades, 342–344
overview of, 340
Transaction costs
in competitive markets, 247–248
perfect price discrimination and, 418
resale difficult when transaction costs are high, 411
Transitivity property, consumer preferences, 100
Transitivity (rationality) assumption, n. 357
Trigger strategy, in dynamic games, 506
Trusts. *See also* Cartels
Two-part pricing
with identical consumers, 429–430
with nonidentical consumers, 430–431, 725–726
overview of, 429
types of nonuniform pricing, 407
Two-period monopoly model, 399–400

U

Ultimatum games, 519
Uncertainty
assessing risk, 587
behavioral economics of, 608
biases in assessing risk probability, 608
diversification as risk reduction method, 598–600
example (British Petroleum), 585, 613–614
example (flight insurance), 603–604
expected utility theory and, 591–592
expected value in assessing risk, 588–590
information in risk reduction, 598
insurance as risk reduction method, 601–605
investing and, 605–606
natural disaster insurance, 604–605
overview of, 586
probability of risk, 587–588
prospect theory, 611–612
reducing risk, 597–598
risk attitudes, 591
risk-averse investing, 607
risk aversion, 592–594
risk-neutral investing, 606
risk neutrality, 595–596
risk preferring, 596–597
variance and standard deviation from expected value, 590–591
violations of expected utility theory, 609–611
Unified Carrier Registration Agreement, 245
Uniform pricing
example (iTunes), 432–433
overview of, 406
Unions, as monopoly, 542
Unit taxes. *See* Specific taxes (unit taxes)
Unitary elasticity, of supply, 78
Universal coverage, restricting opportunistic behavior, 660
Utility
defined, 98
expected utility, 591–592
indifference curves and, 109–110, 710–712
in labor-leisure choice, 715–716
marginal rate of substitution (*MRS*), 111–112
marginal utility, 110–111

maximizing in mutually beneficial trades, 342–344
maximizing subject to budget constraint, 117–118, 712–714
ordinal preferences, 108–109
overview of, 108
risk aversion and, 592–594
risk neutrality and, 595–596
risk preferring and, 596
social welfare functions maximizing, 360–361
utility function, 108
Utility function, 108–112

V

Variable costs (*VC*)
overview of, 208
shape of variable cost curve, 212–213
short-run cost curves, 210–211
Variable costs (*VC*) (*continued*)
shutdown decisions in short-run competitive firms, 259–261
taxes impacting short-run cost curves, 215–216
Variable inputs, 177
Variance, from expected value, 590–591
VC. *See* Variable costs (*VC*)
Vertical integration, as means of preventing resale, 411
von Neumann, John, 591
Voting, as allocation method, 357–359

W

Wages
compensating differentials for occupational safety, 649–650
efficiency wages to prevent moral hazard, 694–697
factor demand impacted by change in, 531–532, 730–731
income and substitution impacted by change in, 160–162
layoffs vs. pay cuts, 696–698
minimum wages with incomplete coverage, 337–340
occupational licensing and, 52
price floors, 58–59
shape of labor supply curves, 162–163
unions ability to raise, 542
urban flight resulting from tax on, 340
Warranties, as signal of quality, 657
Water quotas, impacting opportunity set, 114–115
Wealth
government programs in redistribution of, 353
inequitable distribution of, 354–355
Welfare
applying producer surplus and, 299
barriers to entry or exit reducing, 306
block pricing and, 427–428
competition maximizing, 300–301
of consumers. *See* Consumer surplus (*CS*)
deadweight loss and, 302
deadweight loss (Christmas presents example), 304–305
differentiated products and, 477
in economists' terminology, 287
effects of monopsonies, 546–548
effects of nonoptimal price regulation, 393
effects of optimal price regulation, 390–391
effects of pollution, 622–623, 732–734
effects of price ceilings, 316–317
effects of price changes, 295–297
effects of price discrimination, 426
effects of price floors, 312
effects of sales taxes, 307–309
effects of subsidies, 310–312
entry barrier, 306–307
government policies and, 356, n. 305
group price discrimination and, 426
maximizing, 357
measuring consumer surplus with demand curve, 292–295
measuring producer surplus with supply curve, 298–299
monopolies vs. competitive welfare with externalities, 630–631
overproduction causing decrease in, 302–303
perfect price discrimination and, 414–416
of producers. *See* Producer surplus (*PS*)
Welfare economics, 287–288
Willingness-to-accept (WTA), deadweight loss of Christmas presents and, n. 304
Willingness-to-pay (WTP), deadweight loss of Christmas presents and, n. 304
Willingness to substitute. *See* Marginal rate of substitution (*MRS*)
Winner's curse, 517–518
World Trade Organization (WTO)
aircraft subsidies and, 448
pressure on governments to reduce/eliminate barriers to trade, 395

Z

Zero long-run profit
for competitive firms with free entry, 288–289
for competitive firms with limited entry, 289–290
Zoning laws, as barrier to entry by hotel chains, 482–483

Credits

Cover photograph: FOTOGRIN/Shutterstock
p. 29, cartoon: Pearson Education, Inc.
p. 33, photograph: AGE Fotostock/SuperStock
p. 35, photograph: Lev/Fotolia
p. 37, cartoon: Zach Weinersmith/smbc-comics
p. 51, cartoon: Pearson Education, Inc.
p. 55, photograph: United States Government Printing Office
p. 57, photograph: Carlos Garcia Rawlins/Reuters
p. 67, photograph: Jeffrey M. Perloff
p. 71, photograph: lurin/123RF
p. 80, photograph: United States Geological Survey (USGS)
p. 87, photograph: Pearson Education, Inc.
p. 97, photograph: Jeffrey M. Perloff
p. 99, cartoon: Pearson Education, Inc.
p. 104, cartoon: Pearson Education, Inc.
p. 119, photograph: Kesu/Shutterstock
p. 124, photograph: United States Department of Agriculture (USDA)
p. 134, photograph: Jim West/Alamy Stock Photo
p. 145, photograph: Elena Shashkina/Shutterstock
p. 158, cartoon: Pearson Education, Inc.
p. 172, photograph: vario images GmbH & Co. KG/Alamy Stock Photo
p. 183, photograph: Micheline Pelletier Decaux/Sygma/Getty Images
p. 183, photograph: Pearson Education, Inc.
p. 191, cartoon: Pearson Education, Inc.
p. 196, cartoon: Pearson Education, Inc.
p. 197, photograph: Wang Jianwei/Xinhua News Agency/Newscom
p. 203, photograph: Christian Delbert/Shutterstock
p. 208, cartoon: Pearson Education, Inc.
p. 209, photograph: Pearson Education, Inc.
p. 245, photograph: Nomad_Soul/Fotolia
p. 261, photograph: jaddingt/Fotolia
p. 287, photograph: Monkey Business/Fotolia
p. 290, photograph: Jeffrey M. Perloff
p. 294, photograph: Jeffrey M. Perloff
p. 304, cartoon: Pearson Education, Inc.
p. 320, photograph: JackF/Fotolia
p. 320, photograph: Jeffrey M. Perloff
p. 331, photograph: National Oceanic and Atmospheric Administration (NOAA)
p. 354, Abraham Lincoln photograph: Abraham Lincoln (5 Feb. 1865), Alexander Gardner. Gelatin silver print. Library of Congress Prints and Photographs Division [LC-USZ62-11896]
p. 354, George Washington photograph: George Washington (Vaughan Portrait) (1795), Gilbert Stuart. Oil on canvas, 73 x 60.5 cm (28 3/4 x 23 13/16 in.). Andrew W. Mellon Collection, National Gallery of Art, Washington, DC
p. 355, Breaker Boys photograph: Breaker boys working in Ewen Breaker of Pennsylvania Coal Co. (1911), Lewis Wickes Hine. Library of Congress Prints and Photographs Division [LC-USZ62-12876]
p. 355, Migrant Mother photograph: Destitute pea pickers in California. Mother of seven children. Age thirty-two. Nipomo, California ("Migrant Mother") (1936), Dorothea Lange. Library of Congress Prints and Photographs Division [LC-USZ62-95653]
p. 355, Ronald Reagan photograph: Library of Congress Prints and Photographs Division [LC-USZ62-13040]
p. 355, Barack Obama photograph: Pete Souza/Library of Congress Prints and Photographs Division [LC-DIG-ppbd-00358]
p. 367, photograph: Jeffrey M. Perloff
p. 380, photograph: Jeffrey M. Perloff
p. 382, cartoon: Pearson Education, Inc.
p. 389, photographs: PHILADELPHIA INQUIRER/KRT/Newscom
p. 406, photograph: Jeffrey M. Perloff
p. 410, photograph: Jeffrey M. Perloff
p. 425, cartoon: Pearson Education, Inc.
p. 434, photograph: Paul Broadbent/Shutterstock
p. 448, photograph: Tom Mareschal/Alamy Stock Photo
p. 454, cartoon: Pearson Education, Inc.
p. 468, photograph: Hayley Chouinard
p. 476, photograph: Lee Bennack
p. 477, photograph: Jeffrey M. Perloff
p. 478, photograph: Jeffrey M. Perloff
p. 490, photograph: Fotos593/Shutterstock
p. 499, photograph: Dorling Kindersley Limited
p. 501, cartoon: Pearson Education, Inc.
p. 510, cartoon: Pearson Education, Inc.
p. 528, photograph: vivalapenler/Fotolia

p. 553, photograph: Jeffrey M. Perloff
p. 574, photograph: Jeffrey M. Perloff
p. 585, photograph: United States Coast Guard
p. 587, photograph: Boeri, SRL
p. 592, photograph: Jeffrey M. Perloff
p. 597, photograph: Jeffrey M. Perloff
p. 604, cartoon: Pearson Education, Inc.
p. 609, photograph: Uryadnikov Sergey/Fotolia
p. 619, photograph: Barry Sweet/AP Images
p. 621, photograph: Feng Yu/Fotolia
p. 624, cartoon: Pearson Education, Inc.
p. 633, photograph: Doral Chenoweth III/The Columbus Dispatch/AP Images
p. 637, photograph: Pearson Education, Inc.
p. 649, photograph: Davis Turner/EPA/Newscom
p. 652, photograph: Pearson Education, Inc.
p. 657, cartoon: Pearson Education, Inc.
p. 677, cartoon: Pearson Education, Inc.
p. 685, photograph: Damian Dovarganes/AP Images
p. 689, photograph: Andy Sheppard/Redferns/Getty Images

Symbols Used in This Book

v = *ad valorem* tax (or tariff) rate, or an exponent in a Cobb-Douglass production function

Δ [capital delta] = change in the following variable (for example, the change in p between Periods 1 and 2 is $\Delta p = p_2 - p_1$, where p_2 is the value of p in Period 2 and p_1 is the value in Period 1)

[epsilon] = the price elasticity of demand

η [eta] = the price elasticity of supply

$\mathscr{L}$ = lump-sum tax

π [pi] = profit = revenue − total cost = $R - C$

σ [sigma] = standard deviation

θ [theta] = probability or share

ξ [xi] = the income elasticity of demand

Abbreviations, Variables, and Function Names

AFC = average fixed cost = fixed cost divided by output = F/q

AVC = average variable cost = variable cost divided by output = VC/q

AC = average cost = total cost divided by output = C/q

AP_Z = average product of input Z (for example, AP_L is the average product of labor)

C = total cost = variable cost + fixed cost = $VC + F$

CRS = constant returns to scale

CS = consumer surplus

CV = compensating variation

D = market demand curve

D_r = *residual* demand curve

DRS = decreasing returns to scale

DWL = deadweight loss

F = fixed cost

i = interest rate

I = indifference curve

IRS = increasing returns to scale

K = capital

L = labor

LR = long run

m = constant marginal cost

M = materials

MC = marginal cost = $\Delta C/\Delta q$

MP_Z = marginal (physical) product of input Z (for example, MP_L is the marginal product of labor)

MR = marginal revenue = $\Delta R/\Delta q$

MRS = marginal rate of substitution

$MRTS$ = marginal rate of technical substitution

MU_Z = marginal utility of good Z

n = number of firms in an industry

p = price

PPF = production possibility frontier

PS = producer surplus = revenues − variable costs = $R - VC$

Q = market (or monopoly) output

$\overline{Q}$ = output quota

q = firm output

R = revenue = pq

r = price of capital services

s = per-unit subsidy

S = market supply curve

S_o = supply curve of all the other firms in the market

SC = a market of economies of scope

SR = short run

t = specific or unit tax (or tariff)

T = tax revenue ($\alpha\, pQ$, τQ, $\rho\pi$)

U = utility

VC = variable cost

w = wage

W = welfare

Y = income or budget

Featured Applications in This Book

Chapter 1:
Twinkie Tax 26
Income Threshold Model and China 27

Chapter 2:
Calorie Counting 38
Aggregating Corn Demand Curves 41
Occupational Licensing 52
Venezuelan Price Ceilings and Shortages 57

Chapter 3:
The Demand Elasticities for Google Play and Apple Apps 71
Amazon Prime 75
Anti-Smoking Policies May Reduce Drunk Driving 77
Oil Drilling in the Arctic National Wildlife Refuge 80
Taxes to Prevent Obesity 87
Subsidizing Ethanol 90

Chapter 4:
You Can't Have Too Much Money 100
Indifference Curves Between Food and Clothing 107
Substituting Alcohol for Marijuana 119
Benefiting from Food Stamps 124
Opt In Versus Opt Out 126
Unaware of Taxes 128

Chapter 5:
Smoking Versus Eating and Phoning 138
Fast-Food Engel Curve 145
What's Your Smart Phone Worth to You? 152
Paying Employees to Relocate 156
Working After Winning the Lottery 163

Chapter 6:
Chinese State-Owned Enterprises 174
Malthus and the Green Revolution 182
A Semiconductor Integrated Circuit Isoquant 187
Returns to Scale in Various Industries 192
Robots and the Food You Eat 197
A Good Boss Raises Productivity 198

Chapter 7:
The Opportunity Cost of an MBA 205
The Sharing Economy and the Short Run 209
A Beer Manufacturer's Short-Run Cost Curves 215
3D Printing 230
A Beer Manufacturer's Long-Run Cost Curves 232
Should You Buy an Inkjet or a Laser Printer? 233
Medical Economies of Scope 238

Chapter 8:
Fracking and Shutdowns 261
The Size of Ethanol Processing Plants 269
Entry and Exit of Solar Power Firms 270
Upward-Sloping Long-Run Supply Curve for Cotton 273
Reformulated Gasoline Supply Curves 277

Chapter 9:
What's a Name Worth? 291
Willingness to Pay on eBay 294
Goods with a Large Consumer Surplus Loss from Price Increases 296
Deadweight Loss of Christmas Presents 304
Welfare Effects of Allowing Fracking 306
The Deadweight Loss from Gas Taxes 309
How Big Are Farm Subsidies and Who Gets Them? 315
The Social Cost of a Natural Gas Price Ceiling 317
Russian Food Ban 320

Chapter 10:
Urban Flight 340
The 1% Grow Wealthier 354
How You Vote Matters 359

Chapter 11:
Apple's iPad 376
Cable Cars and Profit Maximization 380
Botox Patent Monopoly 388
Natural Gas Regulation 394
Movie Studios Attacked by 3D Printers! 395
eBay's Critical Mass 399

Chapter 12:
Disneyland Pricing 410
Preventing Resale of Designer Bags 411
Botox and Price Discrimination 416
Google Uses Bidding for Ads to Price Discriminate 417
Warner Brothers Sets Prices for a *Harry Potter* DVD 419
Reselling Textbooks 422
Buying Discounts 425
iTunes for a Song 432
Ties That Bind 433
Super Bowl Commercials 439

Chapter 13:
The Apple-Google-Intel-Adobe-Intuit-Lucasfilms-Pixar Wage Cartel 455
Cheating on the Maple Syrup Cartel 456
Mergers to Monopolize 458
Mobile Number Portability 465
Bottled Water 477
Monopolistically Competitive Food Truck Market 478
Zoning Laws as a Barrier to Entry by Hotel Chains 482

Chapter 14:
Strategic Advertising 499
Tough Love 503
Keeping Out Casinos 513
Bidder's Curse 518
GM's Ultimatum 519

Chapter 15:
Unions and Profits 542
Monopsony and the Gender Wage Gap 546

Chapter 16:
Power of Compounding 557
Saving for Retirement 561
Winning the Lottery 564
Falling Discount Rates and Self-Control 569
Redwood Trees 574

Chapter 17:
Stocks' Risk Premium 595
Gambling 596
Failure to Diversify 600
Flight Insurance 603
Limited Insurance for Natural Disasters 604
Biased Estimates 609
Loss Aversion Contracts 613

Chapter 18:
Negative Externalities from Spam 621
Global Warming 624
Why Tax Drivers 628
Protecting Babies 628
Buying a Town 633
Acid Rain Cap-and-Trade Program 634
Road Congestion 636
Software Piracy 637
Free Riding on Measles Vaccinations 640
What's Their Beef? 641

Chapter 19:
Discounts for Data 657
Adverse Selection and Remanufactured Goods 659
Reducing Consumers' Information 661
Cheap Talk in eBay's Best Offer Market 666

Chapter 20:
Honest Cabbie? 683
Health Insurance and Moral Hazard 685
Sing for Your Supper 688
Capping Oil and Gas Bankruptcies 694
Layoffs Versus Pay Cuts 698